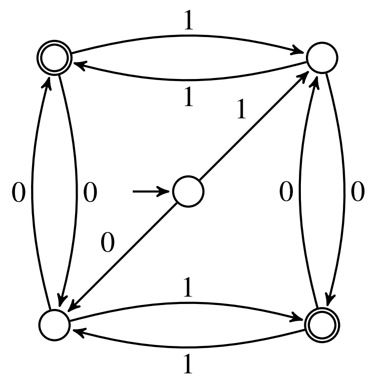
CS 181 EXAM #1 **NAME** ______________________________

SPRING 2015 **UCLA ID** ______________________________

You have 90 minutes to complete this exam. You may assume without proof any statement proved in class.

(3 pts) **1** Give a simple verbal description of the language recognized by the following DFA.

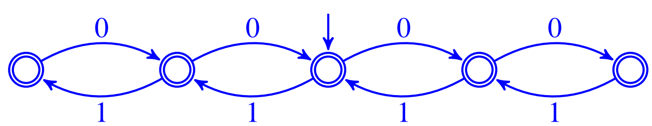


Solution. Nonempty strings of even length.

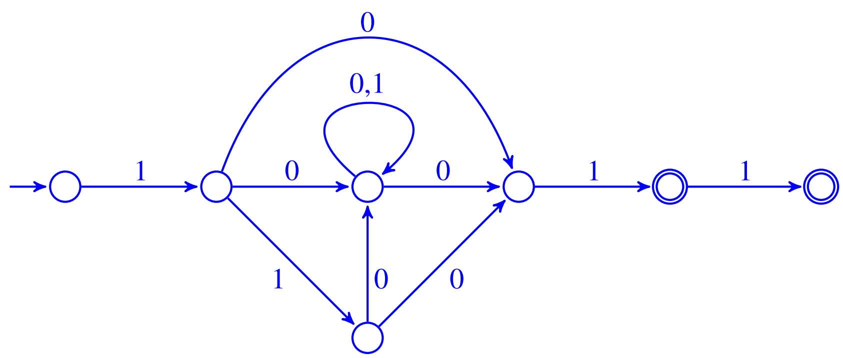
2 Draw NFAs for the following languages over $\{0, 1\}$, taking full advantage of nondeterminism:

(2 pts) **a.** strings such that in every prefix, the numbers of zeroes and ones differ by at most 2;

(2 pts) **b.** strings that begin with 10 or 110, and end with 01 or 011.



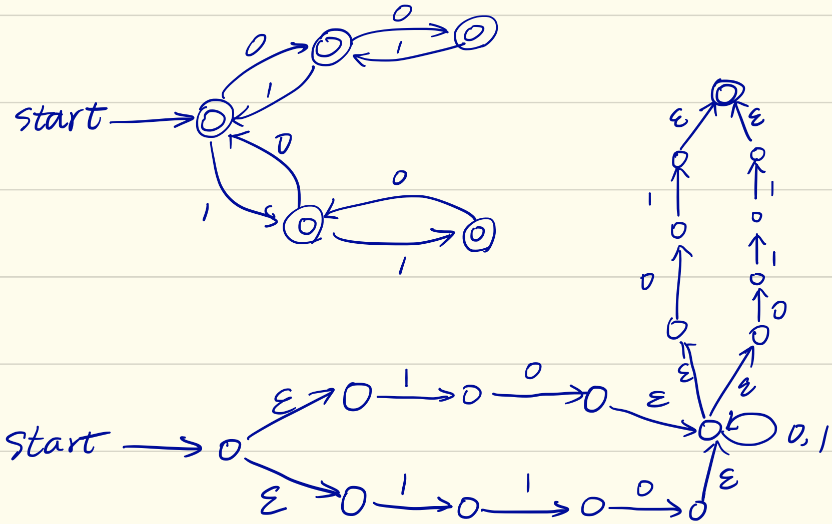
a.



b.

2 Draw NFAs for the following languages over $\{0, 1\}$, taking full advantage of nondeterminism:

- **a.** strings such that in every prefix, the numbers of zeroes and ones differ by at most 2;
- **b.** strings that begin with 10 or 110, and end with 01 or 011.



3 Prove that the following languages over a given alphabet Σ are regular:

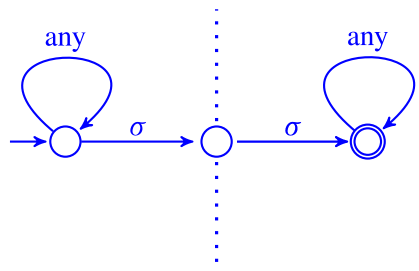
(2 pts) **a.** strings in which no pair of consecutive characters are identical;

(2 pts) **b.** binary strings in which every even-numbered character is a 0;

(2 pts) **c.** the language $\{3, 6, 9, 12, 15, 18, 21, \ldots\}$ over the decimal alphabet, corresponding to natural numbers that are divisible by 3.

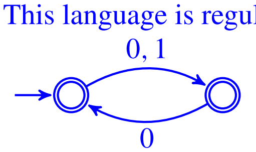
Solution.

a. Let L be the language in the problem statement. Then $\overline{L}$ is the set of all strings that contain a pair of consecutive characters that are identical, corresponding to the following NFA:



This diagram features a branch for each symbol $\sigma \in \Sigma$. Since $\overline{L}$ is regular and regular languages are closed under complement, L must be regular as well.

b. This language is regular because it is recognized by the following NFA:



c. Recall that an integer is divisible by 3 if and only if the sum of its digits is divisible by 3. This suggests the automaton $(\{0,1,2\}, \{0,1,2,\ldots,9\}, \delta, 0, \{0\})$ where $\delta(q,\sigma) = (q+\sigma) \bmod 3$. This automaton *almost* works, except that it accepts syntactically invalid strings such as ϵ or 003. To fix this, modify the automaton to only accept strings that start with a nonzero digit: $(\{\epsilon, 0, 1, 2, \text{FAIL}\}, \{0,1,2,\ldots,9\}, \delta, \epsilon, \{0\})$ where

$$\delta(q,\sigma) = \begin{cases} (q+\sigma) \bmod 3 & \text{if } q = 0,1,2, \\ \sigma \bmod 3 & \text{if } q = \epsilon \text{ and } \sigma \neq 0, \\ \text{FAIL} & \text{otherwise.} \end{cases}$$

(3 pts) **4** The symmetric difference of two languages L' and L'', denoted $L' \Delta L''$, is the set of strings that belong to L' or L'' but not both. Prove that regular languages are closed under symmetric difference.

Solution. Let L' and L'' be regular. By definition, $L' \Delta L'' = (L' \cap \overline{L''}) \cup (\overline{L'} \cap L'')$. Since regular languages are closed under complement, intersection, and union, it follows that $L' \Delta L''$ is regular.

(3 pts) **5** Prove or argue to the contrary: adding a finite number of strings to a regular language necessarily results in a regular language.

Solution. As shown in class, every finite language is regular. Since regular languages are closed under union, it follows that the union of a regular language with a finite language is regular.

(3 pts) **6** The *circular shift* of a language L is defined as $L^{\circlearrowleft} = \{uv : vu \in L\}$, where u and v stand for arbitrary strings. For example, $\{1234\}^{\circlearrowleft} = \{1234, 2341, 3412, 4123\}$. Prove that regular languages are closed under circular shift.

Let $D = (Q, \Sigma, \delta, q_0, F)$ be a DFA for L. Here is a nondeterministic procedure for deciding whether a given string w is in $L^{\circlearrowleft}$:

Stage 1. Choose a state $q \in Q$ nondeterministically.

Stage 2. Launch the DFA on the input string, starting in state q rather than q_0. No need to wait for the DFA to process all of w; whenever the DFA is in an accept state, you may choose to advance to the next stage.

Stage 3. Run the DFA on the unprocessed portion of w, starting in state q_0 and accepting if you end up in state q.

For any fixed $q \in Q$, stages 2 and 3 can be implemented as an NFA D_q which consists of two copies of the original DFA: the first copy has all states marked as rejecting and has ϵ-transitions added from any state in F to the state q_0 of the second copy, and the second copy has q marked as the only accept state. Now for each $q \in Q$, create a copy of that composed automaton D_q. To obtain an automaton for $L^{\circlearrowleft}$, it remains to introduce a new start state that has, for each $q \in Q$, an ϵ-transition to state q of the copy D_q.

(3 pts) **7** Describe an algorithm that takes as input a DFA and determines whether the automaton recognizes the empty language, $\varnothing$.

We can view a DFA as a directed graph, with the DFA's arrows and states corresponding to edges and vertices. The algorithm is simply to check if the graph contains a path from the start state to an accept state. This can be done either by trying out all candidate paths in a brute force manner, or by using an efficient graph algorithm you may have encountered in CS 180, such as depth- or breadth-first search.

CS 181 EXAM #2 | **NAME** ______________________
SPRING 2015 | **UCLA ID** ______________________

You have 90 minutes to complete this exam. You may assume without proof any statement proved in class.

1 Find a regular expression for each of the following languages over $\{0, 1\}$:

(1 pt) **a.** nonempty strings in which the first and last symbols are different;

(1 pt) **b.** strings in which the number of 0s is even;

(2 pts) **c.** strings not containing the substring 01;

(3 pts) **d.** strings in which the number of 0s and the number of 1s are either both even or both odd.

Solution.

a. $0\Sigma^*1 \cup 1\Sigma^*0$

b. $1^*(01^*01^*)^*$

c. 1^*0^*

d. $(\Sigma\Sigma)^*$

1 Find a regular expression for each of the following languages over $\{0, 1\}$:

a. nonempty strings in which the first and last symbols are different;

b. strings in which the number of 0s is even;

c. strings not containing the substring 01;

d. strings in which the number of 0s and the number of 1s are either both even or both odd.

a. $1(0+1)^*0 + 0(0+1)^*1$

b. $(01^*0+1)^*$

c. 1^*0^*

d. $((0+1)(0+1))^*$

0/0D

(2 pts)

2 Prove or disprove:

(2 pts) **a.** for any regular languages $L_1 \subseteq L_2 \subseteq \ldots \subseteq L_n \subseteq \ldots$, the union $\bigcup_{n=1}^{\infty} L_n$ is regular;

(3 pts) **b.** if L_1 and L_2 are two languages such that the equivalence classes of $\equiv_{L_1}$ are exactly the same as those of $\equiv_{L_2}$, then $L_1 = L_2$.

Solution.

a. False. Let $L_n = 01 \cup 0011 \cup \cdots \cup 0^n 1^n$. Then each L_n is finite and hence regular, whereas their union is the language $\{0^n 1^n : n \geqslant 1\}$, which was shown in class to be nonregular.

b. False. The languages $L_1 = \varnothing$ and $L_2 = \Sigma^*$ have the same set of equivalence classes, namely, a single class Σ^*.

(3 pts) **3** Construct a language that can be recognized by a DFA with 2015 states but not with 2014 states. Prove both claims.

Solution. Let L be the language of binary strings whose length is a multiple of 2015. Then L is recognized by the following DFA with 2015 states: $(\{0, 1, 2, \ldots, 2014\}, \{0, 1\}, \delta, 0, \{0\})$, where $\delta(q, \sigma) = (q + \sigma) \bmod 2015$. We will now show that no smaller DFA exists. For any distinct $i, j \in \{0, 1, 2, \ldots, 2014\}$, we have $0^i \mathbf{0}^{\mathbf{2015-i}} \in L$ and $0^j \mathbf{0}^{\mathbf{2015-i}} \notin L$. As a result, the 2015 strings $\varepsilon, 0, 00, 000, \ldots, 0^{2014}$ are each in a different equivalence class of $\equiv_L$. By the Myhill-Nerode theorem, any DFA for L must have at least 2015 states.

4 For each of the following languages, determine whether it is regular, and prove your answer:

(2 pts) **a.** binary strings with five times as many 0s as 1s; $1^N 0^{5N}$

(2 pts) **b.** binary strings of the form uvu, where u and v are nonempty strings; 0^*

(3 pts) **c.** strings over the decimal alphabet $\{0, 1, 2, \ldots, 9\}$ with characters in sorted order;

(3 pts) **d.** binary strings such that in every suffix, the number of 0s and the number of 1s differ by at most 2.

Solution.

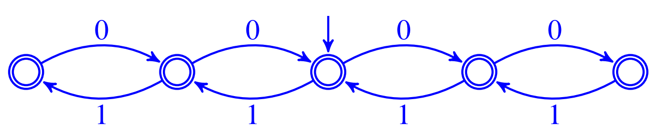
In each part, L stands for the language in question.

a. Nonregular. Let p be arbitrary, and consider the string $0^{5p}1^p$. Let x, y, z be any strings such that y is nonempty, $|xy| \leqslant p$, and $xyz = 0^{5p}1^p$. Then y is a nonempty string of 0s, and therefore the number of 0s in xz is less than five times the number of 1s. Since $xz \notin L$, the language is nonregular by (the contrapositive of) the pumping lemma.

b. Nonregular. We claim that the infinite collection of strings $\varepsilon, 0, 00, \ldots, 0^n, \ldots$ are each in a different equivalence class of $\equiv_L$. Indeed, for $n < N$, the language contains $0^n\mathbf{10}^N\mathbf{1}$ but not $0^N\mathbf{10}^N\mathbf{1}$. By the Myhill-Nerode theorem, L is nonregular.

c. Regular. This language is given by the regular expression $0^*1^*2^*3^* \cdots 9^*$.

d. Regular. Observe that L^R, the reverse of L, is the language of strings with the property that in every *prefix*, the number of 0s and the number of 1s differ by at most 2. This language is regular because it is recognized by the following NFA:



Since regular languages are closed under the reverse operation, L must be regular as well.

CS 181 EXAM #3 NAME ________________________________
SPRING 2015 UCLA ID ________________________________

You have 90 minutes to complete this exam. You may assume without proof any statement proved in class.

$a^{N_1} b^{n_1} \cdots a^{N_k} b^{n_k}$

1 Consider the following context-free grammar:

$N_i \geq n_i$

$$S \to SS \mid T$$
$$T \to aT \mid aTb \mid ab \mid a.$$

(1 pts) **a.** Describe the language generated by this grammar.

(1 pts) **b.** Prove that this grammar is ambiguous. 歧义

(2 pts) **c.** Give an equivalent unambiguous grammar.

$S \to SS \mid TF$

不能多了a

ab

$T \to aT \mid a$

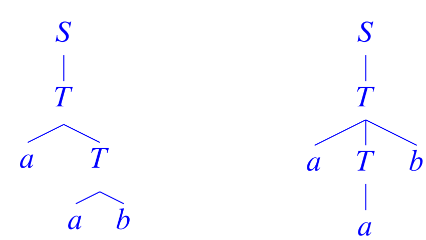
$F \to aFb \mid ab$

Solution.

a. Nonempty strings of the form $a^{N_1} b^{n_1} a^{N_2} b^{n_2} \ldots a^{N_k} b^{n_k}$, where

$$N_1 \geqslant n_1,$$
$$N_2 \geqslant n_2,$$
$$\vdots$$
$$N_k \geqslant n_k.$$

b. The string aab has at least two parse trees:



c. $S \to ATS \mid TS \mid AT \mid A \mid T$
$A \to aA \mid a$
$T \to aTb \mid ab.$

2 Give context-free grammars for the following languages over the binary alphabet:

(2 pts) **a.** nonempty even-length strings with the two middle symbols equal

(2 pts) **b.** strings with twice as many 0s as 1s

(3 pts) **c.** $\{0^n 1^m : n < m < \frac{2015}{2014} n\}$.

$\{0^n 1^m : n < m < 2n\}$

$S \to 0S1 \mid 0S11 \mid 00111$

Solution:

a. $S \to \Sigma S \Sigma \mid 00 \mid 11$
$\Sigma \to 0 \mid 1$

$S \to 0S1 \mid 0^{2015} S 1^{2016} \mid 0^{2014} 1^{2015}$

b. $S \to 1S0S0S \mid 0S1S0S \mid 0S0S1S \mid \varepsilon$

$2015 < 2016 < 2015 \times \frac{2015}{2014}$

c. $S \to 0S1 \mid 0T1$
$T \to 0^{2014} T 1^{2015} \mid 0^{2014} 1^{2015}$

$m = \frac{2015}{2014} n$

$2014m = 2015n$

3 True or false? Prove your answer.

(2 pts) **a.** If L is context-free, then the set of all substrings of strings in L is a context-free language.

(3 pts) **b.** If L is not context-free and F is finite, then $L \setminus F$ is not context-free.

Solution.

a. True. The set of all substrings of strings in L is prefix(suffix(L)), which is context-free whenever L is context-free (by the closure of context-free languages under prefix and suffix).

b. True. We will prove the contrapositive: if F is finite and $L \setminus F$ context-free, then L is context-free. For this, write

$$L = (L \setminus F) \cup (L \cap F).$$

For any finite F, the language $L \cap F$ is also finite, hence regular, hence context-free. We conclude that, with F finite and $L \setminus F$ context-free, L is the union of two context-free languages and is therefore itself context-free (by the closure of context-free languages under union).

2 Give context-free grammars for the following languages over the binary alphabet:

a. nonempty even-length strings with the two middle symbols equal

b. strings with twice as many 0s as 1s

c. $\{0^n 1^m : n < m < \frac{2015}{2014} n\}$.

a. $S \to A \mid B$

$A \to XAX \mid aa$

$2014n < 2014m < 2015n$

$B \to XBX \mid bb$

$X \to a \mid b$

b. $S \to 0S0S1 \mid 0S1S0 \mid 1S0S0 \mid SS \mid \varepsilon$

c. $S \to 0S1 \mid 0T1$

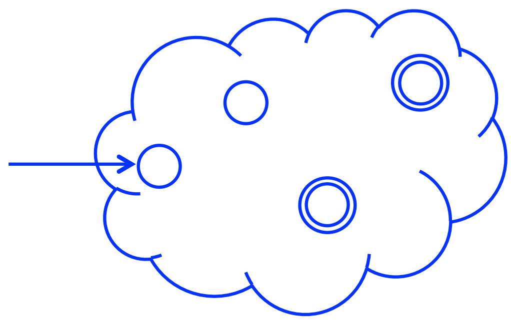
$T \to 0^{2014} T 1^{2015} \mid 0^{2014} 1^{2015}$

$S \to 0S1 \mid 0^{2014} S 1^{2015} \mid 0^{2015} 1^{2016}$

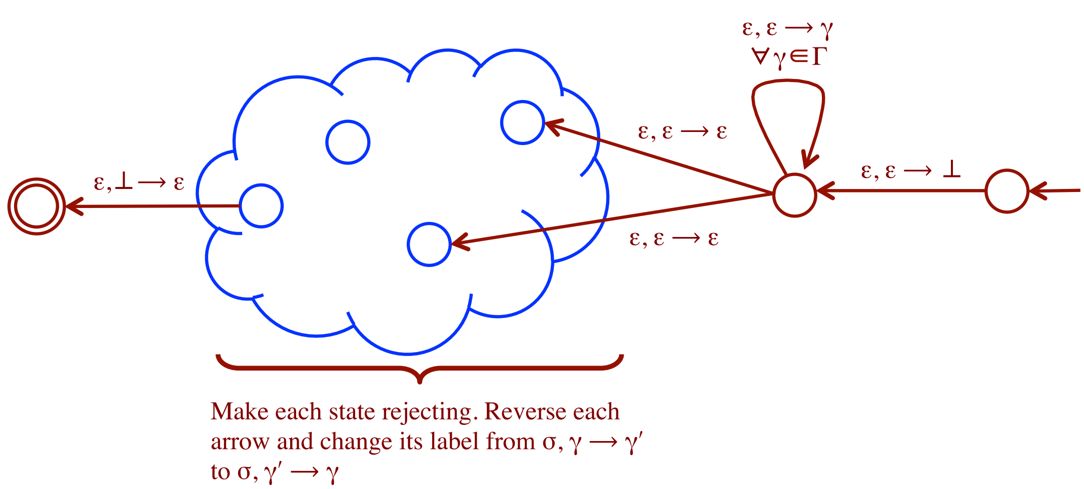
(3 pts) **4** Let L be a given context-free language. Explain how to obtain a PDA for reverse(L) from a PDA for L.

Your solution must not involve context-free grammars in any way. In particular, the following argument must not be used: convert the PDA to a grammar, reverse each rule, and convert back to a PDA.

Solution. The transformation is similar to that for DFAs and NFAs, except that one must now be careful not to forget about the stack. Suppose that, schematically, the original PDA $(Q, \Sigma, \Gamma, \delta, q_0, F)$ looks like this:



Fix a symbol, say $\perp$, that is not currently in Γ. To obtain a PDA for reverse(L), we make the structural changes shown in red:



The added loop plays a vital role in this construction. Its purpose is to "guess" the final contents of the stack for some accepting computation and to populate the stack accordingly.

5 For each of the following languages over the binary alphabet, determine whether it is context-free and prove your answer:

(2 pts) **a.** $\{wvw : w \in \{0,1\}^+, v \in \{0,1\}^*\}$

(2 pts) **b.** $\{0^n 1^m 0^k 1^{n+m} : n, m, k \geqslant 0\}$

(2 pts) **c.** palindromes with equally many 0s and 1s.

Solution. In all parts, L stands for the language is question.

a. Not context-free. Take an arbitrary integer $p \geqslant 1$ and consider the string $w = 0^p 1^p 0^p 1^p \in L$. Fix any decomposition $w = uvxyz$ for some strings u, v, x, y, z with $|v| + |y| \neq 0$ and $|vxy| \leqslant p$. There are two cases to examine: (i) if vxy is contained entirely within the first p symbols or entirely within the last p symbols, then $uv^2xy^2z \notin L$ (here, it is crucial that we pump *up* rather than *down*); (ii) if vxy overlaps with the middle $2p$ characters of w, then $uxz \notin L$. By the pumping lemma, L is not context-free.

b. Context-free, with grammar

$$\begin{aligned} S &\to 0S1 \mid T \\ T &\to 1T1 \mid U \\ U &\to 0U \mid \varepsilon. \end{aligned}$$

c. Not context-free. Take an arbitrary integer $p \geqslant 1$ and consider the string $1^p 0^{2p} 1^p \in L$. Fix any decomposition $1^p 0^{2p} 1^p = uvxyz$ for some strings u, v, x, y, z with $|v|+|y| \neq 0$ and $|vxy| \leqslant p$. There are two cases to examine: (i) if $v \in 0^*$ and $y \in 0^*$, then uv^2xy^2z contains more 0s than 1s and hence is not in L; (ii) if v or y contains a 1, then the length restriction $|vxy| \leqslant p$ implies that uxz contains unequal numbers of 1s on the left and on the right and therefore is not a palindrome: $uxz \notin L$. By the pumping lemma, L is not context-free.

5 For each of the following languages over the binary alphabet, determine whether it is context-free and prove your answer:

a. $\{wvw : w \in \{0,1\}^+, v \in \{0,1\}^*\}$

b. $\{0^n 1^m 0^k 1^{n+m} : n, m, k \geq 0\}$

c. palindromes with equally many 0s and 1s.

a. 假设 $L = \{wvw \mid w \in \{0,1\}^+, v \in \{0,1\}^*\}$ 是 CFL.

则 $N > 0$, L 满足泵引理.

取 $l = 0^N 1^N 0^N 1^N$. 易知 $w = 0^N 1^N$. $v = \varepsilon$.

$|l| > N$. $l = uvxyz$. 其中 $|vxy| < N$. $vy \neq \varepsilon$.

i. 若 vxy 属于前 N 个 0 或后 N 个 1 有.

$uv^2xy^2z \notin L$.

ii. 若 vxy 属于中间 $2N$ 个元素

则 $uxz \notin L$.

综上 l 不满足 pump lemma. 假设不成立

b. $S \to 0S1 \mid T$

$T \to 1T1 \mid P$

$P \to 0P \mid \varepsilon$

5 For each of the following languages over the binary alphabet, determine whether it is context-free and prove your answer:

a. $\{wvw : w \in \{0,1\}^+, v \in \{0,1\}^*\}$

b. $\{0^n 1^m 0^k 1^{n+m} : n, m, k \geqslant 0\}$

c. palindromes with equally many 0s and 1s.

c. 假设 $L = \{w \mid w \in \{0,1\}^*$. palindromes 回文 $\cdots\}$.

取 $z = 0^N 1^{2N} 0^N$ $|z| > N$

从 $|z|$ 满足泵引理. $z = uvwxy$

其中 $|vwx| \leq N$. $vx \neq \varepsilon$.

i. 若 vwx 属于前 N 或后 N 个元素. 易有 $uv^2wx^2y \notin L$

ii. 若 vwx 属于中间 $2N$ 个元素. 则 $uwy \notin L$ (0, 1 数量不等)