

CS/ECE 374 A ✧ Spring 2018

🌀 Homework 2 🌀

Due Tuesday, February 6, 2018 at 8pm

1. Prove that the following languages are *not* regular.

(a) $\{0^a 1 0^b 1 0^c \mid a + b = c\}$

(b) $\{w \in (0 + 1)^* \mid \#(0, w) \leq 2 \cdot \#(1, w)\}$

(c) $\{0^m 1^n \mid m + n > 0 \text{ and } \gcd(m, n) = 1\}$

Here $\gcd(m, n)$ denotes the *greatest common divisor* of m and n : the largest integer d such that both m/d and n/d are integers. In particular, $\gcd(1, n) = 1$ and $\gcd(0, n) = n$ for every positive integer n .

2. For each of the following regular expressions, describe or draw two finite-state machines:

- An NFA that accepts the same language, constructed from the given regular expression using Thompson's algorithm (described in class and in the notes).
- An equivalent DFA, constructed from your NFA using the incremental subset algorithm (described in class and in the notes). For each state in your DFA, identify the corresponding subset of states in your NFA. Your DFA should have no unreachable states.

(a) $(0 + 1)^* \cdot 0001 \cdot (0 + 1)^*$

(b) $(1 + 01 + 001)^* 0^*$

3. For each of the following languages over the alphabet $\Sigma = \{0, 1\}$, either prove that the language is regular (by constructing an appropriate DFA, NFA, or regular expression) or prove that the language is not regular (by constructing an infinite fooling set). Recall that Σ^+ denotes the set of all *nonempty* strings over Σ .

(a) Strings in which the substrings 00 and 11 do not appear the same number of times. For example, $1100011 \notin L$ because both substrings appear twice, but $01000011 \in L$.

(b) Strings in which the substrings 01 and 10 do not appear the same number of times. For example, $1100011 \notin L$ because both substrings appear twice, but $01000011 \in L$.

(c) $\{wxw \mid w, x \in \Sigma^*\}$

(d) $\{wxw \mid w, x \in \Sigma^+\}$

[Hint: Exactly two of these languages are regular. Strings can be empty.]

Solved problem

4. For each of the following languages, either prove that the language is regular (by constructing an appropriate DFA, NFA, or regular expression) or prove that the language is not regular (by constructing an infinite fooling set).

Recall that a *palindrome* is a string that equals its own reversal: $w = w^R$. Every string of length 0 or 1 is a palindrome.

- (a) Strings in $(0 + 1)^*$ in which no prefix of length at least 2 is a palindrome.

Solution: Regular: $\epsilon + 01^* + 10^*$. Call this language L_a .

Let w be an arbitrary non-empty string in $(0 + 1)^*$. Without loss of generality, assume $w = 0x$ for some string x . There are two cases to consider.

- If x contains a 0, then we can write $w = 01^n0y$ for some integer n and some string y ; but this is impossible, because the prefix 01^n0 is a palindrome of length at least 2.
- Otherwise, $x = 1^n$ for some integer n . Every prefix of w has the form 01^m for some integer $m \leq n$. Any palindrome that starts with 0 must end with 0, so the only palindrome prefixes of w are ϵ and 0, both of which have length less than 2.

We conclude that $0x \in L_a$ if and only if $x \in 1^*$. A similar argument implies that $1x \in L_a$ if and only if $x \in 0^*$. Finally, trivially, $\epsilon \in L_a$. ■

Rubric: $2\frac{1}{2}$ points = $\frac{1}{2}$ for “regular” + 1 for regular expression + 1 for justification. This is more detail than necessary for full credit.

- (b) Strings in $(0 + 1 + 2)^*$ in which no prefix of length at least 2 is a palindrome.

Solution: Not regular. Call this language L_b .

I claim that the infinite language $F = (012)^+$ is a fooling set for L_b .

Let x and y be arbitrary distinct strings in F .

Then $x = (012)^i$ and $y = (012)^j$ for some positive integers $i \neq j$.

Without loss of generality, assume $i < j$.

Let z be the suffix $(210)^i$.

- $xz = (012)^i(210)^i$ is a palindrome of length $6i \geq 2$, so $xz \notin L_b$.
- $yz = (012)^j(210)^i$ has no palindrome prefixes except ϵ and 0, because $i < j$, so $yz \in L_b$.

We conclude that F is a fooling set for L_b , as claimed.

Because F is infinite, L_b cannot be regular. ■

Rubric: $2\frac{1}{2}$ points = $\frac{1}{2}$ for “not regular” + 2 for fooling set proof (standard rubric, scaled).

(c) Strings in $(0 + 1)^*$ in which no prefix of length at least 3 is a palindrome.

Solution: Not regular. Call this language L_c .

I claim that the infinite language $F = (001101)^+$ is a fooling set for L_c .

Let x and y be arbitrary distinct strings in F .

Then $x = (001101)^i$ and $y = (001101)^j$ for some positive integers $i \neq j$.

Without loss of generality, assume $i < j$.

Let z be the suffix $(101100)^i$.

- $xz = (001101)^i(101100)^i$ is a palindrome of length $12i \geq 2$, so $xz \notin L_b$.
- $yz = (001101)^j(101100)^i$ has no palindrome prefixes except ε and 0 , because $i < j$, so $yz \in L_b$.

We conclude that F is a fooling set for L_c , as claimed.

Because F is infinite, L_c cannot be regular. ■

Rubric: $2\frac{1}{2}$ points = $\frac{1}{2}$ for “not regular” + 2 for fooling set proof (standard rubric, scaled).

(d) Strings in $(0 + 1)^*$ in which no *substring* of length at least 3 is a palindrome.

Solution: Regular. Call this language L_d .

Every palindrome of length at least 3 contains a palindrome substring of length 3 or 4. Thus, the complement language $\overline{L_d}$ is described by the regular expression

$$(0 + 1)^*(000 + 010 + 101 + 111 + 0110 + 1001)(0 + 1)^*$$

Thus, $\overline{L_d}$ is regular, so its complement L_d is also regular. ■

Solution: Regular. Call this language L_d .

In fact, L_d is *finite*! Appending either 0 or 1 to any of the underlined strings creates a palindrome suffix of length 3 or 4.

$$\varepsilon + 0 + 1 + 00 + 01 + 10 + 11 + 001 + \underline{011} + \underline{100} + 110 + \underline{0011} + \underline{1100}$$

Rubric: $2\frac{1}{2}$ points = $\frac{1}{2}$ for “regular” + 2 for proof:

- 1 for expression for $\overline{L_d}$ + 1 for applying closure
- 1 for regular expression + 1 for justification

Standard fooling set rubric. For problems worth 5 points:

- 2 points for the fooling set:
 - + 1 for explicitly describing the proposed fooling set F .
 - + 1 if the proposed set F is actually a fooling set for the target language.
 - No credit for the proof if the proposed set is not a fooling set.
 - No credit for the *problem* if the proposed set is finite.
- 3 points for the proof:
 - The proof must correctly consider *arbitrary* strings $x, y \in F$.
 - No credit for the proof unless both x and y are *always* in F .
 - No credit for the proof unless x and y can be *any* strings in F .
 - + 1 for correctly describing a suffix z that distinguishes x and y .
 - + 1 for proving either $xz \in L$ or $yz \in L$.
 - + 1 for proving either $yz \notin L$ or $xz \notin L$, respectively.

As usual, scale partial credit (rounded to nearest $\frac{1}{2}$) for problems worth fewer points.