CS/ECE 374 A (Spring 2022)
Homework 4 Solutions

Problem 4.1: For each of the following languages, determine whether it is regular or not, and give a proof. To prove that a language is not regular, you should use the fooling set method. (To prove that a language is regular, you are allowed to use known facts about regular languages, e.g., closure properties, all finite languages are regular, ...)

(a) $\{x(110)^n x^R : x \in \{0,1\}^*,\ n \geq 1\}$

(b) $\{0^i 1^j 0^k : i+k$ is divisible by 3, and k is divisible by j, and $i,j,k \geq 1\}$

(c) $\{yxx^R z : x,y,z \in \{0,1\}^*,\ |x| \geq 374\}$

(d) $\{y0^n 1^n 0^n z : y,z \in \{0,1\}^*,\ n \geq 374\}$

Solution: In each of the parts below, let L be the language in question.

(a) We prove that L is **not regular** by the fooling set method.
Choose $F = \{0^i : i \geq 0\}$.

Let x and y be two arbitrary distinct strings in F.
Then $x = 0^i$ and $y = 0^j$ for some $i \neq j$.
Choose $z = 110 \cdot 0^i$.

Then $xz = 0^i \cdot 110 \cdot 0^i = 0^i \cdot 110 \cdot (0^i)^R \in L$.
On the other hand, $yz = 0^j \cdot 110 \cdot 0^i \notin L$, because $i \neq j$ (in more detail: yz has only one occurrence of a substring of the form $(110)^n$ with $n \geq 1$, and that substring has $n = 1$; the part before the substring is 0^j and the part after the substring is 0^i, but $0^i \neq (0^j)^R$ if $i \neq j$).

Thus, F is a fooling set.
Since F is infinite, L cannot be regular.

(b) We prove that L is **not regular** by the fooling set method.
Choose $F = \{01^{3n-1} : n \geq 1\}$.
Let x and y be two arbitrary distinct strings in F.
Then $x = 01^{3m-1}$ and $y = 01^{3n-1}$ for some $m, n \geq 1$ with $m \neq n$. Without loss of generality, assume $m < n$ (the other case is symmetric).
Choose $z = 0^{3m-1}$.
Then $xz = 01^{3m-1}0^{3m-1} \in L$, since $1 + (3m-1) = 3m$ is divisible by 3 and $3m-1$ is divisible by $3m-1$.
On the other hand, $yz = 01^{3n-1}0^{3m-1} \notin L$, because $3m-1$ is not divisible by $3n-1$ since $m < n$.
Thus, F is a fooling set.
Since F is infinite, L cannot be regular.
[Note: $F = \{0^n : n \geq 1\}$ won't work here.]

(c) We prove that L is **regular**.

By definition, L consists of all strings that contain xx^R as a substring for some string x of length at least 374, i.e., all strings that contain an even-length palindrome of length at least $2 \cdot 374 = 748$.

Observe that if a string w contains an even-length palindrome of length at least 748, i.e., a substring of the form $a_\ell \cdots a_1 \cdot a_1 \cdots a_\ell$ with $a_1, \ldots, a_\ell \in \{0,1\}$ and $\ell \geq 374$, then it must contain a palindrome of length exactly 748, namely, $a_{374} \cdots a_1 \cdot a_1 \cdots a_{374}$.

Let A be the set of all palindromes of length exactly 748. Since A is finite, A is regular. Since $L = (0+1)^* A (0+1)^*$, we conclude that L is also regular.

(d) We prove that L is **not regular** by the fooling set method.

Choose $F = \{0^i : i \geq 374\}$.

Let x and y be two arbitrary distinct strings in F.

Then $x = 0^i$ and $y = 0^j$ for some $i, j \geq 374$ with $i \neq j$. Without loss of generality, assume $i > j$ (the other case is symmetric).

Choose $z = 1^i 0^i$.

Then $xz = 0^i 1^i 0^i \in L$, since xz trivially contains $0^i 1^i 0^i$ as a substring and $i \geq 374$.

On the other hand, $yz = 0^j 1^i 0^i \notin L$ by the following argument: if yz is in L, then it contains a substring of the form $0^n 1^n 0^n$ for some $n \geq 374$, then the middle block 1^n must be equal to 1^i, and so $n = i$; and the left block 0^n must be contained in 0^j, implying that $i = n \leq j$, which contradicts the $i > j$ assumption.

Thus, F is a fooling set.

Since F is infinite, L cannot be regular.

Problem 4.2: Give a context-free grammar (CFG) for each of the following languages. You must provide explanation for how your grammar works, by describing in English what is generated by each non-terminal. (Formal proofs of correctness are not required.)

(a) (30 pts) $\{x(110)^n x^R : x \in \{0,1\}^*,\ n \geq 1\}$

(b) (30 pts) $\{1^i 0^j 1^k : j = 2i + 3k,\ i, j, k \geq 0\}$

(c) (40 pts) $\{1^i 0^j 1^k : i + k \text{ is divisible by 3 and } 0 \leq j \leq k\}$

Solution:

(a)

$$\begin{aligned} S &\rightarrow 0S0 \mid 1S1 \mid A \\ A &\rightarrow 110A \mid 110 \end{aligned}$$

Explanation:

- A generates $\{(110)^n : n \geq 1\}$.
- S generates $\{x(110)^n x^R : x \in \{0,1\}^*\}$.

(b)

$$
\begin{aligned}
S &\to AB \\
A &\to 1A00 \mid \varepsilon \\
B &\to 000B1 \mid \varepsilon
\end{aligned}
$$

Explanation:

- A generates all strings of the form $1^i 0^{2i}$ $(i \geq 0)$.
- B generates all strings of the form $0^{3k} 1^k$ $(k \geq 0)$.
- S generates all strings of the form $1^i 0^{2i} \cdot 0^{3k} 1^k = 1^i 0^j 1^k$ with $j = 2i + 3k$ $(i, k \geq 0)$, as desired.

(c)

$$
\begin{aligned}
S &\to A_0B_0A_0 \mid A_0B_1A_2 \mid A_0B_2A_1 \mid \\
&\quad\; A_1B_0A_2 \mid A_1B_1A_1 \mid A_1B_2A_0 \mid \\
&\quad\; A_2B_0A_1 \mid A_2B_1A_0 \mid A_2B_2A_2 \\
A_0 &\to 111A_0 \mid \varepsilon \\
A_1 &\to 1A_0 \\
A_2 &\to 11A_0 \\
B_0 &\to 000B_0111 \mid \varepsilon \\
B_1 &\to 0B_01 \\
B_2 &\to 00B_011
\end{aligned}
$$

Explanation: Observe that $L = \{1^i 0^j 1^k : i+k \equiv 0 \bmod 3,\ i \geq 0,\ k \geq j\} = \{1^i \cdot (0^j 1^j) \cdot 1^\ell : i + j + \ell \equiv 0 \bmod 3,\ i, \ell \geq 0\}$.

- for each $p \in \{0, 1, 2\}$: A_p generates all strings of the form 1^i with $i \equiv p \bmod 3$.
- for each $q \in \{0, 1, 2\}$: B_q generates all strings of the form $0^j 1^j$ with $j \equiv q \bmod 3$.
- S generates $L = \{1^i \cdot (0^j 1^j) \cdot 1^\ell : i + j + \ell \equiv 0 \bmod 3,\ i, \ell \geq 0\}$, since S goes to $A_p B_q A_r$ over all combinations of $p, q, r \in \{0, 1, 2\}$ with $p + q + r \equiv 0 \bmod 3$.

[Note: There are other equivalent solutions. For example:

$$
\begin{aligned}
S &\to AC \mid 1AC_2 \mid 11AC_1 \\
A &\to 111A \mid \varepsilon \\
C_0 &\to 000C_0111 \mid 00C_0111 \mid 0C_0111 \mid C_0111 \mid \varepsilon \\
C_1 &\to 000C_1111 \mid 00C_1111 \mid 0C_1111 \mid C_1111 \mid 01 \mid 1 \\
C_2 &\to 000C_2111 \mid 00C_2111 \mid 0C_2111 \mid C_2111 \mid 0011 \mid 011 \mid 11
\end{aligned}
$$

Here, A generates all strings of the form 1^i with $i \equiv 0 \bmod 3$, and C_p generates all strings of the form $0^j 1^k$ with $j \leq k$ and $k \equiv p \bmod 3$.]