Automata Theory

Theory Assignment 2

Moida Praneeth Jain, 2022010193

Question 1

Pumping lemma for context free languages states that if L is a context-free language then $\exists p$ (pumping length) such that $\forall s \in L$, s can be divided into 5 substrings $s = uvxyz$ such that
- $uv^t xy^t z \in L \ \forall t \geq 0$
- $|vy| > 0$
- $|vxy| \leq p$

(a)

Given: $L = \{a^i b^j c^k \mid 0 \leq i \leq j \leq k\}$

To Prove: L is not a context-free language

Proof: Assume L is a context-free language.

$\therefore$ L satisfies the pumping lemma for context-free languages. Let p be its pumping length.

Consider the string $s = a^p b^p c^p$. $|s| = 3p > p$, $\therefore s = uvxyz$

Define a function $f : L \to \mathbb{N}$ that takes in a string from the language and outputs the number of unique characters in it. For example, $f(\text{aabc}) = 3$, $f(\text{aaa}) = 1$, $f(\varepsilon) = 0$.

- Case 1: $f(v) \leq 1 \land f(y) \leq 1$

  The maximum number of unique symbols in both v and y combined is $\max(f(v) + f(y)) = 2$ in this case. Since we have 3 symbols (a, b, c), atleast one symbol never occurs in both v and y. Let this symbol be m.
  - Subcase 1: $m = a$

    We pump down with $t = 0$, i.e, $s' = uv^0 xy^0 z = uxz$.

    v and y contained either b or c (or both). Before pumping, all three characters had equal count p. After pumping, the count of either b or c (or both) is lower than the count of a.

    $\therefore s' \notin L$
  - Subcase 2: $m = b$

    If a is present in u or v, then we pump up with $t = 2$, i.e, $s' = uv^2 xy^2 z$. The number of b remains same as p while the number of a has increased.

    Otherweise, if c is present in u or v, then we pump down with $t = 0$, i.e, $s' = uv^0 xy^0 z$. The number of b remains same as p while the number of c has decreased.

    $\therefore s' \notin L$
  - Subcase 3: $m = c$

    We pump up with $t = 2$, i.e, $s' = uv^2 xy^2 z$. The number of c remains the same as p while the number of a or b (or both) increases.

    $\therefore s' \notin L$

  In all the subcases, the pumping lemma does not hold.

- Case 2: $f(v) > 1 \vee f(y) > 1$

  We pump up with $t = 2$, i.e, $s' = uv^2xy^2z$.

  WLOG assume $f(v) > 1$, i.e, $v = l_1 l_2 l_3$ in the correct order (l_3 may be ε).

  $v^2 = l_1 l_2 l_3 l_1 l_2 l_3$. But, l_1 cannot appear after l_2. This is a contradiction.

  Pumping lemma does not hold.

In all possible cases, pumping lemma does not hold. This is a contradiction.

$\therefore$ L is not a context free language.

QED

(b)

Given: $L = \{ww \mid w \in \{0,1\}^*\}$

To Prove: L is not a context-free language

Proof: Assume L is a context-free language.

$\therefore$ L satisfies the pumping lemma for context-free languages. Let p be its pumping length.

Consider the string $s = 0^p 1^p 0^p 1^p$. $|s| = 4p > p$, $\therefore$ $s = uvxyz$

- Case 1: vxy occurs in the left half of the string, i.e, in the first w.

  We pump up with $t = 2$, i.e, $s' = uv^2xy^2z$. This pushes a 1 onto the first position of the second half, while the first position of the first half has a 0. Therefore, this string is not of the form ww.

  $\therefore$ $s' \notin L$
- Case 2: vxy occurs in the right half of the string, i.e, in the second w.

  We pump up with $t = 2$, i.e, $s' = uv^2xy^2z$. This pushes a 0 onto the last position of the first half, while the last position of the second half has a 1. Therefore, this string is not of the form ww.

  $\therefore$ $s' \notin L$
- Case 3: vxy contains the midpoint of the string

  We pump down with t = 0, i.e, $s' = uv^0xy^0z = uxz$.

  Clearly, $s' = 0^p 1^{k_1} 0^{k_2} 1^p$. Note that the initial 0^p and final 1^p remain unaffected because $|vxy| \leq p$. For s' to belong in L, $k_1 = p \wedge k_2 = p$. This is not possible as the string has been pumped down and its length can no longer be $4p$. This is a contradiction.

  $s' \notin L$

In all possible cases, pumping lemma does not hold. This is a contradiction.

$\therefore$ L is not a context free language.

QED

(c)

Given: $L = \{a^{n!} \mid n \geq 0\}$

To Prove: L is not a context-free language

Proof: Assume L is a context-free language.

$\therefore$ L satisfies the pumping lemma for context-free languages. Let p be its pumping length.

Consider the string $s = a^{p!}$. $|s| = p! \geq p$, $\therefore s = uvxyz$

On pumping the string, we get $s_i = uv^i xy^i z$

Clearly, $|s_i| = p! + (i-1)|vy|$ (Note that $|vy| > 0$ according to pumping lemma).

$\forall i \geq 0\ s_i \in L$

$\forall i \geq 0 |s_i|$ is a factorial

$\forall i \geq 0\ p! + (i-1)|vy|$ is a factorial

This implies that there exists an arithmetic progression of factorials with common difference $|vy|$.

Since Γ (The gamma function) is convex for positive inputs, no infinite sequence with linear non zero slope can fit it. So, there can exist no infinite arithmetic progression of factorials with non zero common difference,

This is a contradiction. L does not satisfy the pumping lemma.

$\therefore$ L is not a context free language.

QED

Question 2

<u>Given</u>: $F(a, b) = a_0 b_0 a_1 b_1 ... a_n b_n$

Yes, recursively enumerable languages are closed under this operation.

<u>To Prove</u>: RE languages are closed under F

<u>Proof</u>: Consider two arbitrary RE languages L_1 and L_2.

Let the turing machines that recognize them be M_1 and M_2 respectively.

Let $L = \{F(a, b) \mid a \in L_1, b \in L_2\}$

Now, we construct a turing machine M that recognizes L.

Let this turing machine have three tapes, with the original input on the first tape. Note that multi-tape turing machines are equivalent to regular turing machines.

M = " On input string w:
1. Copy over the even indices of tape 1 onto tape 2.
2. Copy over the odd indices of tape 1 onto tape 3.
3. Simulate M_1 on tape 2. If rejected, reject w.
4. Simulate M_2 on tape 3. If rejected, reject w.
5. Accept w.

"

Consider any string $c \in L$.

$c = a_0 b_0 a_1 b_1 ... a_n b_n$ where a is accepted by M_1 and b is accepted by M_2.

M copies over a onto tape 2 and b onto tape 3. These are accepted, so c is accepted.

$\therefore$ Every string in L is accepted by M.

Now consider any string $c \notin L$.

After steps 1 and 2, since $c \notin L$, either a is not accepted by M_1 or b is not accepted by M_2. In either case, c is not accepted.

$\therefore$ Every string not in L is not accepted by M.

Hence, M recognizes L. So, L is a recursively enumerable language.

$\therefore$ Recursively enumerable languages are closed under F.

QED

Question 3