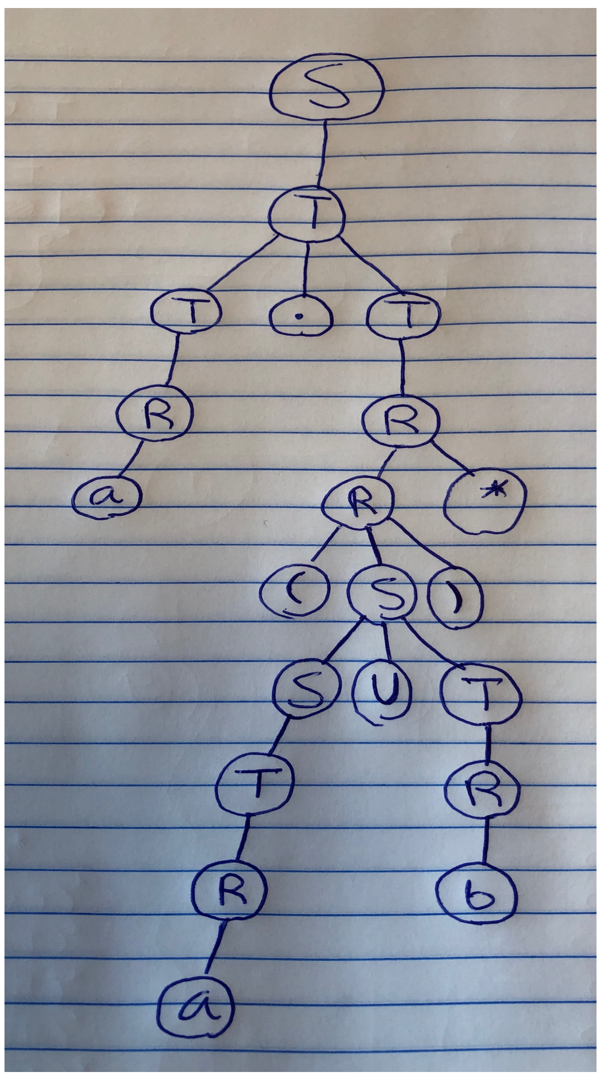
Name: Firas Aboushamalah
Course: CS3331
Student#: 250920750

Assignment 2

1. Consider the alphanumeric alphabet $\Sigma = \{a, b, \ldots, z, A, B, \ldots, Z, 0, 1, \ldots, 9\}$ and let L be the language of all regular expressions over Σ: $L = \{w \in (\Sigma \cup \{\emptyset, (,), \cup, \cdot, *\})^* \mid w$ is a syntactically legal regular expression over $\Sigma\}$

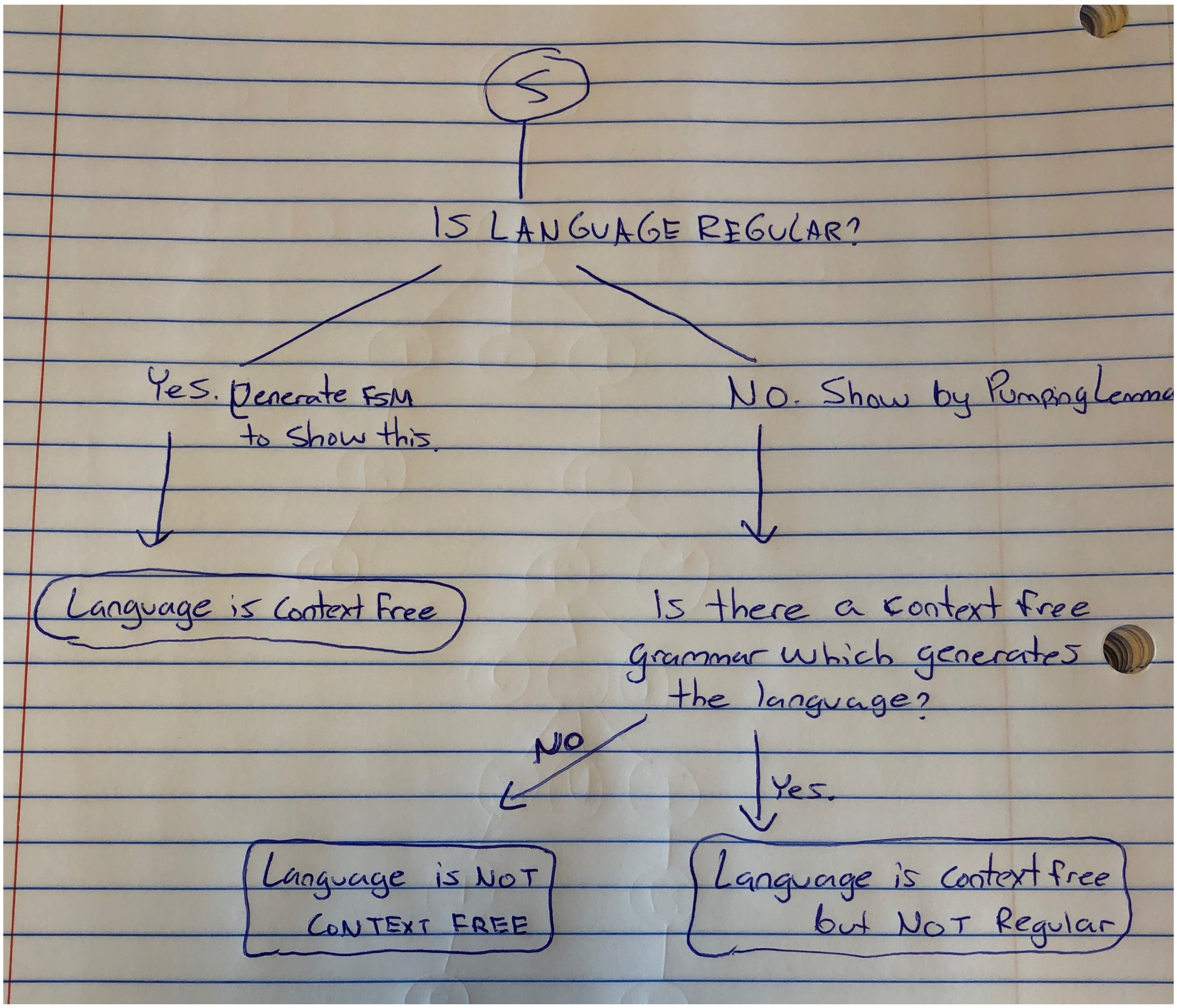
   a) Giving an unambiguous Context Free Grammar:
      $S \to S \cup T$
      $S \to T$
      $T \to T \cdot T$
      $T \to R$
      $R \to R^*$
      $R \to \Sigma$ *(this is equal to the alphanumeric set {a, b,…,z, A, B…Z, 0, 1, …}*
      $R \to \emptyset$
      $R \to (S)$

   b) Show the parse tree that your grammar produces for the string *a(a ∪b)**.



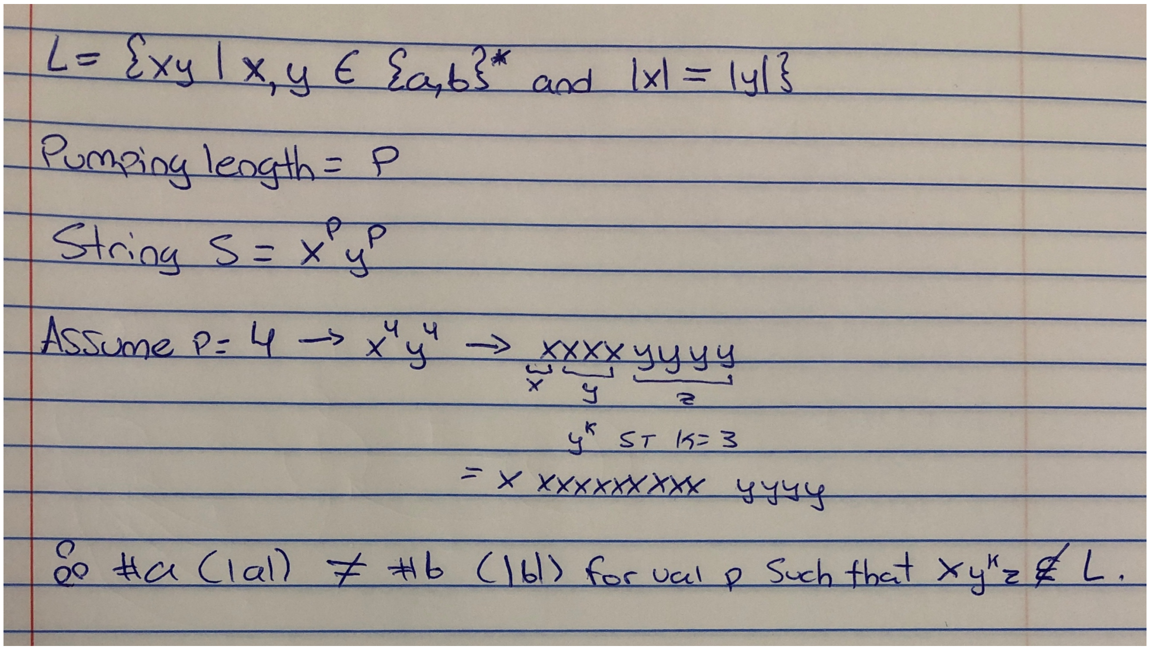
2. For each of the following languages L, prove whether L is regular, context-free but not regular, or not context-free

We know that every regular language is inherently context-free, but not every context-free language is regular. Therefore, if we prove that a language is regular then we have also proved that a language is context free. However, if a language is not regular, then it **may** be context free if it has a regular grammar that can generate the language. I have created a simple decision algorithm below:



(a) L = {xy | x, y ∈ {a, b} ∗ and |x| = |y|}

First, proving that L is not regular:


$L = \{xy \mid x,y \in \{a,b\}^* \text{ and } |x| = |y|\}$

Pumping length = P

String $S = x^P y^P$

Assume $P = 4 \to x^4 y^4 \to \underbrace{x}_{x}\underbrace{xx}_{y}\underbrace{x\,yyyy}_{z}$

y^K ST $K = 3$

$= x\ xxxxxxxxx\ yyyy$

∴ #a (|a|) ≠ #b (|b|) for val p such that $xy^kz \notin L$.

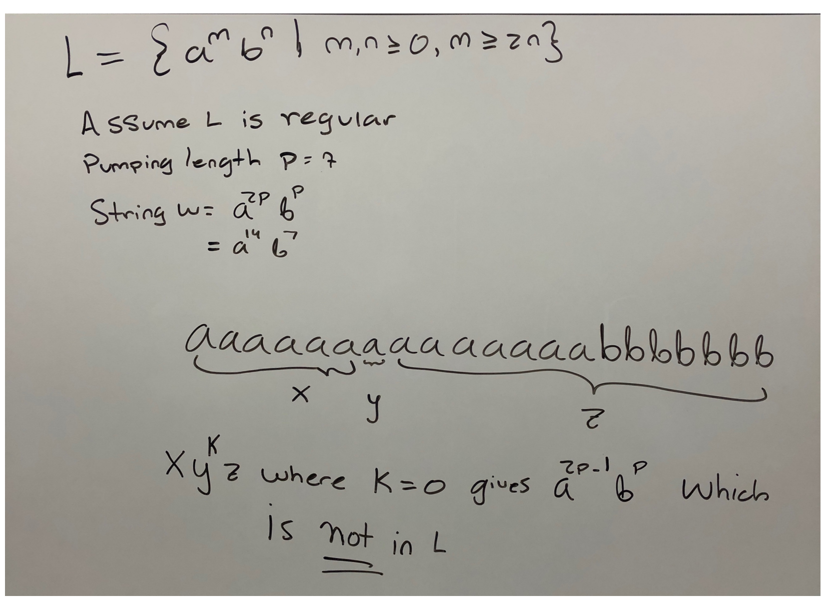
Now, following the decision tree I made – is there a context free grammar which generates L? Yes there is:

$$S \to aRb$$
$$R \to aRb \mid \varepsilon$$

Therefore, L is context-free but not regular.

a) $\{a^m b^n \mid m, n \geq 0 \text{ and } m \geq 2n\}$

The language is not regular. The proof is below:


$L = \{a^m b^n \mid m,n \geq 0, m \geq 2n\}$

Assume L is regular

Pumping length $P = 7$

String $w = a^{2P} b^P$
$= a^{14} b^7$

$\underbrace{aaaaaa}_{x}\underbrace{a}_{y}\underbrace{aaaaaaabbbbbbb}_{z}$

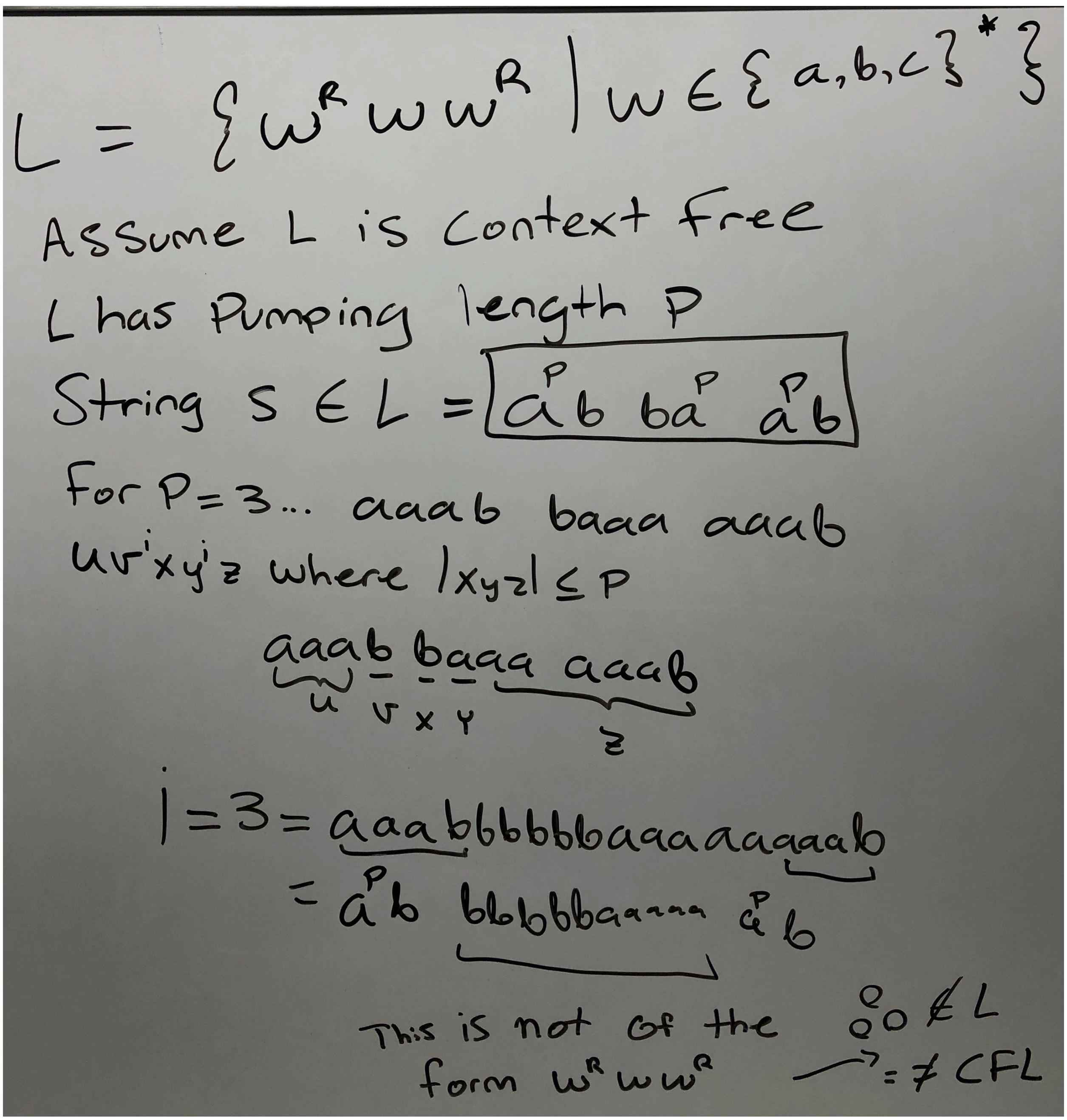
xy^Kz where $K = 0$ gives $a^{2P-1} b^P$ which is not in L

However, it **is** context free. The reason is that there exists a regular grammar as follows:

$$S \rightarrow aaRb \mid R$$
$$R \rightarrow aR \mid aaSb \mid a \mid \varepsilon$$

b) $\{w^R w w^R \mid w \in \{a, b, c\}^*\}$

Using the pumping lemma for a context free language, we can prove this is not context free.



3. Consider the language L = {wwR|w ∈ {a, b} ∗}. Below are two proofs, one showing L is context free, the other showing the opposite. Which proof is correct and why?

The more correct proof is the first one producing a context free grammar which generates L. The reason being, if we can prove that L generates one string then we can prove it generates a maximum of k strings using the same idea as the pumping lemma. A simple string which L(g) where g = CFG, is aabbaa where the following path can be taken to produce the string:

$$
\begin{aligned}
S &\rightarrow aA \\
&\rightarrow aSa \\
&\rightarrow aaAa \\
&\rightarrow aaSaa \\
&\rightarrow aabBaa \\
&\rightarrow aabSbaa \\
&\rightarrow aabbaa
\end{aligned}
$$

However, instead I would like to prove why the example using the pumping lemma to prove it is *not* context free is **incorrect:**

When producing string $s = a^p bba^p$, we may choose an arbitrary pumping length – let's use 4.

This produces the string aaaabbaaaa. Clearly, if we break down s into uvxyz, we notice that there are three possible cases which vxy takes on variables a or b. (because $|vxy| \leq p$) These are:

Case 1) $vxy = a^j b^k$ for $j, k \leq p$
Case 2) $vxy = b^j a^k$ for $j, j \leq p$
Case 3: $vxy = a^j ba^j$ for $j \leq p$

In the example done in the instructions, the pumping lemma was used incorrectly by giving both a's a different number (where at least one of p/q is not 0). However, the law of the pumping lemma shows that uvxyz must have an i for which $uv^i xy^i z$. This means both v and y get the same i. As a result, we notice that for any value $i \neq 1$, the pumping lemma does not provide a contradiction because the resulting string is always $a^j bba^j \in L$.

4. (30pt) Show that the following problem is decidable: Given a context-free grammar G, does G generate any odd-length, nonempty strings?

This can be solved by looking at the intersection and automata constructions. Firstly, recall that for a given *CFL L* and a *regular language K*, $L \cap K$ is also a context free language. Therefore, if we look at the construction of L:

$$\{w \mid w \text{ is odd length}\}$$

We can create a DFSM for L to prove its regularity and further its validity by intersecting it with the CFG to produce the language :

$$\{w \mid w \text{ is odd length, and } w \in L(G)\}.$$

Then, we can note that the language is nonempty iff the CFG G generates a word of odd-length, which it does.