(a) Prove, using the Pumping lemma for Regular Languages, that there is no Universal Regular Expression.

The universal regular language (URE) takes in the form of R\$x where R is the regular expression and x a string which is accepted by R. This must hold for all R and input x that are accepted by R

To show that URE is not a regular language:

Assume that URE is a regular language. Therefore for any R that accepts an input x, R\$x belongs to the URE Language. The alphabet is : $\{a,b,U,*,(,),\epsilon,\emptyset,\$\}$

Therefore, we can use pumping lemma for a string accepted by URE. If all cases of pumping lemma results in a string that is no longer part of URE, then we know URE is not regular language.

Pumping lemma

If URE language is regular, then there must exist a pumping length p such that any string s in URE language of length |s| >= N can be split into w = xyz

- 1. |xy| <= N
- 2. |y| > 0,
- 3. For all i >= 0, xy^iz is also a part of URE L

Let's consider the regular expression that describes a string that is static in length, such as a string of b's to some arbitrary length N.

than we have 2 possible scanning as 1241 EN and y = E

Scenario 1:

$$y = bbb...b =>$$
 We pump y then we get:
 $w_z = bbb...b$ A length $w_z = bbb...b$

For this scenario Wz is not accepted by U as R does not match 2 as R is looking for b's of longth 2N while x is only length N

$$y = bbb...b$$
 => pumping y give us:
 $k \text{ longth}$ $w_z = bbb...b$ $k \text{ bbb...b}$
where $k \text{ LN}$ $w_z = bbb...b$ $w_z = bbb...b$ $w_z = bbb...b$

For this scenario, wis is not accepted as R is looking for b's of length Ntk while x is b's of N length

As pumping lemma has failed for all possible y's in a string accepted by U, URE is not regular.

Therefore URE U regular expression does not exist

(b) Prove, using the Pumping lemma for Context-Free Languages, that there is no Universal Context-Free Grammar.

This question has the same set up as the previous question, but in the context of Context-Free Language rather than regular language

Let k be the number of nonterminal symbols in Universal Context-Free Grammar U.

For any U, consider the following Context-Free Grammar G:

$$S \rightarrow T_1$$

 $T_1 \rightarrow \alpha$
 $T_1 \rightarrow T_2$
 $T_2 \rightarrow \alpha \alpha$
 $T_2 \rightarrow T_3$
 $T_3 \rightarrow \alpha \alpha \alpha$
 $T_N \rightarrow \alpha^N$

Where any terminal Tn can be expanded to a^n or Tn+1 until the final terminal TN -> a^N

And pick the x string accepted by this as a^N In the language U, we can pick the string:

$$W = S \rightarrow T_1; T_1 \rightarrow \alpha; T_1 \rightarrow T_2; \dots; T_N \rightarrow \alpha^N \leqslant \alpha^N$$

$$N \text{ (apetition) of}$$

$$T_1 \rightarrow \alpha^n; T_n \rightarrow T_{n+1} \quad n \ge 1$$

It's important to understand that each terminal Tn+1 relies on Tn for n = 1. Once we do some pumping lemma, we can see that removing any terminals within G will break the input string a^N (which is the largest string possible)

Take $N > 2^{k-1}/3$

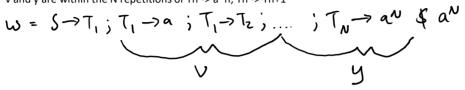
Consider any u,v,w,y,z such that

- vy =! empty string
- |vxy| <= 2^k
- w = uvxyz

We must test whether:

uxz, uvxyz, uv^ixy^iz for i=0 -> infinity are all in U?

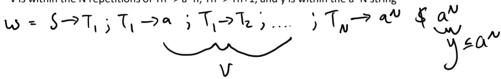
v and y are within the N repetitions of Tn -> a^n; Tn -> Tn+1



If we consider removing the v and y, and consider the string uxz: $\underline{\hspace{1cm}}$ discorrect between T_2 and T_3

This is no longer part of the language because there is now a disconnect between the terminals. Without the connection, S -> T1 cannot be expanded to TN -> a^N, making a^N invalid

v is within the N repetitions of Tn -> a^n; Tn -> Tn+1, and y is within the a^N string



If we pump v and y to produce some variation of uv^ixy^iz, then:

- The repetition of v causes a repetition of the set of context grammar rules within G of the
- The repetition of y causes a^N to be expanded to a^N+n

The repetition of the rules do not accommodate the new larger a^N+n string, therefore the string uv^ixy^iz is not accepted by U