prob8

Tuesday, 1 October 2024 4:49 PM

(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$x belongs to the URE Language. The alphabet is : {a,b,∪,∗,(,),ε,∅,$}

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.

$S = R\,\$\,x$ where $R: \underbrace{bbb\ldots b}_{N \text{ length}}$

$\therefore\ x: \underbrace{bbb\ldots b}_{N \text{ length}}$

So U accepts w=bbb...b $ bbb...b

both same length N

then we have 2 possible scenarios as $|xy| \leq N$ and $y \neq \varepsilon$

$\underbrace{bbb\ldots b}_{xy \text{ must be restricted to this range}}\ \$\ bb\ldots b$

Scenario 1:

$y = \underbrace{bbb\ldots b}_{= N \text{ length}} \Rightarrow$ if we pump y then we get:

$w_2 = \underbrace{bbb\ldots b}_{2N \text{ length}}\ \$\ \underbrace{bbb\ldots b}_{N \text{ length}}$

For this scenario w_2 is not accepted by U as R does not match x as R is looking for b's of length 2N while x is only length N

Scenario 2 :

$y = \underbrace{bbb...b}_{\text{k length}}$ where $k < N$ => pumping y gives us :

$w_3 = \underbrace{bbb...b}_{N+k \text{ length}} \$ \underbrace{bbb...b}_{N \text{ length}}$

For this scenario, w_3 is not accepted as R is looking for b's of length N+k 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:

$S \to T_1$

$T_1 \to a$

$T_1 \to T_2$

$T_2 \to aa$

$T_2 \to T_3$

$T_3 \to aaa$

$\vdots$

$T_N \to a^N$

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:

$w = S \to T_1 ; \underbrace{T_1 \to a ; T_1 \to T_2 ; \dots ; T_N \to a^N}_{N \text{ repetitions of } T_n \to a^n \,;\, T_n \to T_{n+1} \quad n \geq 1} \$ a^N$

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?

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

$$w = S \to T_1 ;\ \underbrace{T_1 \to a ;\ T_1 \to T_2 ;}_{v}\ \underbrace{\ldots ;\ T_N \to a^N \ \$}_{y}\ a^N$$

If we consider removing the v and y, and consider the string uxz:

disconnect between T_2 and T_n

$$S \to T_1 ;\ T_1 \to a ;\ T_1 \to T_2 ;\ T_n \to a^n ; \ldots T_N \to a^N ;\ a^N$$

Note: T_2 was just an example of where the disconnect can be it can be anywhere

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

$$uxz \notin U$$

Case 2:
v is within the N repetitions of Tn -> a^n; Tn -> Tn+1, and y is within the a^N string

$$w = S \to T_1 ;\ T_1 \to a\ \underbrace{;\ T_1 \to T_2 ; \ldots}_{v}\ ;\ T_N \to a^N \ \$\ \underbrace{a^N}_{y \le a^N}$$

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 word.
- 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

$$uv^2xy^2z \notin U$$