Name : Zoe Yow Cui Yi Student ID : 33214476

Subject : FIT2014 Assignment 2 Problem 8 Date:

(a) Let L be the language of the Universal Regular Expression (URE), where it would consist of strings in the form:

$R\$c$, where R is any regular expression over $\{a,b\}$ and $c \in \{a,b\}^*$

i.e. $L = \{ R\$c \mid R$ is a regular expression over $\{a,b\}$, $c \in \{a,b\}^*$, R matches $c \}$

Now, using Pumping Lemma for Regular Languages, we will prove there is no URE.

Assume by way of contradiction, L is regular.

Then, there is a Finite Automaton (FA) that recognizes it.

Let N be the number of states in such a FA.

Let w be the string $\underbrace{(((\ldots(}_{N} a \cup b \underbrace{)\ldots)))}_{N} \$ b$

This string has a length more than N.

By Pumping Lemma, w can be divided into three parts, $w = xyz$,

such that y is non-empty, $|xy| \leq N$

And for all $i > 0$, the string $xy^iz \in L$.

The requirement of $|xy| \leq N$, forces the y to fall within $\underbrace{(((\ldots(}_{N}$

Consider the string $xyyz$.

Since y can only contain the open parenthesis, there will be repetition of the open parenthesis.

This would make the string have more open paranthesis than closing parenthesis, which is not in L.

Therefore, regardless of where y is located in w, the string $xyyz \notin L$.

This violates the conclusion of Pumping Lemma for Regular Languages, so we have a contradiction.

Hence, our initial assumption that L is regular, must be false.

$\therefore$ L is not regular, and there is no URE U over $\{a, b, \cup, *, (,), \epsilon, \emptyset, \$\}$ that matches $R\$c$ if and only if R matches c.

Subject : Date:

(b) Let M be the language described by the Universal Context-Free Grammar (UCFG), which is a set of strings of the form $G\$m$, where :

- G is an encoded context-free grammar over $\{a,b\}$
- c is a string in $\{a,b\}$
- G generates c.

So, the language is :

$$M = \{G\$c \mid G \text{ is a CFG over } \{a,b\},\ c \in \{a,b\},\ G \text{ generates } c\}$$

Now, using Pumping Lemma for Context-Free languages, we will prove that there is no UCFG.

Assume by way of contradiction, M is context-free.

Then, there is a Context-Free Grammar in Chomsky Normal Form.

Let k be the number of non-terminal symbols in the CFG.

Let n be ~~the~~ any positive integer such that $n > 2^k$.

Let w be the string $G\$c$ where

$G = S \to aSb \mid \varepsilon$, and $c = a^n b^n$

i.e. $w = S \to aSb\ ;\ S \to \varepsilon\ \$\ a^n b^n$

This string has the required structure, so it belongs to M.

It is long enough, $|w| > n > 2^k > 2^{k-1}$

Therefore, by Pumping Lemma for Context-Free languages, there exist strings u, v, x, y, z such that $w = uvxyz$ and $|vxy| \leq 2^k$ and $vy \neq \varepsilon$, and for all $i \geq 0$, $uv^ixy^iz \in L$.

Observe that $|vxy| < n$, since $|vxy| \leq 2^k < n$. Consider what this means for the location of vxy with w.

Case 1:

Consider the string $uvvxyyz$.

Since $|vxy| < n$, all v, x, y are forced to be in a^n, so repetition of a's would cause more a's than b's in the string which cannot be generated by G in w. So $uv^2xy^2z \notin M$.

Case 2:

Consider the string uv^0xy^0z.

Similar to case 1, where vxy are forced to be in a^n, with less a's than b's in the string, it cannot be generated by G in w. So, $uv^0xy^0z \notin M$.

Therefore, regardless of pumping up or down vxy in w, we have $uv^ixy^iz \notin M$.

This contradicts our initial assumption. So, our initial assumption is wrong.

$\therefore$ M is not context-free and therefore there is no UCFG U that generates $G\$c$ if and only if G generates c.