CS 601 Spring 2023: Problem Set 2.

Problem 1. (10 points) Define the languages:

$$L_{add} = \{a^i b^{i+j} c^j : i, j \geq 0\}$$
$$L_{mult} = \{a^i b^{ij} c^j : i, j \geq 0\}$$

For each language, what is the smallest class it belongs to (regular, context-free, or TM-decidable)? Justify your answer – for example, if you claim context-free, then give a CFG/PDA for it and prove that it is not regular.

Problem 2. (10 points) If A, B are regular languages, show that the language $A \triangle B = \{xy : x \in A, y \in B, |x| = |y|\}$ is context-free.

Problem 3. (40 points)

(a) (5 points) The pumping lemma for CFLs that we studied in class can be used to prove that some languages are non-context-free, but it is of no help in proving that several other languages are non-context-free. For example, consider the language

$$L = \{a^i b^j c^k d^l : either\ i = 0\ or\ j = k = l\}$$

Let's see why we cannot use the pumping lemma to prove that L is not context-free. You will show that ***every*** string $s \in L$ of length p or greater, where p is the pumping length, can be expressed as $s = uvxyz$ so that all three conditions of the pumping lemma are satisfied.

Note that there are two cases: either $i = 0$ so that $s = b^j c^k d^l$ or else $i > 0$ so that $s = a^i b^j c^j d^j$.

How would you choose vxy in each case so that the pumping lemma is satisfied?

(b) (20 points)

You will now prove a stronger version of the pumping lemma:

Lemma: If L is a CFL then there exists a pumping length p such that every string s of length p or greater in which at least p positions have been marked as *distinguished*, then s can be expressed as $s = uvxyz$ such that:

i. v and y contain at least one distinguished position,
ii. vxy contains at most p distinguished positions, and
iii. $\forall i \geq 0 : uv^i xy^i z \in L$.

To simplify the proof, let's start with a CFG for L that is in Chomsky Normal Form. This means that the parse tree is a binary tree, with each variable at the bottom containing one leaf child labeled with an alphabet symbol. The proof will follow a line of reasoning like the textbook proof, with the difference that you must pay attention to the distinguished positions at the leaves of the tree. Specifically, you must show the existence of a path from root to a leaf that contains a repeated non-terminal symbol, while satisfying the first two conditions.

Hint: construct the path from the root down, always following the child that contains more distinguished positions, breaking ties arbitrarily.

(c) (5 points) Use the lemma of part (b) to prove that the language $L = \{a^i b^j c^k d^l : either\ i = 0\ or\ j = k = l\}$ is not context free.

(d) (10 points) Use the lemma of part (b) to prove that the language $\{a^i b^i c^j : i \neq j\}$ is not context-free.

Hint: consider as counterexample the string $s = a^p b^p c^{p!}$, where p is the pumping length. $p! = p(p-1)\cdots 1$ is the factorial function.