Theory of Computation, CSCI 438, Spring 2022
Essay questions for exam 2, March 11

Answer the following with as many, or as few, words as is needed to thoroughly answer the question. You may use your notes, text or the Internet to answer these questions. You may not discuss these questions with any other students.

1. One of your classmates was finally getting the idea of a pumping lemma for regular languages, when we moved on to a pumping lemma for context free languages. Describe what is similar and what is not similar about these pumping lemmas. Explain this without using the formulas. (10 pts.)

Summary:

Same	Different
Use the finite nature of the descriptions of languages in the class to show that a long string in the language must visit something twice (pigeonhole principle) so can be "pumped".	Regular – state is visited twice (used a DFA) Context-free – variable is visited twice (used a CFG)
Can only be used to show a language is not in the class, not to show it is in the class.	Long string is divided : Regular – 3 pieces Context-free – 5 pieces.
Has a pumping length.	Regular – one repeated part towards the front of the string, Context-free – two repeated parts and can be anywhere in the string.

The pumping lemmas use characteristics of the automaton/grammar to say that certain things must be true of longs string in the language. For regular languages, the pumping lemma gives what must be true due to the finite nature of DFA descriptions. For context free languages the pumping lemma gives what must be true due to the finite nature of CFG descriptions.

In both cases, the pumping lemma shows that there is some length, for which a string longer than that length, must have certain characteristics. Both use the pigeonhole principle that says that if there are only n pigeonholes and more than n pigeons, at least one hole will have more than one pigeon. For regular languages, the DFA has a finite number of states. Therefore, when recognizing long strings, one state must be visited more than once. For context free languages, the CFG has a finite number of variables. Therefore, when yielding long strings, one variable must be visited more than once.

In both cases the pumping lemmas are used to show that a language cannot be in a particular class of languages. The pumping lemmas are not able to show that a language is in a particular class of languages.

The difference is that for regular languages, all long strings in the language, will have a section towards the front of the string which must be able to be repeated zero or more times, due to visiting the state twice.

For context free languages, all long strings in the language, will have two sections fairly close together, which together must be able to be repeated zero or more times, due to repeating a variable twice in the parse tree. There are two sections because of how grammars work. The variable that must be able to be repeated, will have a left, middle and right yield. The left and right yields are what is being repeated. The middle yield will stay constant.

2. A precocious eighth grader has studied context free languages and believes that context free grammars are powerful enough to describe any language. Write an essay convincing this student of the limitations of context free grammars. (5 pts.)

There are many examples of languages for which the pumping lemma for context free languages doesn't hold. Since the pumping lemma for context free languages holds for all context free languages, whenever we have a language for which the pumping lemma does not hold, we have a language which is not context free. This can convince you that context free grammars are not powerful enough to describe any language.

Also, the definition of context free languages is that they have grammars of the form:

$G = (V, \Sigma, R, S)$

V - finite set of variables

Σ - finite set of terminal symbols where V and Σ are disjoint ($V \cap \Sigma = \Phi$)

R – finite set of rules (productions), $x \rightarrow y$ where $x \in V$ and $y \in (V \cup \Sigma_{\varepsilon})^{+}$

$S \in V$ is the start symbol

More powerful languages can be defined by allowing more flexible rules. For instance, context sensitive grammars are just like context free grammars only rules can be:

R – finite set of rules (productions), $x \rightarrow y$ where $x,y \in (V \cup \Sigma\varepsilon)^{+}$ and $|x| \leq |y|$.

The language $L=\{a^n b^n c^n \mid n \geq 0\}$ cannot be described using context free grammars but can be defined using context sensitive grammars as follows:

$S \rightarrow aSBc \mid abc$

$cB \rightarrow Bc$

$bB \rightarrow bb$

This can convince you that context free grammars are not powerful enough to describe any language.