CS350 – Properties of Context Free Languages

Purpose: Context free languages are used to describe the syntax of programming languages. The purpose of this module is to prove some properties of context free languages and also prove that some languages are not context free.

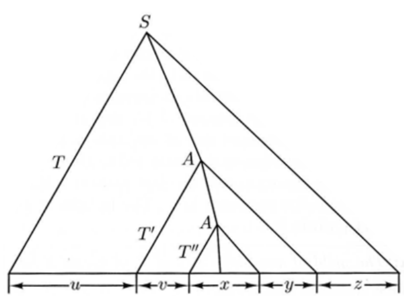
Knowledge: This module will help you become familiar with the following content knowledge:

- How to use prove closure properties of context free languages.
- How to use the context free pumping lemma to show that a language is not context free.

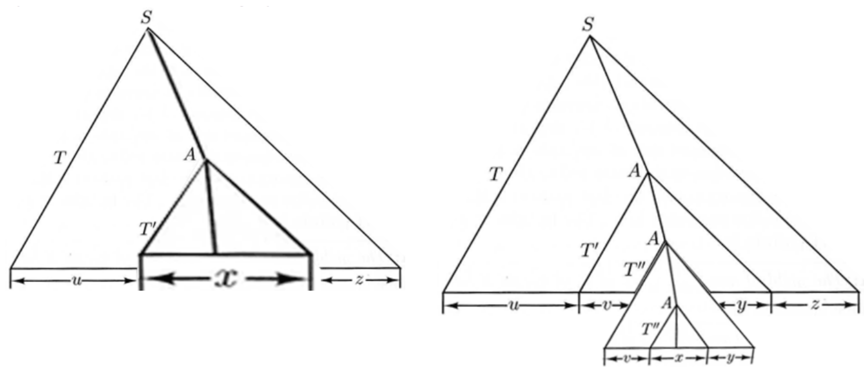
Activity: With your group perform the following tasks and answer the questions. You will be reporting your answers back to the class in 60 minutes.

1. We will start by looking at the derivation trees for context free grammars. Consider a CFG with k nonterminal symbols. How long would a string w need to be to insure that the derivation repeated a nonterminal?

2. If we have a repeated nonterminal, like A, the derivation tree would look like the following:



If we have the string $w = uvxyz$ in the language, what other strings must be in the languages?
Hint: Consider the trees on the next page



3. The property that for strings sufficiently long, a nonterminal must be repeated in the derivation is called the "Context Free Pumping Lemma". This is because we can "pump" by repeating the nonterminal as many times as we want to get other strings that will also be accepted. The context free pumping lemma is used to show that languages are *not* context free by getting a contradiction.

   The steps of a context free pumping lemma proof:

   (a) Assume the language L is context free and therefore derived by some cfg with m being the pumping length.

   (b) Choose a string w in the language of length greater than m (so there must be a repeated nonterminal).

   (c) The context free pumping lemma guarantees w can be divided into parts $uvxyz$ such that for any $i \geq 0$, uv^ixy^iz is also in L, and that $|vy| > 0$ and $|vxy| \leq m$.

   (d) For **each** way $uvxyz$ can be chosen, choose a value i such that pumping that many times gives you a string not in the language. This gives a contradiction so L is not context free.

For some practice with the context free pumping lemma, complete the following proof that $L = \{a^n b^n c^n \mid n \geq 0\}$is not context free.

Assume (towards a contradiction) that L is context free. Then the context free pumping lemma applies to L. Let m be the pumping length.

What would be a possible value for w which will make this work?

The context free pumping lemma guarantees w can be divided into parts $uvxyz$ such that for any $i \geq 0$, uv^ixy^iz is also in L, and that $|vy| > 0$ and $|vxy| \leq m$.

Give *all* the possible ways in which vxy might be given. For one example, vxy might contain all a's.

The only way that pumping will result in a string still in L is if vxy contains an equal number of a's, b's, an c's. Why is this impossible?

This leads to a contradiction so the assumption the L is context free is false.

4. Suppose you have two context free grammars. One has a start symbol S_1 and the other has a start symbol S_2. Consider a new context free grammar that contains all the rules of the two previous grammars with the addition of the rule $S \rightarrow S_1 \mid S_2$

   Why does this show that context free languages are closed under the union operation?

   In a similar construction, how would you show that context free languages are closed under concatenation?

   How would you show that context free languages are closed under Kleene star?

5. Consider the context free grammar for the language $L_1 = \{a^n b^n c^k \mid n \geq 0, k \geq 0\}$

   $S_1 \rightarrow AB$
   $A \rightarrow aAB|\lambda$
   $B \rightarrow cB|\lambda$

   Also consider the context free grammar for the language $L_2 = \{a^n b^k c^k \mid n \geq 0, k \geq 0\}$

   $S_2 \rightarrow CD$
   $C \rightarrow aC|\lambda$
   $D \rightarrow bDc|\lambda$

   What is the language $L_1 \cap L_2$?
   Since we have already shown that this language is not context free this shows that context free languages are not closed under intersection.

   By De Morgan's law, $L_1 \cap L_2 = \overline{(\overline{L}_1 \cup \overline{L}_2)}$. How can we use this to also show that context free languages are not closed under complementation?

6. Draw a Venn diagram of the classes of languages that we have studied: recursive, recursively enumerable, regular, and context free. Give an example of a language that appears in each segment of the diagram.

Complete the following assignments for grading. Each should be done individually but you may consult with a classmate to discuss strategies.

Assignment 1:
A homomorphism is a substitution in which single letters in a string are replaced with new strings. For example here is a homomorphism h:

$$h(a) = ab$$

$$h(b) = bb$$

Then $h(aba) = abbbab$.
Show that context free languages are closed under homomorphisms.

Criteria for Success: You have a clear proof on why a homomorphism applied to each of the strings of a context free language would give another context free language.

Assignment 2:
Use the context free pumping lemma to show each of the following languages is not regular. The notation $n_a(w)$ indicates the number of a's in the string w.

1. $L = \{a^n b^j c^k \mid k > n, k > j\}$
2. $L = \{w \in \{a, b, c\}^* \mid n_a(w) + n_b(w) = 2n_c(w), n_a(w) = n_b(w)\}$

Criteria for Success: You have a full context pumping lemma proof for each language.

Submit your answers in Canvas for grading