Assignment 6

1. Consider the following "addition checker" language $\{a^i b^j c^{i+j} : i, j \geq 0\}$ where the number of c's is equal to the sum of the number of a's and b's. Provide a grammar for that language and formally prove that the grammar accepts the language. You may use without proof the fact that the grammar $S \to 0S1|\epsilon$ generates the language $\{0^i 1^i : i \geq 0\}$.

   We have the grammar: $G = (\{S\}, \{a, b, c\}, P, S)$ with the productions:
   $S \to aSc|X$
   $X \to bXc|\epsilon$
   We claim that this grammar is for the above language, which means $L(G) = L$.

   Firstly, we prove $L(G) \subseteq L$. That means, if $w \in \Sigma^\star$ such that $S \Rightarrow^\star w$ then $w \in L$. Prove by induction on number of steps in a derivation.
   Base case: Number of steps is 2.
   If we derive a word in 2 steps, the word that is derived is $S \Rightarrow X \Rightarrow \epsilon$ which is in L.
   Inductive Hypothesis: Assume there exists some $n \geq 2$ such that for all n steps derivations, the word that is derived is $a^i b^j c^{i+j}$ and $i + j = n - 2$.
   Inductive Step: Prove this is true for $n+1$ steps derivation. Suppose that $w \in \Sigma^\star$ such that $S \Rightarrow^{n+1} w$.
   We can start at S and rewrite $w = axc$ such that $S \Rightarrow^n x$. By our induction hypothesis, the word $x = a^i b^j c^{i+j}$. This means that $w = axc = aa^i b^j c^{i+j} c = a^{i+1} b^j c^{i+j+1}$ after $n + 1$ steps, which is in L as needed.
   The other case start at $S \Rightarrow X \Rightarrow^n b^{n-1} c^{n-1}$ which is in L as needed.
   Therefore, $L(G) \subseteq L$.

   Now we prove $L \subseteq L(G)$.
   Suppose that $z \in L$ and $z = a^i b^j c^{i+j}$ for some $i, j > 0$. We want to show that $S \Rightarrow^\star z$.
   Prove by induction on the value of i in the word z.
   Base case: $z = a^0 b^j c^j$. Then $S \Rightarrow b^j c^j$ is a derivation of z.
   Inductive Hypothesis: Assume that there is some value n such that $a^n b^j c^{n+j}$ has a derivation in L.
   Inductive Step: Show that there is a derivation for $a^{n+1} b^j c^{n+1+j}$.
   We start with $S \Rightarrow aSc$. By our Inductive Hypothesis, we have that $S \Rightarrow^\star a^n b^j c^{n+j}$. We use that as the last derivation of S in the sentinel form aSc. Thus, $S \Rightarrow aSc \Rightarrow^n aa^n b^j c^{n+j} c = a^{n+1} b^j c^{n+1+j}$, as needed.

   Therefore, $L \subseteq L(G)$.

   Since $L(G) \subseteq L$ and $L \subseteq L(G)$, then $L(G) = L$, means the grammar accepts the language.

2. (Note: this question has two parts that are each worth one third of the assignment's marks.) In class we showed that the language `prefix`(L) is closed for context-free languages, meaning that if L is context-free then `prefix`(L) is also context free. This does not imply, however, that if L is *not regular* then `prefix`(L) is *not regular*.

   (a) Give an example non-regular context-free language L_1 and *prove* that `prefix`(L_1) is *regular*.
   Let $L_1 = \left\{x \in \{a,b\}^\star : x = x^r\right\}$, a language for palindromes. We proved that palindromes is non regular context-free language.
   The grammar for this language has productions: $P \to \epsilon|a|b|aPa|bPb$.
   The `prefix`$(L_1) = (a+b)^\star$. We prove this as followed:
   We prove that `prefix`$(L_1) \subseteq (a+b)^\star$. This is trivial since a palindrome over the alphabet $\{a,b\}$ will have its prefix generated from some number of a and b.
   Then we prove that $(a+b)^\star \subseteq$ `prefix`(L_1). For every word $w \in (a+b)^\star$, there is a word p in L_1 that $p = ww^r$. Therefore, w will be a prefix of some words in L_1.
   Since the prefix of L_1 is regular expression, therefore the prefix is also regular.

   (b) Give an example non-regular context-free language L_2 and *prove* that `prefix`(L_2) is *not regular*.
   Let $L_2 = \left\{0^i 1^i : i \geq 0\right\}$. We have proven that this language is not regular in class.
   We have the grammar that generates the language L_2 is $G = (V, \Sigma, P, S)$ with productions $S \to 0S1|\epsilon$.
   `prefix`$(L_2) = 0^i 1^j$ such that $j \leq i$.
   We will use pumping lemma to show L_2 is not regular. We prove this by contradiction.
   Assume L_2 is regular. Therefore there exists some pumping length n.
   We choose a word $z = 0^n 1^n$ which is in `prefix`L_2 and it is also longer than n. We consider $z = uvw$ where $|uv| \leq n$ and $|v| \geq 1$.
   We have $uv = 0^{i+j}$ and let $u = 0^i$ and $v = 0^j$. Then $w = 0^{n-i-j}1^n$.
   Now we consider $w^0 = uv^0w = 0^i 0^{n-i-j} 1^n$ where $j \geq 1$. But we have $i + n - i - j = n - j < n$ thus the word is not an element of `prefix`(L_2).
   Therefore, the language `prefix`(L_2) is not regular.

   Note that you may specify L_1 or L_2 using math notation and give a grammar *without* doing a proof that the grammar generates the language. If you use a language that we have not shown was non-regular for L_1 or L_2, however, you must prove that the language you use is non-regular (i.e., with pumping lemma, etc.).