Assignment 17, Authomata Theory

Oleg Sivokon

<2016-01-23 Sat>

Contents

1 Problems

1.1 Problem 1

Prove that language $L = \{a^i b^{i+j} c^j \mid 1 \leq i \leq j\}$ is not context-free.

1.1.1 Answer 2

Suppose, for contradiction, L is context-free, then, according to pumping lemma, the following applies:

1. p is the "pumping length".
2. For every word $z \in L$, $z = uvwxy$, s.t.
3. $|vwx| \leq p$.
4. $|vx| \geq 1$.
5. $uv^n wx^n y \in L$.

Consider $p = i$, then there are five distinct ways to decompose w into $uvwxy$. Of them three will decompose in a way such that both v and x are the same symbol, i.e. both v and x are either a, b or c.

It is easy to see none of the above can be pumped: if $v = a^r$ and $x = a^s$ then by pumping a, eventually there will be more as in z than cs, which contradicts $i \leq j$. Similarly, if we pump bs, eventually there will be more bs than as and cs together. Similarly for cs.

Another two possible decompositions are:

1. $v = a^r$ and $x = b^s$. However, again, if we pump as, i.e. $r \neq 0$, then eventually there will be more as than cs. And similarly for bs. When we pump as and bs together, eventually there will be more as than cs, again, contradicting $i \leq j$.

2. Thus the only case worth considering is where $v = b^r$ and $x = c^s$. Consider the word $z = a^p b^{2p} c^p \in L$ with this decomposition. If either $r = 0$ or $s = 0$, we proceed as above, however, if $r = s \neq 0$, then it must be the case that for all words $z' = a^p b^{p+p-r+r*i} c^{p-r+r*i}$, $z' \in L$. but it is not the case for $i = 0$. Since $|a^p| > \left|c^{p-r}\right|$ contrary to the required $i \leq j$.

These are all the possible decompositions of z, since neither can be pumped, it must be the case that L is not context-free.

1.2 Problem 2

Prove that context-free languages are not closed under *max* operation.

1.2.1 Answer 2

Recall the definition of *max*:

$$max(L) = \{u \in L \mid \forall v \in \Sigma^* : uv \in L \implies v = \epsilon\} \ .$$

Let's take $L = \{a^n b^m c^k \mid n \leq k \vee m \leq k\}$. L is context-free, since we can give a context-free grammar $L(G) = L$ as follows:

$$\begin{aligned}
&S \to X \mid Y \\
&X \to aXC \mid C \\
&C \to bCc \mid Cc \mid bBCc \mid c \\
&B \to bB \mid b \\
&Y \to AZ \\
&A \to aA \mid \epsilon \\
&Z \to bZc \mid Q \\
&Q \to cQ \mid c \ .
\end{aligned}$$

However, the $max(L) = \{a^n b^n c^n\}$, which is known to be non-context-free.

1.3 Problem 3

Prove $L = \{a^i b^2 c^j \mid i = 2j\}$ is constext free using closure properties and some language from assignment 16.

1.3.1 Answer 3

Recall that we proved language $M = \{a^k b^i c^j d^{j-i} e^k \mid 1 \leq i \leq j, k \geq 2\}$ to be context-free. We can define homomorophism:

$$h(x) = \begin{cases} aa & \text{for } x = a \\ b & \text{for } x = b \lor x = c \lor x = d \\ c & \text{for } x = e \end{cases}.$$

Now, $M' = h(M) = \{a^{2k} b^{i+j+j-i} c^k\}$, where $2j$ is any even integer, thus could be rewritten as $\{a^{2k} b^{2j} c^k\}$. Due to closure of context-free languages under homomorphism, M' is context-free.

Next, we can intersect M' with a regular language $a^* b^2 c^*$ to get L. Since context-free languages are closed under intersection with regular languages we proved that L is context-free.

1.4 Problem 4

Prove or disprove each of the following statements:

1. L is a irregular context-free language. G is a context-sensitive language. $L \cap G$ is not context-free.

2. L_1 and L_2 are irregular context-free languages s.t. $L_1 \cap L_2 \neq \emptyset$. $L_1 \cap L_2$ is irregular context-free language.

3. L is a regular language over Σ. G is a context-sensitive language. Define substitution f s.t. $\forall \sigma \in \Sigma : f(\sigma) = G$. $f(L)$ is context-sensitive.

1.4.1 Anwser 4

An interesection of a context-free and a context-sensitive languages may be context-free. For instance, $\{a^n b^n\} \cap \{a^n b^n c^n\} = \{a^n b^n\}$, where $n \geq 0$ is context-free.

1.4.2 Answer 5

An interesection of two context-free languages isn't necessarily irregular. For instance $\{a^n b^n\} \cap \{a^n c^n\} = \{a^n\}$ where $n \geq 0$ is regular.

1.4.3 Answer 6

The language $L = \{\epsilon\}$ is regular. $f(L) = L$ since no substitution took place, hence this claim is false.

1.5 Problem 5

Let L be a context-free language over the alphabet $\Sigma = \{a, b, c, \dots z\}$. Prove that L' is also context-free, when defined as follows:

$$L' = \{w \mid |w| \equiv 0 \pmod 2 \ \wedge |w| \geq 4 \ \wedge \ \mathbf{Sub}(w)\}$$

Where $\mathbf{Sub}(w)$ is true whenever

$$w = \begin{cases} x\mathbf{a}y\mathbf{z}z & \text{for } xpyqz \in L \land p \neq \mathbf{a} \land q \neq \mathbf{z} \\ x\mathbf{z}y\mathbf{a}z & \text{for } xpyqz \in L \land p \neq \mathbf{z} \land q \neq \mathbf{a} \end{cases} |p| = |q| = 1$$

1.5.1 Answer 7

1. Provided L is regular, we can bring its grammar G to Greibach normal form.

2. Now, for every rule of the form $A \to xA_1A_2A_3 \dots A_n$ we introduce new rules: $A \to aA'_1A'_2A'_3 \dots A'_n$ whenever $x \neq a$ and $A' \to zA''_1A''_2A''_3 \dots A''_n$ whenever $x \neq z$.

3. We replace the rules of the form $A \to x$ with $A'' \to x$.

The resulting grammar G' will nondeterministically substitute a for some terminal, which does not equal a and z for some terminal which does not equal z. It can only terminate the derivation when both substitutions took place. Using the same technique we can construct grammar G'' which first replaces z and then a. The union of G' and G'' (still a context-free grammar, since context-free languages are closed under union) will take care of $\mathbf{Sub}(w)$ condition.

Now, we can take $G''' = (G' \cup G'') \cap R$, where $R = \{r \mid r \in \Sigma^+ \land |r| \geq 4\}$. Since R is regular, and intersection of context-free and regular languages is known to be context-free, G''' must be context-free. This completes the proof.