MA2C03: TUTORIAL 8 PROBLEMS
FORMAL LANGUAGES AND GRAMMARS

1) Describe the formal language over the alphabet $\{a, b, c\}$ generated by the context-free grammar whose only non-terminal is $\langle S \rangle$, whose start symbol is $\langle S \rangle$, and whose production rules are the following:
(1) $\langle S \rangle \to b$
(2) $\langle S \rangle \to c$
(3) $\langle S \rangle \to a\langle S \rangle$
In other words, describe the structure of the strings generated by this grammar.

Solution: This context-free grammar produces strings of the type $a^m b$ or $a^m c$ for $m \geq 0$.

2) Let L be the language over the alphabet $\{0, 1\}$ consisting of all words where the string 00 occurs as a substring. Prove from the definition of a regular language that the language L is regular.

Solution: Let the alphabet $A = \{0, 1\}$. Recall that the definition of a regular language allows for finite subsets of A^*, the Kleene star, concatenations, and unions. Note that

$$L = \{w \in A^* \mid w = u \circ 00 \circ v \quad u, v \in A^*\}.$$

Therefore, we can let $L_1 = \{00\}$ be the language consisting of just the string 00 of interest. L_1 is a finite set, so it is allowed in the definition of a regular language. Let $L_2 = \{0, 1\}$. L_2 is finite, hence likewise allowed. Let $L_3 = L_2^*$, the Kleene star applied to L_2. The language $L_3 = A^*$, i.e. it is the set of all words that can be formed over the alphabet $A = \{0, 1\}$. Set $L_4 = L_3 \circ L_1$, and then $L_5 = L_4 \circ L_3$. Note that the words in L_5 have exactly the structure of the words in L, and in fact, $L = L_5$. Note also that the solution here is by no means unique. No two of you will necessarily have arrived at the same exact expression, order or labelling of the intermediate languages L_i that come into the definition of a regular language as applied to L.