Discussion Problems 5

Problem One: Nonregular Languages

Let $L = \{ w \in \{\texttt{0}, \texttt{1}, \texttt{2}\}^* \mid w$ contains the same number of copies of the substrings **01** and **10** $\}$. This language is similar to the one in Problem Set Five, except that the alphabet is now $\{\texttt{0}, \texttt{1}, \texttt{2}\}$ instead of $\{\texttt{0}, \texttt{1}\}$.

Prove that L is not a regular language. This shows that whether a language is regular or not might depend on what the alphabet of the language is.

Problem Two: Designing CFGs

Below are a list of alphabets and languages over those alphabets. For each language, design a context-free grammar that generates that language.

i. Let $\Sigma = \{ \texttt{p}, \wedge, \vee, \neg, \to, \leftrightarrow, \texttt{(}, \texttt{)}, \top, \bot \}$ and let $PL = \{ w \in \Sigma^* \mid w$ is a legal propositional logic formula using just the variable p $\}$. Write a CFG for PL.

ii. Let $\Sigma = \{ \texttt{0}, \texttt{1} \}$ and consider the regular expression R = **(0|(10)*)*|10***. Write a CFG G such that $\mathscr{L}(R) = \mathscr{L}(G)$.

Problem Three: Uncertainty about Ambiguity

In this question, you'll explore some properties of ambiguous grammars. Consider the language following language defined over the alphabet $\Sigma = \{\texttt{1}, \geq\}$

$$GE = \{ \texttt{1}^n \geq \texttt{1}^m \mid n \geq m \}$$

Here is one possible context-free grammar for GE:

$$S \to \texttt{1}S \mid \texttt{1}S\texttt{1} \mid \geq$$

i. Show that this grammar is ambiguous.

ii. Find a different grammar for GE that is unambiguous. Briefly explain, but do not formally prove, why your grammar is unambiguous.