Finite Automata - HW5

Erica Cole

March 1, 2012

Collaborator: Will Laurance

Problem 2.6(b)

Give a context-free grammar that generates the language: The complement of the language $\{a^n b^n | n \geq 0\}$.

Let's define $L = \{a^n b^n | n \geq 0\}$. Then the complement of L is $\sum^* - L$. For this language, call it M, we know that there are three cases:

1. $m > n$: $A_1 = \{a^m b^n | m > n \geq 0\}$
2. $m < n$: $A_2 = \{a^m b^n | 0 \leq m < n\}$
3. contains substring ba: $A_3 = \{w \in \{a, b\}^* | ba \text{ is a substring of } w\}$

Then we know that $M = A_1 \cup A_2 \cup A_3$. So let's define a context-free grammar that generates the three cases. Let $G = \{V, \sum, S, R\}$ where:

- V: $\{S, X, Y, Z, W\}$
- $\sum$: $\{a, b\}$
- S: S
- R:

$$
\begin{aligned}
\mathrm{S} &\rightarrow \mathrm{X} \mid \mathrm{Y} \mid \mathrm{Z} \\
\mathrm{X} &\rightarrow \mathrm{a} \mid \mathrm{aX} \mid \mathrm{aXb} \\
\mathrm{Y} &\rightarrow \mathrm{b} \mid \mathrm{Yb} \mid \mathrm{aYb} \\
\mathrm{Z} &\rightarrow \mathrm{WbaW} \\
\mathrm{W} &\rightarrow \mathrm{WW} \mid \mathrm{a} \mid \mathrm{b} \mid \varepsilon
\end{aligned}
$$

Problem 2.6(d)

Give a context-free grammar that generates the language: $\{x_1 \# x_2 \# \cdots \# x_k | k \geq 1$ each $x_1 \in \{a, b\}^*$ and for some i and j, $x_i = x_j^{\mathcal{R}}\}$

For this language there are two cases:

1. $i = j$: This creates a palindrome for x_i
2. $i \neq j$

So we define a context-free grammar that generates these two cases. Let $G = \{V, \sum, S, R\}$ where:

- V: $\{S, A, B, C, D, E, F\}$
- $\sum$: $\{a, b\}$
- S: S
- R:

$$\begin{aligned}
&\text{S} \rightarrow \text{ABC} \mid \text{ADC} \\
&\text{B} \rightarrow \text{aBa} \mid \text{bBb} \mid \text{a} \mid \text{b} \mid \varepsilon \\
&\text{D} \rightarrow \text{aDa} \mid \text{bDb} \mid \#\text{A} \\
&\text{C} \rightarrow \#\text{E} \mid \varepsilon \\
&\text{E} \rightarrow \#\text{E} \mid \text{Ea} \mid \text{Eb} \mid \varepsilon \\
&\text{A} \rightarrow \text{F}\# \mid \varepsilon \\
&\text{F} \rightarrow \text{F}\# \mid \text{aF} \mid \text{bF} \mid \varepsilon
\end{aligned}$$

Problem 2.9

Give a context-free grammar that generates the language: $A = \{a^i b^j c^k | i = j \text{ or } j = k \text{ where } i, j, k \geq 0\}$. Is your grammar ambiguous? Why or why not?

Consider the following CFG for A.

$$\begin{aligned}
&\text{S} \rightarrow \text{AB} \mid \text{CD} \\
&\text{A} \rightarrow \text{aA} \mid \varepsilon \\
&\text{B} \rightarrow \text{bBc} \mid \varepsilon \\
&\text{C} \rightarrow \text{aCb} \mid \varepsilon \\
&\text{D} \rightarrow \text{cD} \mid \varepsilon
\end{aligned}$$

Yes, this grammar is ambiguous. This can be seen when considering the derivations for abc. Consider:

- AB $\Rightarrow$ aAB $\Rightarrow$ aB $\Rightarrow$ abBc $\Rightarrow$ abc
- CD $\Rightarrow$ aCbD $\Rightarrow$ abD $\Rightarrow$ abcD $\Rightarrow$ abc

These two derivations correctly construct the string abc using this grammar, but they are not equivalent. Thus the grammar is ambiguous.

Problem 2.13

Let $G = (V, \Sigma, R, S)$ be the following grammar. $V = \{S, T, U\}$; $\Sigma = \{0, \#\}$; and R is the set of rules:

$$\begin{aligned}
&\text{S} \rightarrow \text{TT} \mid \text{U} \\
&\text{T} \rightarrow 0\text{T} \mid \text{T}0 \mid \# \\
&\text{U} \rightarrow 0\text{U}00 \mid \#
\end{aligned}$$

1. Describe $L(G)$ in English.
   To be able to describe $L(G)$ we construct a regular expression for it. Consider $L(G) = \{0^*\#0^*\#0^*\} \cup \{0^n\#0^{2n}\}$. Then there are two parts that create $L(G)$.
   (a) $0 \cdots 0\#0 \cdots 0\#0 \cdots 0$ where each set of 0s has length of 0 or greater.

(b) $0 \cdots 0\#0 \cdots 0$ where the second set of 0s has length 2 times the length of the first set of 0s, or both sets can have length 0.

2. Prove that $L(G)$ is not regular.

To prove that $L(G)$ is not regular, we will use the Pumping Lemma. Assume that $L(G)$ is a regular language, then the Pumping Lemma applies to $L(G)$. Let p be the constant in the Pumping Lemma. Select $s = 0^p\#0^{2p}$. Then wek now that $|s| = 2p + p + 1 > p$. Then $s = xyz$ with $|y| > 0$ and $|xy| \leq p$. So we know that xy contains only 0s from the first set of 0s, and say that y contains i of those 0s. Then consider $xy^0z = 0^{p-i}\#0^{2p}$. By the Pumping Lemma, this string must be in $L(G)$, however there are $p - i$ 0s in the first set and more than $2(p - i)$ 0s in the second set. So the string is not in $L(G)$. Thus, $L(G)$ is not a regular language.

Problem 2.14

Convert the following CFG into an equivalent CFG in Chomsky Normal Form, using the procedure given in Theorem 2.9.

$$\begin{aligned} &\mathrm{A} \to \mathrm{BAB} \mid \mathrm{B} \mid \varepsilon \\ &\mathrm{B} \to 00 \mid \varepsilon \end{aligned}$$

First we start by creating a new start variable S_0.

$$\begin{aligned} &S_0 \to \mathrm{A} \\ &\mathrm{A} \to \mathrm{BAB} \mid \mathrm{B} \mid \varepsilon \\ &\mathrm{B} \to 00 \mid \varepsilon \end{aligned}$$

Then we elminate the ε rules, starting with $\mathrm{A} \to \varepsilon$, then $\mathrm{B} \to \varepsilon$.

$$\begin{aligned} &S_0 \to \mathrm{A} \mid \varepsilon \\ &\mathrm{A} \to \mathrm{BAB} \mid \mathrm{B} \mid \mathrm{BB} \\ &\mathrm{B} \to 00 \mid \varepsilon \end{aligned}$$

$$\begin{aligned} &S_0 \to \mathrm{A} \mid \varepsilon \\ &\mathrm{A} \to \mathrm{AB} \mid \mathrm{BA} \mid \mathrm{BAB} \mid \mathrm{B} \mid \mathrm{BB} \mid \mathrm{A} \\ &\mathrm{B} \to 00 \end{aligned}$$

Then we eliminate the unit rules, starting with A, then S_0.

$$\begin{aligned} &S_0 \to \mathrm{A} \mid \varepsilon \\ &\mathrm{A} \to \mathrm{AB} \mid \mathrm{BA} \mid \mathrm{BAB} \mid 00 \mid \mathrm{BB} \\ &\mathrm{B} \to 00 \end{aligned}$$

$$\begin{aligned} &S_0 \to \mathrm{AB} \mid \mathrm{BA} \mid \mathrm{BAB} \mid 00 \mid \mathrm{BB} \mid \varepsilon \\ &\mathrm{A} \to \mathrm{AB} \mid \mathrm{BA} \mid \mathrm{BAB} \mid 00 \mid \mathrm{BB} \\ &\mathrm{B} \to 00 \end{aligned}$$

Then we create a new rule $\mathrm{C} \to 0$ to adjust the double terminal for B.

$$\begin{aligned} &S_0 \to \mathrm{AB} \mid \mathrm{BA} \mid \mathrm{BAB} \mid \mathrm{CC} \mid \mathrm{BB} \mid \varepsilon \\ &\mathrm{A} \to \mathrm{AB} \mid \mathrm{BA} \mid \mathrm{BAB} \mid \mathrm{CC} \mid \mathrm{BB} \\ &\mathrm{B} \to \mathrm{CC} \\ &\mathrm{C} \to 0 \end{aligned}$$

Finally we create a new rule D $\rightarrow$ BA to adjust the triple nonterminals in S_0 and A.

$$
\begin{aligned}
S_0 &\rightarrow \text{AB} \mid \text{BA} \mid \text{DB} \mid \text{CC} \mid \text{BB} \mid \varepsilon \\
\text{A} &\rightarrow \text{AB} \mid \text{BA} \mid \text{DB} \mid \text{CC} \mid \text{BB} \\
\text{B} &\rightarrow \text{CC} \\
\text{C} &\rightarrow 0 \\
\text{D} &\rightarrow \text{BA}
\end{aligned}
$$

Problem 5

Give CFGs for following languages:

1. $A = \{a^i b^j c^k | i + j = k\}$
   Consider the following context-free grammar:

$$
\begin{aligned}
\text{S} &\rightarrow \text{W} \mid \text{N} \mid \varepsilon \\
\text{W} &\rightarrow \text{aWc} \mid \text{N} \mid \varepsilon \\
\text{N} &\rightarrow \text{bNc} \mid \varepsilon
\end{aligned}
$$

2. $B = \{a^i b^j c^k | i + j \neq k\}$
   Consider the following context-free grammar:

$$
\begin{aligned}
\text{S} &\rightarrow \text{H} \mid \text{G} \\
\text{H} &\rightarrow \text{aJ} \mid \text{Jb} \\
\text{J} &\rightarrow \text{aJ} \mid \text{Jb} \mid \text{aJb} \mid \text{N} \mid \varepsilon \\
\text{N} &\rightarrow \text{aJc} \mid \text{P} \mid \varepsilon \\
\text{P} &\rightarrow \text{bPc} \mid \varepsilon \\
\text{G} &\rightarrow \text{Kc} \\
\text{K} &\rightarrow \text{Kc} \mid \text{aKc} \mid \text{M} \quad \varepsilon \\
\text{M} &\rightarrow \text{bMc} \mid \varepsilon
\end{aligned}
$$