CS481F01 HW 4 – CFGs

A. Demers

10 Oct – due 17 Oct

Please remember to turn in each problem on a separate page, *put your name on each page*, and turn in the pages in three separate piles!

1. For each of the following languages, give a context-free grammar that generates the language.

(**a**) $\{ 0^i 1^i \mid i \geq 0 \}^*$
(**b**) $\{ 0^i 1^j 0^k \mid i = j \vee j = k \}$
(**c**) $\{ 0^i 1^j \mid i \leq j \leq 2i \}$
(**d**) $\{ w \in \{0,1\}^* \mid \sharp 1(w) = \sharp 0(w) \}$
(**e**) $\{ x \in \{0,1\}^* \mid \neg((\exists i, j)(x = (0^i 1^i)^j)) \}$

The language described in part (e) is bit subtle: it is strings over $\{0, 1\}$ that do *not* consist of j repetitions of $0^i 1^i$ for any fixed i. Examples: the empty string is not in the language, since it corresponds to the case $i = j = 0$. The strings $1w$ and $w0$ are in the language for any w. The string 001100110011 is not in the language, since it corresponds to the case $i = 2, j = 3$. The strings 0110011 and 0011000111 are both in the language. This is getting near the limit of what can be checked by a CFG.

In each case, make sure the grader can understand *why* your answer generates the specified language. Your answers to (e), (d) and possibly (c) may require informal proofs.

2. Recall from lecture that a *right linear* grammar is one in which the productions are of the form

$$S \rightarrow \varepsilon$$
$$A \rightarrow aB \quad \text{or}$$
$$A \rightarrow a$$

These grammars are often called *regular* grammars. In this question you show why.

(a) Suppose you are given an arbitrary union-dot-star regular expression $\mathcal{E}$. Show (by induction on the structure of $\mathcal{E}$) how to construct a right linear grammar G such that $L(G) = L(\mathcal{E})$. Explain your answer.

(b) Suppose you are given an arbitrary right linear grammar G. Show how to construct an NFA M such that $L(M) = L(G)$. Argue that your solution is correct.

Now you can conclude that the right linear languages are exactly the regular sets.

3. Say grammar G is a *symmetric linear* grammar if its productions are of the form

$$\begin{aligned} A &\rightarrow aBc \\ A &\rightarrow a \qquad \text{or} \\ A &\rightarrow \varepsilon \end{aligned}$$

A language L is a *symmetric linear language* if $L = L(G)$ for some symmetric linear grammar G.

(a) Give an example of a symmetric linear language that is not regular.

(b) Prove every regular language is a symmetric linear language. For example, given a DFA M, show how to construct a symmetric linear grammar G generating $L(M)$.

(c) Show that every symmetric linear language over a single-letter alphabet is regular.

Hint: Part (b) may be harder than you think. Note that states (or sets of states, or pairs of states, or ...) of a FA can be encoded into the nonterminal symbols of a grammar.