CpSc 421 Homework 8 **Due: Nov. 10, 4pm**

Extra Credit

Attempt up to three of the problems below.

1. (**15 points**) Let $G = (V, \Sigma, R, Expr)$ be a CFG with variables $V = \{Expr, Factor, Term\}$, and terminals $\Sigma = \{\mathsf{CONSTANT}, \mathsf{IDENTIFIER}, \mathsf{PLUS}, \mathsf{TIMES}, \mathsf{LPAREN}, \mathsf{RPAREN}\}$ and rules:

$$\begin{array}{rcllll} Expr & \to & Term & \mid & Expr\,\mathsf{PLUS}\,Term & \\ Term & \to & Factor & \mid & Term\,\mathsf{Times}\,Factor & \\ Factor & \to & \mathsf{IDENTIFIER} & \mid & \mathsf{CONSTANT} & \mid \mathsf{LPAREN}\,Expr\,\mathsf{RPAREN} \end{array}$$

Here are regular expressions for the terminals:

$$\begin{array}{rcl} \mathsf{CONSTANT} & \equiv & (\mathtt{0} \cup \mathtt{1} \cup \ldots \cup \mathtt{9})^+ \\ \mathsf{IDENTIFIER} & \equiv & (\mathtt{a} \cup \mathtt{b} \cup \ldots \cup \mathtt{z} \cup \mathtt{A} \cup \mathtt{B} \cup \ldots \cup \mathtt{Z})^* \\ \mathsf{PLUS} & \equiv & \mathtt{+} \\ \mathsf{TIMES} & \equiv & \mathtt{*} \\ \mathsf{LPAREN} & \equiv & \mathtt{(} \\ \mathsf{RPAREN} & \equiv & \mathtt{)} \end{array}$$

Whitespace between terminals is ignored.

For each string below, either show that it is generated by G by drawing a parsetree or showing a derivation, or explain why it is not generated by G.

(a) `2*a+b`

(b) `a+2*b`

(c) `a-1`

(d) `(aardvark+2)*antelope`

(e) `2x + 3*(y+z)`

2. (**20 points**), Sipser, problem 2.27
Let $G = (V, \Sigma, R, S)$ be the following grammar:

$$\begin{array}{rcl} STMT & \to & ASSIGN \mid IfThen \mid IfThenElse \\ IfThen & \to & \mathtt{if\ condition\ then}\ STMT \\ IfThenElse & \to & \mathtt{if\ condition\ then}\ STMT\ \mathtt{else}\ Stmt \\ ASSIGN & \to & \mathtt{a:=1} \end{array}$$

$$\begin{array}{rcl} \Sigma & = & \{\mathtt{if},\ \mathtt{condition\ then},\ \mathtt{else},\ \mathtt{a:=1}\} \\ V & = & \{STMT,\ IfThen,\ IfThenElse,\ ASSIGN\} \end{array}$$

G is a natural-looking grammar for a fragment of a programming language, but G is ambiguous.

(a) (**1**0 points) Show that G is ambiguous.

(b) (**1**0 points) Give a new, unambiguous grammar for the same language.

3. (**32 points**) For each language below, either show that it is contex-free or prove that it is not. Please give a short explanation of how any CFG or PDA that you use for your solution works.

$$\begin{array}{rcl} C_1 & = & \{\mathtt{a}^i\mathtt{b}^j\mathtt{c}^k \mid i \le j \le k\} \\ C_2 & = & \overline{C_1} \\ C_3 & = & \{\mathtt{a}^i\mathtt{b}^j \mid i \in \{j, 2j\}\} \\ C_4 & = & \overline{C_3} \end{array}$$

4. (**40 points**) Let Σ be any finite alphabet with $|\Sigma| \geq 2$. Let

$$D \quad = \quad \{s \in \Sigma^* \mid \exists w \in \Sigma^*.\ s = ww\}$$

   (a) (**10 points**) Prove that D is not context-free.

   (b) (**30 points**) Prove that $\overline{D}$ is context-free.

5. (**40 points**) A *type 0 grammar* is like a context-free grammar, except that the rules are of the form $\alpha \rightarrow \beta$ where α and β can be arbitrary strings of variables and terminals.

   (a) (**10 points**) Write a type-0 grammar that generates the language

$$\{s \in \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}^* \mid \exists n \in \mathbb{Z}^{\geq 0}.\ s = \mathsf{a}^n \mathsf{b}^n \mathsf{c}^n\}$$

   (b) (**10 points)** Show that every language that is generated by a type 0 grammar is Turing recognizable.

   (c) (**20 points)** Show that every language that is Turing recognizable is generated by a type 0 grammar.

6. (**50 points**) Let $\Sigma = \{\mathtt{1}\}$.

   (a) (**15 points**) Show a language, $F_1 \subseteq \Sigma^*$ such that F_1 is not Turing decidable.

   (b) (**15 points**) Let $F_2 \subseteq \Sigma^*$ be context-free. Show that F_2 is regular.

   (c) (**20 points**) Let $F_3 \subseteq \Sigma^*$ be *any* language. Show that F_3^* is regular.