COMP 330A 2019, Assignment 3
Due Thursday, November 7th 2019 23:59

[10%]

2.2 **a.** Use the languages $A = \{\mathtt{a}^m\mathtt{b}^n\mathtt{c}^n \mid m, n \geq 0\}$ and $B = \{\mathtt{a}^n\mathtt{b}^n\mathtt{c}^m \mid m, n \geq 0\}$ together with Example 2.36 to show that the class of context-free languages is not closed under intersection.

b. Use part (a) and DeMorgan's law (Theorem 0.20) to show that the class of context-free languages is not closed under complementation.

[8%]

2.15 Give a counterexample to show that the following construction fails to prove that the class of context-free languages is closed under star. Let A be a CFL that is generated by the CFG $G = (V, \Sigma, R, S)$. Add the new rule $S \to SS$ and call the resulting grammar G'. This grammar is supposed to generate A^*.

[10%]

2.27 Let $G = (V, \Sigma, R, \langle\text{STMT}\rangle)$ be the following grammar.

$$\begin{aligned}
\langle\text{STMT}\rangle &\to \langle\text{ASSIGN}\rangle \mid \langle\text{IF-THEN}\rangle \mid \langle\text{IF-THEN-ELSE}\rangle \\
\langle\text{IF-THEN}\rangle &\to \mathtt{if\ condition\ then}\ \langle\text{STMT}\rangle \\
\langle\text{IF-THEN-ELSE}\rangle &\to \mathtt{if\ condition\ then}\ \langle\text{STMT}\rangle\ \mathtt{else}\ \langle\text{STMT}\rangle \\
\langle\text{ASSIGN}\rangle &\to \mathtt{a{:}{=}1}
\end{aligned}$$

$$\begin{aligned}
\Sigma &= \{\mathtt{if}, \mathtt{condition}, \mathtt{then}, \mathtt{else}, \mathtt{a{:}{=}1}\} \\
V &= \{\langle\text{STMT}\rangle, \langle\text{IF-THEN}\rangle, \langle\text{IF-THEN-ELSE}\rangle, \langle\text{ASSIGN}\rangle\}
\end{aligned}$$

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

a. Show that G is ambiguous.

b. Give a new unambiguous grammar for the same language.

[10%]

2.28 Give unambiguous CFGs for the following languages.

b. $\{w \mid$ the number of $\mathtt{a}$'s and the number of $\mathtt{b}$'s in w are equal$\}$

c. $\{w \mid$ the number of $\mathtt{a}$'s is at least the number of $\mathtt{b}$'s in $w\}$

[8%]

2.30 Use the pumping lemma to show that the following languages are not context free.

a. $\{\mathtt{0}^n\mathtt{1}^n\mathtt{0}^n\mathtt{1}^n \mid n \geq 0\}$

d. $\{t_1\#t_2\#\cdots\#t_k \mid k \geq 2$, each $t_i \in \{\mathtt{a}, \mathtt{b}\}^*$, and $t_i = t_j$ for some $i \neq j\}$

[8%]

2.36 Give an example of a language that is not context free but that acts like a CFL in the pumping lemma. Prove that your example works. (See the analogous example for regular languages in Problem 1.54.)

[8%]

2.47 Let $\Sigma = \{0,1\}$ and let B be the collection of strings that contain at least one 1 in their second half. In other words, $B = \{uv| \; u \in \Sigma^*, \, v \in \Sigma^* 1 \Sigma^* \text{ and } |u| \geq |v|\}$.

a. Give a PDA that recognizes B.

b. Give a CFG that generates B.

[8%]

2.48 Let $\Sigma = \{0,1\}$. Let C_1 be the language of all strings that contain a 1 in their middle third. Let C_2 be the language of all strings that contain two 1s in their middle third. So $C_1 = \{xyz| \; x, z \in \Sigma^* \text{ and } y \in \Sigma^* 1 \Sigma^*, \text{ where } |x| = |z| \geq |y|\}$ and $C_2 = \{xyz| \; x, z \in \Sigma^* \text{ and } y \in \Sigma^* 1 \Sigma^* 1 \Sigma^*, \text{ where } |x| = |z| \geq |y|\}$.

a. Show that C_1 is a CFL.

b. Show that C_2 is not a CFL.

2.999 Let ∑ = { 0,1,∪,∘,*,(,),ø,€ }

(We used the euro sign "€" for the empty string to distinguish it from the empty string itself "**ε**",i.e. |**ε**|=0 while |€|=1).

Define L_{REG} = { w∈∑* | w is a valid REGULAR EXPRESSION }.

[10%] a. Show L_{REG} ∉ REG

[10%] b. Show L_{REG} ∈ CFL

Let ∑ = { 0,1,A,B,…,Y,Z,(,),⟪,⟫,{,},⇨,€, **,**}

(We used the euro sign "€" for the empty string to distinguish it from the empty string itself "**ε**",i.e. |**ε**|=0 while |€|=1, the special comma sign "**,**" to distinguish it from the normal comma "," , the brackets "⟪,⟫" to distinguish them from the normal brackets "⟨,⟩" and the special arrow "⇨" to distinguish it from the normal "→").

Define L_{CFG} = { w∈∑* | w is a valid Context-Free Grammar in implicit form }.

c. Show L_{CFG} ∈ CFL. HINT: the first rules of your grammar should be something like

[10%]

⟨CFG⟩→{⟨RULES⟩}
⟨RULES⟩→⟨RULE⟩ | ⟨RULE⟩,⟨RULES⟩
⟨RULE⟩→ ⟪ ⟨NAME⟩ ⟫ ⇨ ⟨STRING⟩
…

FORMAL DEFINITION OF A REGULAR EXPRESSION

DEFINITION 1.52

Say that R is a ***regular expression*** if R is

1. a for some a in the alphabet Σ,
2. ε,
3. $\emptyset$,
4. $(R_1 \cup R_2)$, where R_1 and R_2 are regular expressions,
5. $(R_1 \circ R_2)$, where R_1 and R_2 are regular expressions, or
6. (R_1^*), where R_1 is a regular expression.

€ acts as $\boldsymbol{\varepsilon}$ in the alphabet above to avoid confusion.

In items 1 and 2, the regular expressions a and ε represent the languages $\{a\}$ and $\{\varepsilon\}$, respectively. In item 3, the regular expression $\emptyset$ represents the empty language. In items 4, 5, and 6, the expressions represent the languages obtained by taking the union or concatenation of the languages R_1 and R_2, or the star of the language R_1, respectively.

A context-free grammar is in implicit form if it is only an R. The set V is extracted from R as every sub-string of the form ⟨ {A,B,…,Z}$^+$ ⟩ such as ⟨VAR⟩, or ⟨EXPRESSION⟩. The terminals are simply {0,1}, and the start variable is the leftmost part of the first rule of R.

DEFINITION 2.2

A ***context-free grammar*** is a 4-tuple (V, Σ, R, S), where

1. V is a finite set called the ***variables***,
2. Σ is a finite set, disjoint from V, called the ***terminals***,
3. R is a finite set of ***rules***, with each rule being a variable and a string of variables and terminals, and
4. $S \in V$ is the start variable.

Examples of valid strings in L_{CFG} are

(the grammar G_1 of page 102) {⟪A⟫⇨0⟪A⟫1,⟪A⟫⇨⟪B⟫,⟪B⟫⇨#}

(the grammar of example 2.3) {⟪S⟫⇨0⟪S⟫1,⟪S⟫⇨⟪S⟫⟪S⟫,⟪S⟫⇨€}