

COMP 330A 2019, Assignment 3

Due Thursday, March 17th 2022 23:59

[10%]

- 2.2
- Use the languages $A = \{a^m b^n c^n \mid m, n \geq 0\}$ and $B = \{a^n b^n c^m \mid m, n \geq 0\}$ together with Example 2.36 to show that the class of context-free languages is not closed under intersection.
 - 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 $S \rightarrow SS$ and call the resulting grammar G' . This grammar is supposed to generate A^* .

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

[10%]

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

$$\Sigma = \{\text{if, condition, then, else, a:=1}\}$$

$$V = \{\langle \text{STMT} \rangle, \langle \text{IF-THEN} \rangle, \langle \text{IF-THEN-ELSE} \rangle, \langle \text{ASSIGN} \rangle\}$$

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

- Show that G is ambiguous.
- Give a new unambiguous grammar for the same language.

[10%]

- 2.28 Give unambiguous CFGs for the following languages.

- $\{w \mid \text{the number of a's and the number of b's in } w \text{ are equal}\}$
- $\{w \mid \text{the number of a's is at least the number of b's in } w\}$

[8%]

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

- $\{0^n 1^n 0^n 1^n \mid n \geq 0\}$
- $\{t_1 \# t_2 \# \dots \# t_k \mid k \geq 2, \text{ each } t_i \in \{a, b\}^*, \text{ and } t_i = t_j \text{ for some } i \neq j\}$

[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, $B = \{uv \mid u \in \Sigma^*, v \in \Sigma^*1\Sigma^* \text{ and } |u| \geq |v|\}$.

- Give a PDA that recognizes 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 $C_1 = \{xyz \mid x, z \in \Sigma^* \text{ and } y \in \Sigma^*1\Sigma^*, \text{ where } |x| = |z| \geq |y|\}$ and $C_2 = \{xyz \mid x, z \in \Sigma^* \text{ and } y \in \Sigma^*1\Sigma^*1\Sigma^*, \text{ where } |x| = |z| \geq |y|\}$.

- Show that C_1 is a CFL.

[8%]

2.55 Let G_1 be the following grammar that we introduced in Example 2.45. Use the DK-test to show that G_1 is not a DCFG.

$$\begin{aligned} R &\rightarrow S \mid T \\ S &\rightarrow aSb \mid ab \\ T &\rightarrow aTbb \mid abb \end{aligned}$$

2.999 Let $\Sigma = \{0,1,\cup, \circ, *, (,), \emptyset, \epsilon\}$

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

Define $L_{\text{REG}} = \{w \in \Sigma^* \mid w \text{ is a valid REGULAR EXPRESSION}\}$.

[10%]

- Show $L_{\text{REG}} \notin \text{REG}$

[10%]

- Show $L_{\text{REG}} \in \text{CFL}$

Let $\Sigma = \{0,1,A,B,\dots,Y,Z,\langle,\rangle,\{\},\Rightarrow,\epsilon, ,\}$

(We used the euro sign "€" for the empty string to distinguish it from the empty string itself "ε", i.e. $|\epsilon|=0$ while $|\epsilon|=1$, the special comma sign "," to distinguish it from the normal comma ",", the brackets " \langle,\rangle " to distinguish them from the normal brackets " \langle,\rangle " and the special arrow " \Rightarrow " to distinguish it from the normal " \rightarrow ").

Define $L_{\text{CFG}} = \{w \in \Sigma^* \mid w \text{ is a valid Context-Free Grammar in implicit form}\}$.

[10%]

- Show $L_{\text{CFG}} \in \text{CFL}$. HINT: the first rules of your grammar should be something like

```

<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.

ϵ acts as ϵ in the alphabet above to avoid confusion.

In items 1 and 2, the regular expressions a and ϵ represent the languages $\{a\}$ and $\{\epsilon\}$, 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 $\langle \{A, B, \dots, Z\}^+ \rangle$ such as $\langle \text{VAR} \rangle$, or $\langle \text{EXPRESSION} \rangle$. 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) $\{\langle A \rangle \Rightarrow 0 \langle A \rangle 1, \langle A \rangle \Rightarrow \langle B \rangle, \langle B \rangle \Rightarrow \# \}$

(the grammar of example 2.3) $\{\langle S \rangle \Rightarrow 0 \langle S \rangle 1, \langle S \rangle \Rightarrow \langle S \rangle \langle S \rangle, \langle S \rangle \Rightarrow \epsilon \}$