2.27 Let $G = (V, \sum, R, \langle STMT \rangle)$ be the following grammar.

$$\langle \text{STMT} \rangle \to \langle \text{ASSIGN} \rangle | \langle \text{IF-THEN} \rangle | \langle \text{IF-THEN-ELSE} \rangle$$
$$\langle \text{IF-THEN} \rangle \to \textbf{if condition then } \langle \text{STMT} \rangle$$
$$\langle \text{IF-THEN-ELSE} \rangle \to \textbf{if condition then } \langle \text{STMT} \rangle \textbf{ else } \langle \text{STMT} \rangle$$
$$\langle \text{ASSIGN} \rangle \to \textbf{a:=1}$$

$$\sum = \{\textbf{if, condition, then, else, a:=1}\}$$
$$V = \{\langle \text{STMT} \rangle, \langle \text{IF-THEN} \rangle, \langle \text{IF-THEN-ELSE} \rangle, \langle \text{ASSIGN} \rangle\}$$

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

a. Show that G is ambiguous.

Solution. G is ambiguous with the string "**if condition then if condition then a:=1 else a:=1**". One derivation to get the string is:

$$\langle \text{STMT} \rangle \to \langle \text{IF-THEN} \rangle$$
$$\langle \text{IF-THEN} \rangle \to \textbf{if condition then } \langle \text{STMT} \rangle$$
$$\textbf{if condition then } \langle \text{IF-THEN-ELSE} \rangle$$
$$\textbf{if condition then if condition then } \langle \text{STMT} \rangle \textbf{ else } \langle \text{STMT} \rangle$$
$$\textbf{if condition then if condition then } \langle \text{ASSIGN} \rangle \textbf{ else } \langle \text{ASSIGN} \rangle$$
$$\textbf{if condition then if condition then a:=1 else a:=1}$$

Another derivation to get the string is:

$$\langle \text{STMT} \rangle \to \langle \text{IF-THEN-ELSE} \rangle$$
$$\textbf{if condition then } \langle \text{STMT} \rangle \textbf{ else } \langle \text{STMT} \rangle$$
$$\textbf{if condition then } \langle \text{IF-THEN} \rangle \textbf{ else } \langle \text{ASSIGN} \rangle$$
$$\textbf{if condition then if condition then } \langle \text{STMT} \rangle \textbf{ else a:=1}$$
$$\textbf{if condition then if condition then } \langle \text{ASSIGN} \rangle \textbf{ else a:=1}$$
$$\textbf{if condition then if condition then a:=1 else a:=1}$$

■

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

Solution. An unambiguous grammar $G' = (V', \sum, R', \langle STMT \rangle)$,
where $V' = \{\langle \text{STMT} \rangle, \langle \text{NIT} \rangle, \langle \text{NITE} \rangle, \langle \text{ASSIGN} \rangle\}$,
$\sum = \{$**if, condition, then, else, a:=1**$\}$,
and R' is described as:

$$\langle \text{STMT} \rangle \to \langle \text{ASSIGN} \rangle | \langle \text{NIT} \rangle | \langle \text{NITE} \rangle$$
$$\langle \text{NIT} \rangle \to \textbf{if condition then } \langle \text{STMT} \rangle | \textbf{if condition then } \langle \text{NITE} \rangle \textbf{ else } \langle \text{NIT} \rangle$$
$$\langle \text{NITE} \rangle \to \textbf{if condition then } \langle \text{NITE} \rangle \textbf{ else } \langle \text{NITE} \rangle | \langle \text{STMT} \rangle$$
$$\langle \text{ASSIGN} \rangle \to \textbf{a:=1}$$

■

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

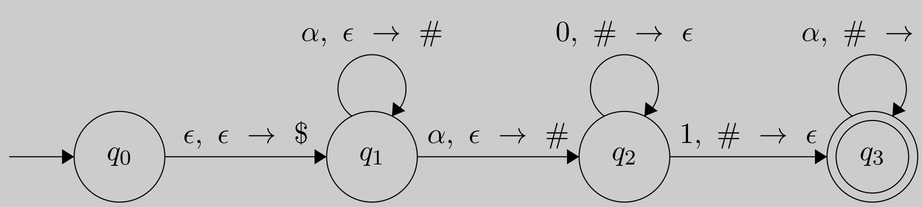
d. $\{t_1\#t_2\#\dots\#t_k | k \geq 2, \text{each } t_i \in \{a,b\}^*, \text{and } t_i = t_j \text{ for some } i \neq j\}$

Solution. Assume $L = \{t_1\#t_2\#\dots\#t_k | k \geq 2, \text{each } t_i \in \{a,b\}^*, \text{and } t_i = t_j \text{ for some } i \neq j\}$ is context-free and let p be its pumping length. Let $s = 0^p1^p\#0^p1^p \in L$ and by the pumping lemma, s can be written as $uvxyz$ where $|vxy| \leq p$ and $|vy| > 1$. Suppose vxy lies on one side of the # symbol. Then, pumping once to uv^2xy^2z gives us a result in a string where $t_1 \neq t_2$, so $uv^2xy^2z \notin L$. Now suppose that vxy contains the # symbol. If either v or y contains the # symbol, then s can be pumped down to uv^0xy^0z, which won't contain the # symbol and will not be in L. Otherwise, the # symbol is contained in x, v is a substring of 1^p, and y is a substring of 0^p, and pumping s down to uv^0xy^0z reduces either the total number of ones in t_1 or the total number of zeroes in t_2. This means that $t_1 \neq t_2$ for uv^0xy^0z and the string is not in L. ■

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

a. Give a PDA that recognizes B.

Solution.



■

b. Give a CFG that generates B.

Solution. $S \to UV$
$U \to AB$
$V \to A1A|A1B|A1U|B1U|U1U$
$A \to 00*$
$B \to 11*$ ■