You may use JFLAP to help yourself work on these exercises. You may wish to go through the tutorial sections: "Context-free Grammar" and "Pushdown Automata" available at http://www.jflap.org/modules/ (accessible through the left yellow navigation pane).

(1) For each of the Context-Free Grammars (CFGs) given below, give answers to the accompanying questions (together with short justifications where needed)

a)

$$
\begin{aligned}
R &\rightarrow XRX \mid S \\
S &\rightarrow \mathtt{a}T\mathtt{b} \mid \mathtt{b}T\mathtt{a} \\
T &\rightarrow XTX \mid X \mid \varepsilon \\
X &\rightarrow \mathtt{a} \mid \mathtt{b}
\end{aligned}
$$

1. What are the variables (non-terminals)? $V = \{\square, \square, \square, \square\}$
2. What are the terminals? $\Sigma = \{\square, \square, \square\}$
3. What is the start variable? $S = \square$
4. Give three strings in $L(G)$ $\square$, $\square$, $\square$
5. Give three strings not in $L(G)$ $\square$, $\square$, $\square$
6. True or False:
   (a) $T \rightarrow \mathtt{aba}$
   (b) $T \xrightarrow{*} \mathtt{aba}$
   (c) $T \rightarrow T$
   (d) $T \xrightarrow{*} T$
   (e) $XXX \xrightarrow{*} \mathtt{aba}$
   (f) $X \xrightarrow{*} \mathtt{aba}$
   (g) $T \xrightarrow{*} XX$
   (h) $T \xrightarrow{*} XXX$
   (i) $S \xrightarrow{*} \varepsilon$

Notation: "$\rightarrow$" means derivable in *one* step; "$\xrightarrow{*}$" means derivable in *zero or more* steps

b)

$$
\begin{aligned}
A &\rightarrow \mathtt{bb}A\mathtt{b} \mid B \\
B &\rightarrow \mathtt{a}B \mid \varepsilon
\end{aligned}
$$

Use the grammar to derive the following strings

$\mathtt{bbab}$ $\quad$ $\mathtt{bbb}$ $\quad$ $\mathtt{a}^6$ $\quad$ $\mathtt{b}^4\mathtt{a}^3\mathtt{b}^2$

c)

$$
\begin{aligned}
S &\rightarrow \mathtt{a}A\mathtt{bb} \mid \mathtt{b}B\mathtt{aa} \\
A &\rightarrow \mathtt{a}A\mathtt{bb} \mid \varepsilon \\
B &\rightarrow \mathtt{b}B\mathtt{aa} \mid \varepsilon
\end{aligned}
$$

Use the grammar to derive the following strings (where possible):

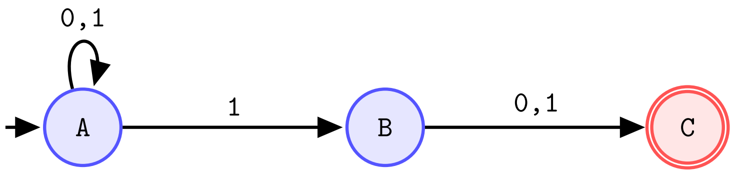
aabbbb $\quad$ bbaaaa $\quad$ aabb $\quad$ baa d)

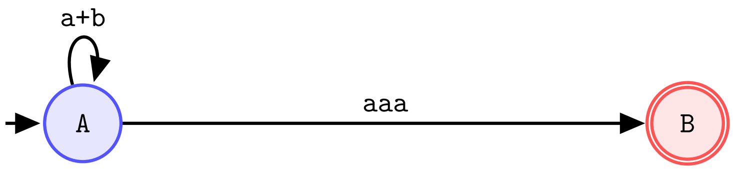
$$\begin{aligned} E &\rightarrow E + T \mid T \\ T &\rightarrow T \times F \mid F \\ F &\rightarrow (E) \mid \mathtt{a} \end{aligned}$$

Give parse trees for each of the following strings

$$\mathtt{a} \qquad \mathtt{a} + \mathtt{a} \qquad \mathtt{a} + \mathtt{a} + \mathtt{a} \qquad (\mathtt{a}) + (\mathtt{a} + \mathtt{a}) \qquad ((\mathtt{a}))$$

(2) Convert the following (G)NFAs into Regular Grammars.





(3) Design a PDA and a CFG for the following language over $\Sigma = \{\mathtt{a}, \mathtt{b}\}$

$$L = \{w \mid w = (\mathtt{ab})^n \text{ or } w = \mathtt{a}^{4n}\mathtt{b}^{3n} \text{ for } n \geq 0\}.$$

Do this in two steps:

a) Explain the idea used.

b) Design a state diagram for the PDA.

c) Design a CFG.

(4) **(Ambiguity)** Sometimes a grammar can generate the same string in several different ways, with several different parse trees, and likely several different meanings. If this happens, we say that the string is derived *ambiguously* in that grammar, which is then qualified as being an **ambiguous** grammar.

Consider the CFG

$$E \rightarrow E + E \mid E \times E \mid (E) \mid a$$

Derive the string $a + a \times a$ in two different ways using parse trees, and explain their (different) meanings.

Now note that the following alternative CFG is *not* ambiguous:

$$\begin{aligned} E &\rightarrow E + T \mid T \\ T &\rightarrow T \times F \mid F \\ F &\rightarrow (E) \mid a \end{aligned}$$

What is the parse tree for the previous example string $(a + a \times a)$?
What is the parse tree for $(a + a) \times a$?

(5) A string w is a *palindrome* if $w = w^R$, where w^R is formed by writing the symbols of w in reverse order, e.g. if $w = \texttt{011}$ then $w^R = \texttt{110}$.

Design PDAs and CFGs for each of the following languages

a) $\{w \mid w = \texttt{b}^n\texttt{ab}^n, \quad n \geq 0\}$

b) $\{w\texttt{c}w^R \mid w \in \{\texttt{a},\texttt{b}\}^*\}$ (so it is defined over the alphabet $\{\texttt{a},\texttt{b},\texttt{c}\}$)

c) $\{ww^R \mid w \in \{\texttt{a},\texttt{b}\}^*\}$

d) The language of palindromes over $\{\texttt{a},\texttt{b}\}$

e) The language of palindromes over $\{\texttt{a},\texttt{b}\}$ whose length is a multiple of 3

Hint: Consider the even and odd length cases first.

(6) Design CFGs generating the following languages.

a) $\{\texttt{a}^i\texttt{b}^j \mid i,j \geq 0 \text{ and } i \geq j\}$

b) $\{\texttt{a}^i\texttt{b}^j \mid i,j \geq 0 \text{ and } i \neq j\}$ (Complement of the language $\{\texttt{a}^n\texttt{b}^n \mid n \geq 0\}$)

c) The language of all strings over $\{\texttt{a},\texttt{b}\}$ with a single symbol ‘b’ located *exactly in the middle* of the string.

$$\{\texttt{b}, \texttt{aba}, \texttt{abb}, \texttt{bba}, \texttt{bbb}, \texttt{aabaa}, \ldots\}$$

d) The language of strings over $\{\texttt{a},\texttt{b}\}$ containing more a’s than b’s. (e.g. abaab)

e) The language of strings over $\{\texttt{a},\texttt{b}\}$ containing an equal number of a’s and b’s.

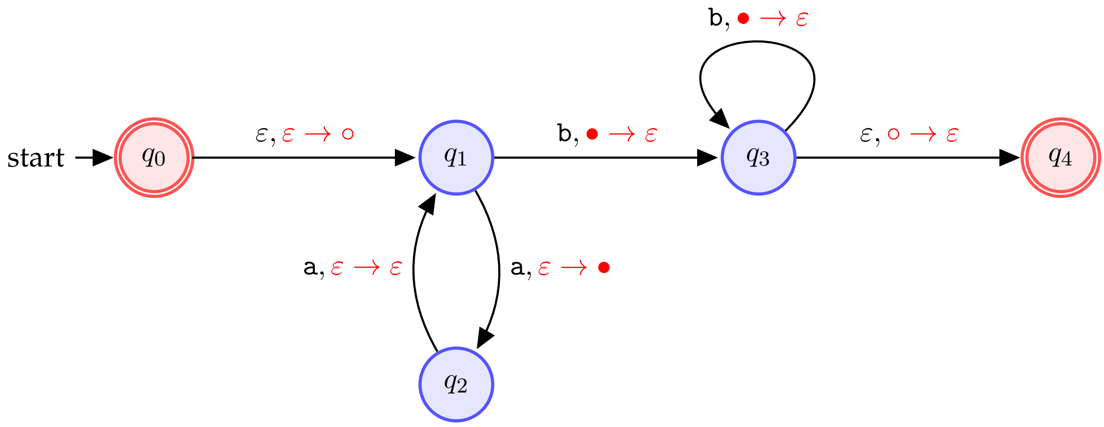
f) The language of strings with twice as many a’s as b’s.

g) $\{w\#x \mid w,x \in \{\texttt{0},\texttt{1}\}^* \text{ and } w^R \text{ is a substring of } x\}$

h) $\{x_1\#x_2\#\cdots\#x_k \mid k \geq 1, \text{ each } x_i \in \{\texttt{a},\texttt{b}\}^*, \text{ and for some } i \text{ and } j,\ x_i = x_j^R\}$

Give informal descriptions of PDAs for the above languages. (How would you use the stack?)

(1) Consider the following PDA



a) Produce the formal definition for the above NFA. This should consist of:

- The set of states $Q = \{\square, \square, \square, \square, \square\}$
- The input alphabet $\Sigma = \{\square, \square\}$
- The stack alphabet $\Gamma = \{\square, \square\}$
- The transition function, $\delta\colon Q \times \Sigma_\varepsilon \times \Gamma_\varepsilon \to 2^{Q \times \Gamma_\varepsilon}$, in table form

Σ_ε :	a			b			ε		
Γ_ε :	$\bullet$	$\circ$	ε	$\bullet$	$\circ$	ε	$\bullet$	$\circ$	ε
q_0									$\{(q_1, \circ)\}$
q_1			$\{(q_2, \bullet)\}$	$\{(q_3, \varepsilon)\}$					
q_2			$\square$						
q_3				$\{(q_3, \varepsilon)\}$				$\square$	
q_4									

The $\emptyset$ entries have been left blank to make the table easier to read.

- The start state $q_{\text{start}} = \square$
- The set of accept states $F = \{\square, \square\}$

b) Simulate the following strings: (For each step record: the state, the symbol just read and the stack contents – you may use JFLAP to help you)

`aaab` `aaaab` `aab` `aabb` `aaaabb`

c) Use set notation to describe the language recognized by this PDA.

(Notations similar to $\{\mathtt{a}^n\mathtt{b}^n \mid n \geq 0\}$ for example).

(2) Let $\Sigma = \{\mathtt{a}, \mathtt{b}\}$ and let B be the language of strings that contain at least one $\mathtt{b}$ in their second half. In other words, $B = \{uv \mid u \in \Sigma^*, v \in \Sigma^*\mathtt{b}\Sigma^* \text{ and } |v| \leq |u|\}$.

a) Give a PDA that recognizes B.

b) Give a CFG that generates B.

(3) Let

$$\begin{aligned} C &= \{x\#y \mid x, y \in \{0,1\}^* \text{ and } x \neq y\} \\ D &= \{x\#y \mid x, y \in \{0,1\}^* \text{ and } |x| = |y| \text{ but } x \neq y\} \end{aligned}$$

Show that C and D are both CFLs by producing PDAs or CFGs for them.

(4) Give a **counter example** 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^*.

*Note: the class of **context-free languages** is actually **closed under the regular operations** (union, concatenation, and star) but the above argument fails to prove closure under star. What is missing?*