Exercise 1.

(a)(i) We prove that δ is well defined by a contrapositive proof i.e. if $\exists a \in \Sigma\ [xa] \neq [ya]$, then $[x] \neq [y]$. Suppose $[xa] \neq [ya]$ for $a \in \Sigma$, then $(xa, ya) \notin R_L$. This means that $\exists w \in \Sigma^*$ such that $xaw \in L$ but $yaw \notin L$ or vice versa. Now let $z = aw$, we have $xz \in L \wedge yz \notin L$ or vice versa. Hence, $(x, y) \notin R_L$ and $[x] \neq [y]$. We conclude by contrapositive that if $[x] = [y]$ then $[xa] = [ya]$ for all $a \in \Sigma$.

(ii) We prove a more general statement i.e. $\forall x \in \Sigma^*\ \hat{\delta}([\epsilon], x) = [x]$.
This implies $\hat{\delta}([\epsilon], x) \in F$ iff $x \in L$. Since if $x \in L$, then by the definition of F, $[x] = \hat{\delta}([\epsilon], x) \in F$. Conversely, if $\hat{\delta}([\epsilon], x) \in F \Leftrightarrow [x] \in F$, then by the definition of F, $x \in L$.
Now we prove the lemma by an induction on $|x|$.

- **Basis case** $x = \epsilon$. $\delta([\epsilon], \epsilon) = [\epsilon]$ by the definition of the transition function.
- **Step case** Assume that $\hat{\delta}([\epsilon], x) = [x]$ for $|x| < k$. We prove for $w = xa$ where $x \in \Sigma^{k-1}, a \in \Sigma$.

$$\begin{aligned}
\hat{\delta}([\epsilon], xa) &= \delta(\hat{\delta}([\epsilon], x), a) && \text{(definition of } \hat{\delta}) \\
&= \delta([x], a) && \text{(IH)} \\
&= [xa] && \text{(definition of } \delta)
\end{aligned}$$

Hence, we have proven the lemma.

(b) Denote the DFA in part (a) as $D = (Q, \Sigma, \delta, q_0, F)$. Let DFA $A = (Q_A, \Sigma, \delta_A, q_{0_A}, F_A)$ with $L(A) = L$ and no unreachable states. Construct $f : Q_A \to Q$

$$f(\hat{\delta}_A(q_{0_A}, w)) = \hat{\delta}(q_0, w) \text{ for } \forall w \in \Sigma^*$$

f is well defined since

- for each $q \in Q_A$, there exists a $f(q) \in Q$. Let $q \in Q_A$, then $\exists w \in \Sigma^*$ such that $q = \hat{\delta}_A(q_{0_A}, w)$ since q is a reachable state by construction. By def of the constructed f, we have $f(q) = \hat{\delta}(q_0, w) = \hat{\delta}([\epsilon], w) \overset{part(b)}{=} [w] \in Q$
- and each $q \in Q_A$ has a unique mapping $f(q) \in Q$. Suppose $\hat{\delta}_A(q_{0_A}, x) = \hat{\delta}_A(q_{0_A}, y)$ for $x \neq y$, we show $\hat{\delta}(q_0, x) = \hat{\delta}(q_0, y)$. Let $w \in \Sigma^*$. $xw \in L \Leftrightarrow \hat{\delta}_A(q_{0_A}, xw) \in F_A \Leftrightarrow \hat{\delta}_A(\hat{\delta}_A(q_{0_A}, x), w) \in F_A \Leftrightarrow \hat{\delta}_A(\hat{\delta}_A(q_{0_A}, y), w) \in F_A \Leftrightarrow \hat{\delta}_A(q_{0_A}, yw) \in F_A \Leftrightarrow yw \in L$. Hence, $\forall w \in \Sigma^*, xw \in L \Leftrightarrow yw \in L$, i.e. $(x, y) \in R_L \Leftrightarrow [x] = [y] \Leftrightarrow \hat{\delta}(q_0, x) = \hat{\delta}(q_0, y)$

If A has fewer states than D, then f means there exists at least one state $[x] \in D$ such that for $\forall w \in \Sigma^*$, $f(\hat{\delta}_A(q_{0_A}, w)) \neq [x]$. However, for x, $f(\hat{\delta}_A(q_{0_A}, x)) = \hat{\delta}(q_0, x) = [x]$, a contradiction. Therefore, A must have at least as many states as D. For A with unreachable states, it must be more. Hence D is the minimal DFA for L.

Exercise 2.

(a) Pick n=7. We consider $w = a^j c^k a^l b^m \in L$ with $|w| \geq 7$ in two separate cases.

- $0 \leq j \leq 5$, $k \geq 2$ and $l \neq m$
  Divide $w = xyz$ as $x = a^j c$, $y = c$, and $z = c^{k-2} a^l b^m$. $|xy| = j + 2 \leq 7$ since $0 \leq j \leq 5$, $|y| = 1 > 0$, and by pumping y, we get
  - $xy^0 z = a^j c c^{k-2} a^l b^m = a^j c^{k-1} a^l b^m$
    if $k = 2$, then $k - 1 = 1$, $l \neq m$ still means $xy^0 z \in L$
    if $k > 2$, then $k - 1 \geq 2$, $l \neq m$ still holds, and $xy^0 z \in L$
  - $xy^i z = a^j c c^i c^{k-2} a^l b^m = a^j c^{i+k-1} a^l b^m$ for $i > 1$
    $i + k - 1 > 2$ for $k \geq 2$, $l \neq m$ still holds, so $xy^i z \in L$ for $i > 1$

  Hence, all w when $k \geq 2$ satisfies the pumping lemma.

- $0 \leq j \leq 5$, $k < 2$ and $k, l, m \in \mathbb{N}$
  Note that for $w \in L$ with $|w| \geq 7$, w should have at least one of l and m not being zero since $j + k < 7$
  - if $l \neq 0 \Leftrightarrow l \geq 1$, then we divide $w = xyz$ such that $x = a^j c^k$, $y = a$, $z = a^{l-1} b^m$. $|xy| = j + k + 1 \leq 7$, $|y| = 1 > 0$, and $xy^i z = a^j c^k a^i a^{l-1} b^m = a^j c^k a^{i+l-1} b^m$ and $i + l - 1 \geq 0$ for $i \in \mathbb{N}$, hence $xy^i z \in L$ for $\forall i \in \mathbb{N}$
  - if $l = 0$, then $m \neq 0 \Leftrightarrow m \geq 1$. We divide $w = xyz$ such that $x = a^j c^k a^0$, $y = b$, $z = b^{m-1}$. $|xy| = j + k + 1 \leq 7$, $|y| = 1 > 0$, and $xy^i z = a^j c^k a^0 b^i b^{m-1} = a^j c^k a^0 b^{i+m-1}$ and $i + m - 1 \geq 0$ for $i \in \mathbb{N}$, hence $xy^i z \in L$ for $\forall i \in \mathbb{N}$

Hence, we have proven L satisfies the pumping lemma.

(b) We show that R_L has an infinite number of equivalence classes.
Consider $u = a^j c^k a^l b^m$ and $v = a^j c^k a^{l'} b^{m'}$ where $0 \leq j \leq 5$, $k \geq 2$, $l > m$ and $l' > m'$, and $l - m \neq l' - m'$. Now pick $w = b^{l-m}$, $uw = a^j c^k a^l b^m b^{l-m} = a^j c^k a^l b^l \notin L$ since $k \geq 2$ but $l = m$; on the other hand, $vw = a^j c^k a^{l'} b^{m'} b^{l-m} = a^j c^k a^{l'} b^{m'+l-m} \in L$ since $m' + l - m \neq l'$. Hence, $(u, v) \notin R_L \Leftrightarrow [u] \neq [v]$. In other words, for any pair of $(u, v) \in \Sigma^* \times \Sigma^*$ with $k \geq 2$ and distinct positive values of $l - m$, we have $[u] \neq [v]$. Since there is an infinite number of distinct positive values of $l - m$ for $l, m \in \mathbb{N}$, there is an infinite number of distinct equivalence classes in R_L. By the Myhill-Nerode Theorem, L is not regular.

Exercise 3.

(a) We prove a lemma i.e. $f(\hat{\delta}(q, w)) = \hat{\delta}'(f(q), w)$ for $\forall w \in \Sigma^*$, $\forall q \in Q$.
Let $q \in Q$, by an induction on $|w|$,

- **Base case** $w = \epsilon$
  $f(\hat{\delta}(q, \epsilon)) = f(q)$ by the def of $\hat{\delta}$ and $\hat{\delta}'(f(q), \epsilon) = f(q)$ by def of $\hat{\delta}'$, hence $f(\hat{\delta}(q, \epsilon)) = \hat{\delta}'(f(q), \epsilon)$

- **Step case** Assume the claim for $\forall w$ with $|w| < n$. Now we prove for $w = xa$ where $x \in \Sigma^{n-1}$, $a \in \Sigma$

$$\begin{aligned}
f(\hat{\delta}(q, xa)) &= f(\delta(\hat{\delta}(q, x), a)) && \text{(def of } \hat{\delta}) \\
&= \delta'(f(\hat{\delta}(q, x)), a) && \text{(def (3) of } f) \\
&= \delta'(\hat{\delta}'(f(q), x), a) && \text{(IH since } |x| < n) \\
&= \hat{\delta}'(f(q), xa) && \text{(def of } \hat{\delta}')
\end{aligned}$$

Hence, we have proven the lemma.
Now we prove $\mathscr{L}(P, q) = \mathscr{L}(P', f(q))$ for $\forall q \in Q$. Let $q \in Q$, $w \in \Sigma^*$

$$\begin{aligned}
w \in \mathscr{L}(P, q) &\Leftrightarrow \hat{\delta}(q, w) \in F \\
&\Leftrightarrow f(\hat{\delta}(q, w)) \in F' && \text{(def (2) of } f) \\
&\Leftrightarrow \hat{\delta}'(f(q), w) \in F' && \text{(lemma)} \\
&\Leftrightarrow w \in \mathscr{L}(P', f(q))
\end{aligned}$$

Hence the claim.

(b)

$$\begin{aligned}
L(P) &= \{w \in \Sigma^* \mid \hat{\delta}(q_0, w) \in F\} \\
&= \mathscr{L}(P, q_0) && \text{(def of } \mathscr{L}) \\
&= \mathscr{L}(P', f(q_0)) && \text{(part (a))} \\
&= \mathscr{L}(P', q_0') && \text{(def (1) of } f) \\
&= \{w \in \Sigma^* \mid \hat{\delta}'(q_0', w) \in F'\} && \text{(def of } \mathscr{L}) \\
&= L(P')
\end{aligned}$$

Exercise 4.

1. We show validity of $(r^*)^* = r^*$ by proving $L((r^*)^*) = L(r^*)$.
$L((r^*)^*) = L(r^*)^* = (L(r)^*)^* = L(r)^* = L(r^*)$ where the third equaltiy uses the algebraic law of Kleene-*.
Now we prove for $\forall L \subseteq \Sigma^*, (L^*)^* = L^*$. Let $L \subseteq \Sigma^*$,

- $L^* \subseteq (L^*)^*$ since $L^* = (L^*)^1 \subseteq (L^*)^*$
- $(L^*)^* \subseteq L^*$
  Let $w \in (L^*)^*$, we can write $w = w_1 w_2 ... w_n$ for $n \geq 0$ where each $w_i \in L^*$. We can also write each $w_i = w_{i1} w_{i2} ... w_{il_i}$ for $l_i \geq 0$ where each $w_{il_i} \in L$. Then $w = w_{11} w_{12} .. w_{1l_1} ... w_{n1} w_{n2} .. w_{nl_n} \in L^{\sum_1^n l_n} \subseteq L^*$. Hence, $(L^*)^* \subseteq L^*$.

2. We prove $L((r + s)^*) = L((r^* s)^* r^*)$. Denote $R = L(r)$ and $S = L(s)$ for clarity.

- $L((r + s)^*) = L(r + s)^* = (L(r) \cup L(s))^* = (R \cup S)^*$
- $L((r^* s)^* r^*) = L((r^* s)^*) L(r^*) = L(r^* s)^* L(r)^* = (L(r^*) L(s))^* R^* = (L(r)^* S)^* R^* = (R^* S)^* R^*$

In other words, we prove $(R \cup S)^* = (R^*S)^*R^*$. Laws that are needed:

L1. $\forall L \subseteq \Sigma^*$, $L^*L \subseteq L^*$

Proof. Suppose $L \subseteq \Sigma^*$. Let $w \in L^*L$. Then we can write $w = xa$ where $x \in L^*$, $a \in L$. x can be further written as $x_1x_2...x_n$ for $n \geq 0$ where each $x_i \in L$. Then $w = x_1x_2...x_na \in L^{n+1} \subseteq L^*$. Hence, $L^*L \subseteq L^*$.

L2. $\forall L, M, N \subseteq \Sigma^*$, $M \subseteq N \Rightarrow LM \subseteq LN$

Proof. Suppose $M, N \subseteq \Sigma^*$ with $M \subseteq N$. Let $w \in LM$. We can write $w = xa$, where $x \in L, a \in M$. Since $M \subseteq N$, then $a \in N$, and $w = xa \in LN$. Hence, $LM \subseteq LN$.

L3. $\forall M, N \in \Sigma^*$, $M \subseteq N \Rightarrow M^* \subseteq N^*$

Proof. Suppose $M, N \in \Sigma^*$ with $M \subseteq N$. Let $w \in M^*$, then we can write $w = w_1w_2...w_n$ where each $w_i \in M$. Since $M \subseteq N$, hence each $w_i \in N$. Then $w = w_1w_2...w_n \in N^*$. Hence, $M^* \subseteq N^*$.

L4. $\forall M, N, X, Y \in \Sigma^*$, $M \subseteq N \wedge X \subseteq Y \Rightarrow MX \subseteq NY$

Proof. Suppose $M, N, X, Y \in \Sigma^*$ with $M \subseteq N$ and $X \subseteq Y$. Let $w \in MX$, then $w = ax$ where $a \in M$, $x \in X$. Since $M \subseteq N$ and $X \subseteq Y$, then $a \in N$ and $x \in Y$. Then $w = ax \in NY$. Hence, $MX \subseteq NY$.

- $(R \cup S)^* \subseteq (R^*S)^*R^*$

  We prove a lemma that $\forall n \geq 0$, $(R \cup S)^n \subseteq (R^*S)^*R^*$. By an induction on n,

  - **Base case** n = 0. $(R \cup S)^0 = \{\epsilon\} = \{\epsilon\}\{\epsilon\} = (R^*S)^0R^0$. Since $(R^*S)^0 \subseteq (R^*S)^*$ and $R^0 \subseteq R^*$, by L4 $(R^*S)^0R^0 \subseteq (R^*S)^*R^*$. Hence $(R \cup S)^0 \subseteq (R^*S)^*R^*$.
  - **Step case** Assume $(R \cup S)^n \subseteq (R^*S)^*R^*$ and prove $(R \cup S)^{n+1} \subseteq (R^*S)^*R^*$

    Let $w \in (R \cup S)^{n+1} = (R \cup S)^n(R \cup S)$. Then we can write $w = xa$ where $x \in (R \cup S)^n$, $a \in (R \cup S)$. By IH, $x \in (R^*S)^*R^*$. a could be either in R or in S.

    * $a \in R$. Then $w = xa \in ((R^*S)^*R^*)R \overset{assoc}{=} (R^*S)^*(R^*R)$. Prove $(R^*S)^*(R^*R) \subseteq (R^*S)^*R^*$

      Proof. By L1, we have $R^*R \subseteq R^*$. Since $(R^*S)^*$ is a regular language in Σ^* as reg langs are closed under concatenation and Kleene-*, then by L2 we have $(R^*S)^*(R^*R) \subseteq (R^*S)^*R^*$.
    * $a \in S$. Then $w = xa \in ((R^*S)^*R^*)S \overset{assoc}{=} (R^*S)^*(R^*S) \overset{L1}{=} (R^*S)^* = (R^*S)^*\{\epsilon\} = (R^*S)^*R^0$. Since $R^0 \subseteq R^*$, and $(R^*S)^*$ is a reg lang, we have by L2 $(R^*S)^*R^0 \subseteq (R^*S)^*R^*$.

    Thus, $w = xa \in (R^*S)^*R^*$ and we have proven the step case.

  Hence the lemma. Then the set union $\cup_{n \geq 0}(R \cup S)^n = (R \cup S)^* \subseteq (R^*S)^*R^*$.

- $(R^*S)^*R^* \subseteq (R \cup S)^*$

  Since $R \subseteq (R \cup S)$, by L3 we have $R^* \subseteq (R \cup S)^*$. Also, $S \subseteq (R \cup S)$, by L4 $R^*S \subseteq (R \cup S)^*(R \cup S) \overset{L1}{\subseteq} (R \cup S)^*$. By L3 again, $(R^*S)^* \subseteq ((R \cup S)^*)^* \overset{part1}{=} (R \cup S)^*$. By L4, $(R^*S)^*R^* \subseteq (R \cup S)^*(R \cup S)^* = ((R \cup S)^*)^2 \subseteq ((R \cup S)^*)^* \overset{part1}{=} (R \cup S)^*$.

Hence, we have formally proven $L((r+s)^*) = L((r^*s)^*r^*)$, thus the validity of $(r+s)^* = (r^*s)^*r^*$.

3. Let $R = L(r)$, $S = L(s)$.

- $L((rs)^*) = L(rs)^* = (L(r)L(s))^* = (RS)^*$
- $L(\epsilon + r(sr)^*s) = L(\epsilon) \cup L(r(sr)^*s) = \{\epsilon\} \cup L(r(sr)^*)L(s) = \{\epsilon\} \cup L(r)L((sr)^*)S = \{\epsilon\} \cup RL(sr)^*S = \{\epsilon\} \cup R(L(s)L(r))^*S = \{\epsilon\} \cup R(SR)^*S$

We prove $(RS)^* = \{\epsilon\} \cup R(SR)^*S$ by equaltiy of the subsets, i.e, $(RS)^0 = \{\epsilon\}$ and $(RS)^n = R(SR)^{n-1}S$ for $\forall n \geq 1$.

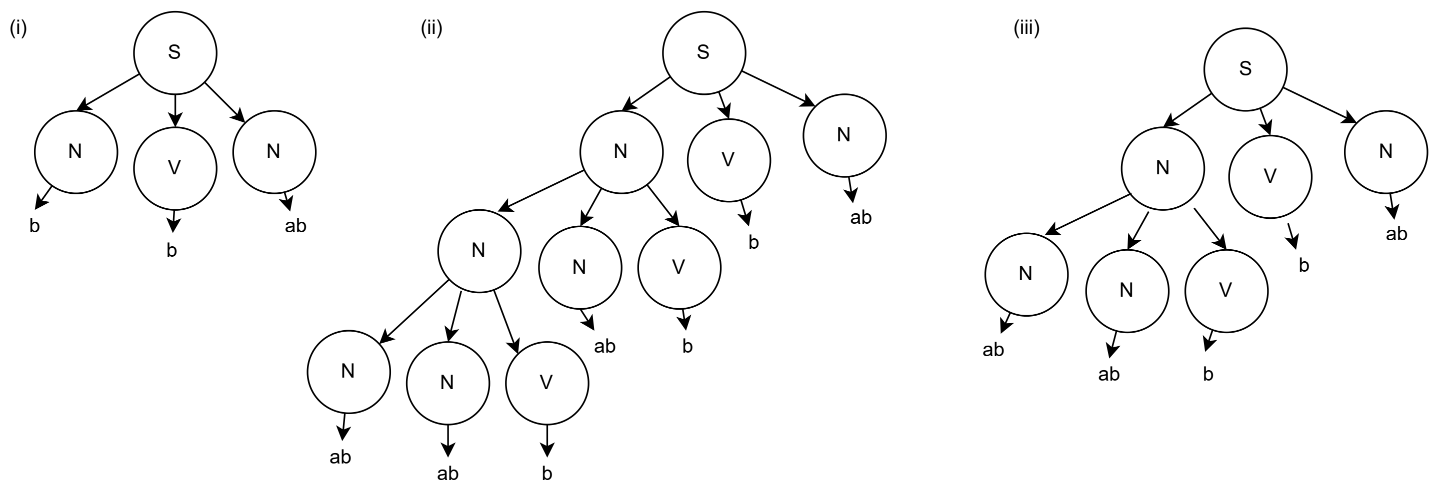
- $(RS)^0 = \{\epsilon\}$ by definition
- we prove by an induction on n that $(RS)^n = R(SR)^{n-1}S$ for $\forall n \geq 1$.
  - **Base case** n = 1. $(RS)^1 = RS = (R\{\epsilon\})S = (R(SR)^0)S = R(SR)^0S$
  - **Step case** $(RS)^{n+1} = (RS)^n(RS) \overset{IH}{=} (R(SR)^{n-1}S)(RS) \overset{assoc}{=} ((R(SR)^{n-1}S)R)S$ $\overset{assoc}{=} (R(SR)^{n-1}(SR))S = (R(SR)^n)S = R(SR)^nS$

  Hence $\cup_{n\geq 1}(RS)^n = \cup_{n\geq 1}R(SR)^{n-1}S = \cup_{n\geq 0}R(SR)^nS$

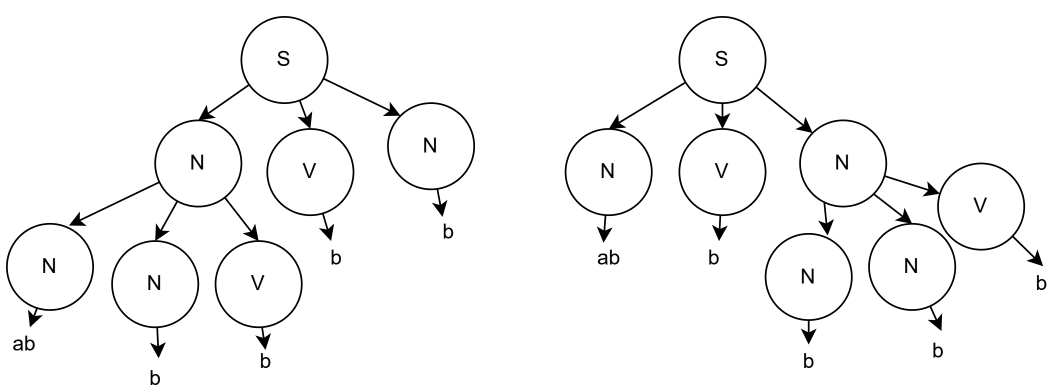
From above, $(RS)^* = \cup_{n\geq 0}(RS)^n = (RS)^0 \cup (\cup_{n\geq 1}(RS)^n) = \{\epsilon\} \cup (\cup_{n\geq 0}R(SR)^nS) = \{\epsilon\} \cup R(SR)^*S$. Hence, $(rs)^* = \epsilon + r(sr)^*s$ is valid.

Exercise 5.

(a) As below



(b) G is ambiguous because there exists $w \in L(G)$ that has two parse trees. Consider $w = abbbbb \in L(G)$. The following two parse trees generate the same string $abbbbb$.



(c) G' is defined by the following productions:

$$
\begin{aligned}
S &\to AVN \\
A &\to ANV \mid ab \mid a \\
N &\to NNV \mid ab \mid b \\
V &\to b
\end{aligned}
$$