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 state. Construct $f : Q_A \to Q$

$$f(\hat{\delta}_A(q_{0_A}, w)) = \hat{\delta}(q_0, w) \text{ for } \forall w \in \Sigma^* \tag{1}$$

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, $f(\hat{\delta}_A(q_{0_A}, x)) = \hat{\delta}(q_0, x) = [x]$, this is 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.

We prove that any DFA $A = (Q_A, \Sigma, \delta_A, q_{0_A}, F_A)$ with $L(A) = L$ has at least as many states as D, by constructing a surjective function f from Q_A to Q. We adapted the proof from solutions of Tutorial 2. Define $S(q) = \{w \in \Sigma^* \mid \hat{\delta}(q_{0_A}, w) = q\}$, and f

$$\begin{aligned} &f : \{q \in Q_A \mid S(q) \neq \emptyset\} \to \{[x] \mid x \in \Sigma^*\} \\ &f(q) = [x] \text{ with } S(q) \subseteq [x] \end{aligned}$$

We show that f is well defined, i.e.

- $\forall q \in Q_A$ such that $S(q) \neq \emptyset$, $\exists [x] \in Q$ with $S(q) \subseteq [x]$
  Proof. For $q \in Q_A$, pick $[x] \in Q$ such that $x \in S(q)$. Now for $\forall u, v \in S(q)$, we proved in the tutorial that $(u, v) \in R_L$. Hence for $\forall u \in S(q)$, $(x, u) \in R_L$. Hence, $u \in [x]$, and $S(q) \subseteq [x]$

- if $f(q) = [x]$ and $f(q) = [y]$ for $x \neq y$, then $[x] = [y]$
  Proof. Suppose $f(q) = [x]$ and $f(q) = [y]$, then $S(q) \subseteq [x]$ and $S(q) \subseteq [y]$. Then $u \in S(q) \Rightarrow u \in [x] \Leftrightarrow (x, u) \in R_L$; and $u \in S(q) \Rightarrow u \in [y] \Leftrightarrow (y, u) \in R_L \Leftrightarrow (u, y) \in R_L$. By transitivity, $(x, u) \in R_L \wedge (u, y) \in R_L \Rightarrow (x, y) \in R_L \Leftrightarrow [x] = [y]$

Now, we show that f is surjective. For $[x] \in Q$, we pick $q = \hat{\delta}(q_{0_A}, x) \in Q_A$. $S(q) \neq \emptyset$ since $x \in S(q)$ by definition. We only need to show $S(q) \subseteq [x]$. By the same argument that $\forall u, v \in S(q), (u, v) \in R_L$. $u \in S(q) \Rightarrow (x, u) \in R_L \Leftrightarrow u \in [x]$.
Since f is surjective, then for every state $[x]$ in DFA D, there are corresponding reachable states in DFA A that are mapped to $[x]$, and the cardinality of Q_A exclusive of unreachable states is at least the cardinality of Q. Hence, any DFA that accepts L would have at least as many states as DFA D. Then D is the minimal DFA.

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 $l' > m'$, $l - m \neq l' - m'$. Since $k \geq 2 \wedge l \neq m \wedge l' \neq m'$, by definition $u, v \in L$. Now pick $w = b^{l-m}$, $uw = a^j c^k a^l b^m b^{l-m} = a^j c^k a^l b^l \notin L$, but $vw = a^j c^k a^{l'} b^{m'} b^{l-m} = a^j c^k a^{l'} b^{m'+l-m} \in L$ since $m' + l - m \neq l'$. Hence, $(u, v) \notin R_L \Leftrightarrow [u] \neq [v]$. In other words, for any pair of $(u, v) \in L$ 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 $w_i = x_{i1} x_{i2} ... x_{il_i}$ for $l_i \geq 0$ where each $x_{il_i} \in L$. Then $w = x_{11} x_{12} ... x_{1l_1} .. x_{n1} x_{n2} ... x_{nl_n} \in L^{\sum_1^n l_n} \subseteq L^*$. Hence, $w \in L^*$.

2. We prove $L((r+s)^*) = L((r^*s)^*r^*)$. Denote $R = L(r)$ and $S = L(s)$ for clarity. By def of reg lang,
$L((r+s)^*) = L(r+s)^* = (L(r) \cup L(s))^* = (R \cup S)^*$
$L((r^*s)^*r^*) = L((r^*s)^*)L(r^*) = L(r^*s)^*L(r)^* = (L(r^*)L(s))^*R^* = (L(r)^*S)^*R^* = (R^*S)^*R^*$
In other words, we prove $(R \cup S)^* = (R^*S)^*R^*$.

- $(R \cup S)^* \subseteq (R^*S)^*R^*$
  We prove the lemma that for $\forall n \geq 0$, $(R \cup S)^n \subseteq (R^*S)^*R^*$, then by def of set union, $(R \cup S)^* = \cup_{n \geq 0}(R \cup S)^n \subseteq (R^*S)^*R^*$. By an induction on n,
  - n $= 0$ $(R \cup S)^0 = \{\epsilon\} = \{\epsilon\}\{\epsilon\} = (R^*S)^0 R^0 \subseteq (R^*S)^*R^*$
  - prove $(R \cup S)^{n+1} \subseteq (R^*S)^*R^*$

$$\begin{aligned}
(R \cup S)^{n+1} &= (R \cup S)^n (R \cup S) && \text{(concat of lang)} \\
&= \{xa \in \Sigma^* \mid x \in (R \cup S)^n, a \in (R \cup S)\} && \text{(set notation of concat)} \\
&\subseteq \{xa \in \Sigma^* \mid x \in (R^*S)^*R^*, a \in (R \cup S)\} && \text{(IH)} \\
&= \{xa \in \Sigma^* \mid x \in (R^*S)^*R^*, a \in R\} \cup \{xa \in \Sigma^* \mid x \in (R^*S)^*R^*, a \in S\} \\
&= ((R^*S)^*R^*)R \cup ((R^*S)^*R^*)S && \text{(set def of concat)}
\end{aligned}$$

  We show both subsets are in $(R^*S)^*R^*$. But first we prove some laws.
  L1. $\forall L \subseteq \Sigma^*$, $L^*L \subseteq L^*$
  Proof. $L^*L = (\cup_{n \geq 0} L^n)L = \{xa \in \Sigma^* \mid x \in \cup_{n \geq 0} L^n, a \in L\} = \cup_{n \geq 0}\{xa \in \Sigma^* \mid x \in L^n, a \in L\} = \cup_{n \geq 0}(L^n L) = \cup_{n \geq 0} L^{n+1} = \cup_{n \geq 1} L^n \subseteq L^*$
  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 $w \in LM$. Write $w = xa$, where $x \in L, a \in M$. Since $M \subseteq N$, $a \in N$. Then $w = xa \in LN$. In short, $LM = \{xa \in \Sigma^* \mid x \in L, a \in M\} \subseteq \{xa \in \Sigma^* \mid x \in L, a \in N\} = LN$
  Back to the proof
    * $((R^*S)^*R^*)R \subseteq (R^*S)^*R^*$

$$\begin{aligned}
& R^*R \subseteq R^* && \text{(L1)} \\
& \Rightarrow (R^*S)^*(R^*R) \subseteq (R^*S)^*R^* && \text{(L2)} \\
& \Leftrightarrow ((R^*S)^*R^*)R \subseteq (R^*S)^*R^* && \text{(associativity)}
\end{aligned}$$

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

$$(R^*S)^*(R^*S) \subseteq (R^*S)^* \tag{L1}$$
$$\Leftrightarrow ((R^*S)^*R^*)S \subseteq (R^*S)^* = (R^*S)^*\{\epsilon\} = (R^*S)^*R^0 \subseteq (R^*S)^*R^*$$
(assoc & concat with $\{\epsilon\}$)

Hence the step case.

Hence the lemma.

- $(R^*S)^*R^* \subseteq (R \cup S)^*$
L3. for $\forall M, N \in \Sigma^*$, $M \subseteq N \Rightarrow M^* \subseteq N^*$
Proof. Suppose $M \subseteq N$. 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. for $\forall M, N, X, Y \in \Sigma^*$, $M \subseteq N \wedge X \subseteq Y \Rightarrow MX \subseteq NY$
Proof. Suppose $M \subseteq N \wedge X \subseteq Y$. 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 $x \in Y$. Hence, $w \in NY$. $MX \subseteq NY$.
Now we can prove $(R^*S)^*R^* \subseteq (R \cup S)^*$.
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 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.