One of the problems on Guided Problem Set 4 turns out to have multiple correct solutions, based on subtly different interpretations of the states of the underlying NFA.

Prove that for every regular language \( L \), the language \( \text{INSERT}01(L) = \{ x01y \mid xy \in L \} \) is regular.

Here I'll describe three distinct solutions. Let me emphasize that this is not an exhaustive list, and that one solution is not necessarily a refinement of any other. For each solution, I'll show the NFA constructed for the following DFA, which accepts the single-string language \( L = \{0100\} \):
Solution (guess early): Let $M = (Q, s, A, \delta)$ be an arbitrary DFA that accepts $L$. We construct an NFA $M' = (Q', s', A', \delta')$ for $\text{INSERT01}(L)$ as follows:

$$Q' = Q \times \{\varepsilon, 0, 01\}$$

$$s' = (s, \varepsilon)$$

$$A' = A \times \{01\}$$

$$\delta'((q, \varepsilon), \theta) = \{(\delta(q, \theta), \varepsilon), (q, \theta)\}$$

$$\delta'((q, 0), \theta) = \emptyset$$

$$\delta'((q, 01), \theta) = \{(\delta(q, 0), 01)\}$$

$$\delta'((q, 01), 1) = \{(\delta(q, 1), 01)\}$$

The states $(q, x)$ have the following interpretations:

- $(q, \varepsilon)$ means $M$ is in state $q$ and $M'$ has not read any part of the inserted $01$.
- $(q, 0)$ means $M$ is in state $q$ and $M'$ has just read the $0$ from the inserted $01$.
- $(q, 01)$ means $M$ is in state $q$ and $M'$ has already read and skipped the inserted $01$.

When $M'$ is in state $(q, \varepsilon)$ and reads $\theta$, $M'$ non-deterministically guesses whether that $\theta$ is part of the inserted $01$.

- If the $\theta$ is not part of the inserted $01$, then $M'$ passes the $\theta$ to $M$, and $M'$ has still not read any part of the inserted $01$. Thus, the new state of $M'$ is $(\delta(q, \theta), \varepsilon)$.
- On the other hand, if the $\theta$ is part of the inserted $01$, then $M'$ does not pass the $\theta$ to $M$, and $M'$ has just read the $\theta$ from the inserted $01$. Thus, the new state of $M'$ is $(q, \theta)$.

Similarly, when $M'$ is in state $(q, 0)$ and reads $\theta$, $M'$ realizes that incorrectly guessed that the previous $\theta$ was part of the inserted $01$, so it destroys the universe.

Here is the NFA that results from applying this transformation to the example DFA for $L = \{0100\}$. For readability, I've omitted transitions to the dump states $(x, \varepsilon), (x, 0)$, and $(x, 01)$. The unique accepting path for the input string $010010 \in \text{INSERT01}(L)$ is highlighted in green.
**Solution (guess late):** Let \( M = (Q, s, A, \delta) \) be an arbitrary DFA that accepts \( L \). We construct an NFA \( M' = (Q', s', A', \delta') \) as follows:

\[
Q' = Q \times \{ \varepsilon, 0, 01 \} \\
A' = A \times \{ 01 \} \\
\delta'((q, \varepsilon), 0) = \{(q, 0)\} \\
\delta'((q, \varepsilon), 1) = \{(\delta(q, 1), \varepsilon)\} \\
\delta'((q, 0), 0) = \{(\delta(q, 0), \varepsilon)\} \\
\delta'((q, 0), 1) = \{(\delta(q(0), 1), 1, \epsilon), (q, 01)\} \\
\delta'((q, 01), 0) = \{(\delta(q, 0), 01)\} \\
\delta'((q, 01), 1) = \{(\delta(q, 1), 01)\}
\]

The states \((q, x)\) have the following interpretations:

- \((q, \varepsilon)\) means \( M \) is in state \( q \), \( M' \) has not read the inserted \( \varepsilon \), and \( M' \) did not just read a \( \varepsilon \).
- \((q, 0)\) means \( M \) is in state \( q \), \( M' \) has not read the entire inserted \( 01 \), and \( M' \) just read a \( 0 \) which may or may not be from the the inserted \( 01 \).
- \((q, 01)\) means \( M \) is in state \( q \) and \( M' \) has already read and skipped the inserted \( 01 \).

When \( M' \) is in state \((q, \varepsilon)\) and reads \( \varepsilon \), \( M' \) remembers the \( \varepsilon \) but does not pass it to \( M' \) (yet).

When \( M' \) is in state \((q, 0)\) and reads \( \varepsilon \), \( M' \) knows that the previous \( \varepsilon \) was not part of the inserted \( 01 \), so it passed the previous \( \varepsilon \) to \( M \) and remembers the new \( \varepsilon \).

When \( M' \) is in state \((q, 0)\) and reads \( 1 \), \( M' \) non-deterministically guesses whether the previous \( \varepsilon \) and this \( 1 \) are the inserted \( 01 \).

- If the \( 1 \) is part of the inserted \( 01 \), then \( M' \) remembers that it has read the inserted \( 01 \), but it does not pass anything to \( M \). So the new state of \( M' \) is \((q, 01)\).
- If the \( 1 \) is not part of the inserted \( 01 \), \( M' \) passes the previous \( \varepsilon \) to \( M \) and then passes the new \( 1 \) to \( M \). In this case, \( M' \) has not read the entire inserted \( 01 \), and \( M' \) did not just read \( 1 \). So the new state of \( M' \) is \((\delta(\delta(q, 0), 1), \varepsilon)\).

Here is the NFA that results from applying this transformation to the example DFA for \( L = \{0100\} \). For readability, I’ve omitted transitions to the dump states \((x, \varepsilon)\), \((x, 0)\), and \((x, 01)\). The unique accepting path for the input string \(010010 \in \text{INSERT}-\!01(L)\) is highlighted in green.
**Solution (FOMO):** Let $M = (Q, s, A, \delta)$ be an arbitrary DFA that accepts $L$. We construct an NFA $M' = (Q', s', A', \delta')$ as follows:

$$Q' = Q \times \{\varepsilon, 0, 01\}$$
$$s' = (s, \varepsilon)$$
$$A' = A \times \{01\}$$

$$\delta'((q, \varepsilon), \emptyset) = \{(\delta(q, \emptyset), \varepsilon), (q, 0)\}$$
$$\delta'((q, \varepsilon), 1) = \{(\delta(q, 1), \varepsilon)\}$$

$$\delta'((q, 0), \emptyset) = \{(\delta(\delta(q, \emptyset), \emptyset), \varepsilon), (\delta(q, \emptyset), 0)\}$$
$$\delta'((q, 0), 1) = \{(\delta(\delta(q, \emptyset), 1), \varepsilon), (q, 01)\}$$

$$\delta'((q, 01), \emptyset) = \{(\delta(q, \emptyset), 01)\}$$
$$\delta'((q, 01), 1) = \{(\delta(q, 1), 01)\}$$

The states $(q, x)$ have the following interpretations:

- $(q, \varepsilon)$ means $M$ is in state $q$, $M'$ has not read any part of the inserted $01$, and $M'$ may or may not have just read a $\emptyset$.
- $(q, 0)$ means $M$ is in state $q$, $M'$ has not read the entire inserted $01$, and $M'$ just read a $\emptyset$ which may or may not be from the the inserted $01$.
- $(q, 01)$ means $M$ is in state $q$ and $M'$ has already read and skipped the inserted $01$.

This machine always hedges its bets. From any state and given any symbol, $M'$ transitions to all possible states that are logically consistent with what it knows. In particular, this solution includes every transition from the two previous solutions, but it includes some additional transitions as well.

For example, when $M'$ is in state $(q, 0)$ and reads $\emptyset$, $M'$ non-deterministically guesses which of two possibilities is the truth:

- If the new $\emptyset$ is definitely not part of the inserted $01$, then $M'$ passes the previous $\emptyset$ to $M$, and then passes the new $\emptyset$ to $M$. In this case, $M'$ has not read any part of the inserted $01$. So the new state of $M'$ is $(\delta(\delta(q, \emptyset), \emptyset), \varepsilon)$.
- If the new $\emptyset$ might be part of the inserted $01$, then $M'$ passes the previous $\emptyset$ to $M$ and remembers the new $\emptyset$. So the new state of $M'$ is $(\delta(q, \emptyset), 0)$.

Here is the NFA that results from applying this transformation to the example DFA for $L = \{0100\}$. For readability, I've omitted transitions to the dump states $(x, \varepsilon)$, $(x, 0)$, and $(x, 01)$. One of the four accepting paths for the input string $010010 \in \text{INSERT}01(L)$ is highlighted in green.