The following problems ask you to prove some "obvious" claims about recursively-defined string functions. In each case, we want a self-contained, step-by-step induction proof that builds on formal definitions and prior results, *not* on intuition. In particular, your proofs must refer to the formal recursive definitions of string length and string concatenation:

$$|w| = \begin{cases} 0 & \text{if } w = \varepsilon \\ 1 + |x| & \text{if } w = ax \text{ for some symbol } a \text{ and some string } x \end{cases}$$

$$w \bullet z := \begin{cases} z & \text{if } w = \varepsilon \\ a \cdot (x \bullet z) & \text{if } w = ax \text{ for some symbol } a \text{ and some string } x \end{cases}$$

You may freely use the following results, which were proved in the lecture notes:

Lemma 1: $w \bullet \varepsilon = w$ for all strings w.

Lemma 2: $|w \bullet x| = |w| + |x|$ for all strings w and x.

Lemma 3: $(w \bullet x) \bullet y = w \bullet (x \bullet y)$ for all strings w, x, and y.

The **reversal** w^R of a string w is defined recursively as follows:

$$w^R := \begin{cases} \varepsilon & \text{if } w = \varepsilon \\ x^R \bullet a & \text{if } w = ax \text{ for some symbol } a \text{ and some string } x \end{cases}$$

For example, $STRESSED^R = DESSERTS$ and $WTF374^R = 473FTW$.

1 Prove that $|w^R| = |w|$ for every string w.

Solution:

We first provide a painfully detailed proof – when you usually write the proof it does not have to be as detailed, but make sure you could write it in this level of detail if you have to – if you can not, then the proof is probably wrong.

Proof:

Statement of what we are proving, and what is the induction done over: We prove that for any string w, such that $n = |w|$, we have that $|w^R| = |w|$. The proof is by induction on the length of w.

Base case: If $n = 0$ then $w = \varepsilon$, then $w^R = \varepsilon^R = \varepsilon$ (by the definition of $\cdot^R$), and $w^R = \varepsilon = w$. In particular, this implies that $|w^R| = 0 = |w|$ (by the definition of $|\cdot|$), as claimed.

Induction hypothesis: Let k be any non-negative integer number. Assume that for any $n \leq k$, and any word w with $n = |w|$, we have $|w^R| = |w|$.

Induction step: We next prove the claim for any word w of length $n = k + 1$.

We can write the given w as $w = ax$ for some symbol a and some string x, where $|x| = k$. We have

$$
\begin{aligned}
\left|w^R\right| &= \left|x^R \bullet a\right| && \text{by definition of } w^R \\
&= \left|x^R\right| + |a| && \text{by Lemma 2} \\
&= \left|x^R\right| + 1 && \text{by definition of } |\cdot| \text{ (twice)} \\
&= |x| + 1 && \text{by the induction hypothesis} \\
&= |w|, && \text{by definition of } |\cdot|.
\end{aligned}
$$

We conclude that $\left|w^R\right| = |w|$, as claimed. ∎

Here is the same proof written in a more natural level of detail.

Proof: The proof is by induction on the length of $|w|$. If $|w| = 0$ then $w = \varepsilon$, then $w^R = \varepsilon^R = \varepsilon$ (by the definition of $\cdot^R$), and $w^R = \varepsilon = w$. In particular, this implies that $|w^R| = 0 = |w|$ (by the definition of $|\cdot|$), as claimed.

Assume that the claim holds for all strings w of length at most k, for some non-negative integer k. And let w be any word of length $k + 1$. We can write w as $w = ax$ for some symbol a and some string x, where $|x| = k$. As such, we have

$$
\begin{aligned}
\left|w^R\right| &= \left|x^R \bullet a\right| && \text{by definition of } w^R \\
&= \left|x^R\right| + 1 && \text{by Lemma 2} \\
&= \left|x^R\right| + 1 && \text{by definition of } |\cdot| \text{ (twice)} \\
&= |x| + 1 && \text{by the induction hypothesis} \\
&= |w| && \text{by definition of } |\cdot|.
\end{aligned}
$$

We conclude that $\left|w^R\right| = |w|$, as claimed. ∎

2 Prove that $(w \bullet z)^R = z^R \bullet w^R$ for all strings w and z.

Solution:

Proof: The proof is by induction on the length of w.

Base case: If $w = \varepsilon$, then

$$
\begin{aligned}
(w \bullet z)^R &= z^R && \text{by definition of } \bullet \\
&= z^R \bullet \varepsilon && \text{by Lemma 1} \\
&= z^R \bullet w^R && \text{by definition of } {}^R,
\end{aligned}
$$

as claimed.

Inductive hypothesis: Let w and z be arbitrary strings, and assume that for any string x where $|x| < |w|$ we have that $(x \bullet z)^R = z^R \bullet x^R$.

Inductive step: Now, $w = ax$ for some symbol a and some string x. As such, we have

$$
\begin{aligned}
(w \bullet z)^R &= (a \cdot (x \bullet z))^R & \text{by definition of } \bullet \\
&= (x \bullet z)^R \bullet a & \text{by definition of } {}^R \\
&= (z^R \bullet x^R) \bullet a & \text{by the induction hypothesis, because } |x| < |w| \\
&= z^R \bullet (x^R \bullet a) & \text{by Lemma 3} \\
&= z^R \bullet w^R & \text{by definition of } {}^R,
\end{aligned}
$$

as claimed.   ∎

> But how did I know that the induction hypothesis needs to change the first string w, but not the second string z? I wrote down the inductive argument first, and then noticed that in the proof for $w \bullet z$, we needed the inductive hypothesis on $x \bullet z$. Same string z, but w changed to x. Alternatively, in light of Lemma 2, I could have inducted on the **sum** of the string lengths with the inductive hypothesis "Assume for all strings x and y such that $|x| + |y| < |w| + |z|$ that $(x \bullet y)^R = y^R \bullet x^R$."

3 Prove that $(w^R)^R = w$ for every string w.

<u>**Solution:**</u>

Proof by induction on the length of w.

Let w be an arbitrary string.

Assume for any string w where $|x| < |w|$ that $(x^R)^R = x$.

There are two cases to consider.

- If $w = \varepsilon$, then $(w^R)^R = \varepsilon^R = \varepsilon$ by definition.
- Otherwise, $w = ax$ for some symbol a and some string x.

$$
\begin{aligned}
(w^R)^R &= (x^R \bullet a)^R & \text{by definition of } {}^R \\
&= a^R \bullet (x^R)^R & \text{by problem 2} \\
&= a \bullet (x^R)^R & \text{by definition of } {}^R \\
&= a \cdot (x^R)^R & \text{by definition of } \bullet \\
&= a \cdot x & \text{by the induction hypothesis} \\
&= w & \text{by assumption}
\end{aligned}
$$

In both cases, we conclude that $(w^R)^R = w$.

Remark: The careful reader would observe that the above proof is written in somewhat confusing fashion. For example, he base case is proved after the inductive hypothesis is stated. As stated above, this proof can be unmangled and written in the tedious explicit way as done in the first proof.

To think about later: Let $\#(a, w)$ denote the number of times symbol a appears in string w. For example, $\#(X, WTF374) = 0$ and $\#(0, 000010101010010100) = 12$.

4 Give a formal recursive definition of $\#(a, w)$.

Solution:

$$\#(a, w) = \begin{cases} 0 & \text{if } w = \varepsilon \\ 1 + \#(a, x) & \text{if } w = ax \text{ for some string } x \\ \#(a, x) & \text{if } w = bx \text{ for some symbol } b \neq a \text{ and some string } x \end{cases}$$

5 Prove that $\#(a, w \bullet z) = \#(a, w) + \#(a, z)$ for all symbols a and all strings w and z.

Solution:

(Induction on w)

Let a be an arbitrary symbol, and let w and z be arbitrary strings.

Assume for any string x such that $|x| < |w|$ that $\#(a, x \bullet z) = \#(a, x) + \#(a, z)$. There are three cases to consider.

- If $w = \varepsilon$, then

$$\begin{aligned} \#(a, \varepsilon \bullet x) &= \#(a, x) & \text{by definition of } \bullet \\ &= \#(a, w) + \#(a, x) & \text{by definition of } \# \end{aligned}$$

- If $w = ax$ for some string x, then

$$\begin{aligned} \#(a, w \bullet z) &= \#(a, ax \bullet z) & \text{by definition of } \bullet \\ &= \#(a, a \cdot (x \bullet z)) & \text{by definition of } \bullet \\ &= 1 + \#(a, x \bullet z) & \text{by definition of } \# \\ &= 1 + \#(a, x) + \#(a, z) & \text{by the induction hypothesis} \\ &= \#(a, ax) + \#(a, z) & \text{by definition of } \# \\ &= \#(a, w) + \#(a, z) & \text{because } w = ax \end{aligned}$$

- If $w = bx$ for some symbol $b \neq a$ and some string x, then

$$\begin{aligned} \#(a, w \bullet z) &= \#(a, b \cdot (x \bullet z)) & \text{by definition of } \bullet \\ &= \#(a, x \bullet z) & \text{by definition of } \# \\ &= \#(a, x) + \#(a, z) & \text{by the induction hypothesis} \\ &= \#(a, bx) + \#(a, z) & \text{by definition of } \# \\ &= \#(a, w) + \#(a, z) & \text{because } w = bx \end{aligned}$$

In every case, we conclude that $\#(a, w \bullet z) = \#(a, w) + \#(a, z)$.

6 Prove that $\#(a, w^R) = \#(a, w)$ for all symbols a and all strings w.

Solution:

(**Induction on** w) Let a be an arbitrary symbol, and let w be an arbitrary string.

Assume for any string x such that $|x| < |w|$ that $\#(a, x^R) = \#(a, x)$.

There are three cases to consider.

- If $w = \varepsilon$, then $w^R = \varepsilon = w$ by definition, so $\#(a, w^R) = \#(a, w)$.
- If $w = ax$ for some string x, then

$$
\begin{aligned}
\#(a, w^R) &= \#(a, x^R \bullet a) && \text{by definition of } ^R \\
&= \#(a, x^R) + \#(a, a) && \text{by problem 5} \\
&= \#(a, x^R) + 1 && \text{by definition of } \# \\
&= \#(a, x) + 1 && \text{by the induction hypothesis} \\
&= \#(a, w) && \text{by definition of } \#
\end{aligned}
$$

- If $w = bx$ for some symbol $b \neq a$ and some string x, then

$$
\begin{aligned}
\#(a, w^R) &= \#(a, x^R \bullet b) && \text{by definition of } ^R \\
&= \#(a, x^R) + \#(a, b) && \text{by problem 5} \\
&= \#(a, x^R) && \text{by definition of } \# \\
&= \#(a, x) && \text{by the induction hypothesis} \\
&= \#(a, w) && \text{by definition of } \#
\end{aligned}
$$

In every case, we conclude that $\#(a, w^R) = \#(a, w)$.