Staff Solutions to Problem Set 4

Reading:
- Section 5.4. State Machines: Invariants
- Chapter 6. Recursive Data Types

Problem 1.
A robot moves on the two-dimensional integer grid. It starts out at $(0, 0)$ and is allowed to move in any of these four ways:

1. $(+2, -1)$: right 2, down 1
2. $(-2, +1)$: left 2, up 1
3. $(+1, +3)$
4. $(-1, -3)$

Prove that this robot can never reach $(1, 1)$.

Solution. Solution 1: Denote the robot’s position by $(x, y)$. Observe that the quantity $z = x + 2y$ does not change when the robot undergoes an atomic movement of type 1 or type 2. A type 3 movement increases $z$ by $1 + 2(3) = 7$, while a type 4 movement decreases $z$ by 7. Thus, the predicate $P(x, y) := 7 \mid x + 2y$

is a preserved invariant for the robot. $P(0, 0)$, and the robot starts at the origin, so by the Invariant Principle, $P$ must hold for all reachable states $(x, y)$ of the robot. Since $P(1, 1)$ is false, the robot cannot reach $(1, 1)$.

For any function, $f$, of the robot’s position, $(x, y)$, let $\Delta f_k(x, y)$ denote the change in $f$ that would result from an atomic movement of type $k$ starting at $(x, y)$. We started to formulate our invariant by defining a function $z$ such that for some $k_0$, $\Delta z_{k_0}(x, y)$ would be zero regardless of the state $(x, y)$. In this way we guaranteed that, as long as all the possible values of $\Delta z_k(x, y)$ for $k \neq k_0$ were not relatively prime, we would be able to formulate a simple preserved invariant based on the divisibility of $z - z(0, 0)$. We developed $P$ by taking $k_0 = 1$ or $k_0 = 2$ and defining $z = x + 2y$. We could just as well have taken $k_0 = 3$ or $k_0 = 4$ instead, defined $z = 3x - y$, and constructed the equally valid and useful preserved invariant $Q(x, y) := 7 \mid 3x - y$.

Solution 2: If the robot starts at the origin and makes $a_k$ atomic movements of type $k$, $k \in \{1, 2, 3, 4\}$, its position will be

$$(2a_1 - 2a_2 + a_3 - a_4, -a_1 + a_2 + 3a_3 - 3a_4)$$

Letting $b_1 = a_1 - a_2$ and $b_2 = a_3 - a_4$, we have $(x, y) = (2b_1 + b_2, -b_1 + 3b_2)$, where $b_1$ and $b_2$ must be integers. Now, note that $x + 2y = 7b_2$. Thus, the robot’s position $(x, y)$ must always be such that
7 | x + 2y. (A preserved invariant true wherever the robot goes!) Since this requirement is not met with 
(x, y) = (1, 1), therefore the robot cannot reach (1, 1).

(Alternatively, we have 3x − y = 7b1, so 7 | 3x − y must also hold for any location of the robot, yet it
does not hold for (x, y) = (1, 1).)

Solution 3: From Solution 2, the robot’s position must always be

\[(2a_1 - 2a_2 + a_3 - a_4, -a_1 + a_2 + 3a_3 - 3a_4)\]

for some \(a_1, a_2, a_3, a_4 \in \mathbb{N}\). Now, the system

\[
\begin{align*}
2a_1 - 2a_2 + a_3 - a_4 &= 1 \\
-a_1 + a_2 + 3a_3 - 3a_4 &= 1
\end{align*}
\]

has no solutions \((a_1, a_2, a_3, a_4) \in \mathbb{N}^4\). (It has solutions in \(\mathbb{R}^4\) given by \((a_1, a_2, a_3, a_4) = (\frac{3}{7}s + s, s, \frac{3}{7} + t, t)\),
for \(s, t \in \mathbb{R}\).) Thus the robot, using only the four permissible atomic movements, cannot possibly reach
\((1, 1)\).

\[\blacksquare\]

Problem 2.

Let \(L\) be some convenient set whose elements will be called labels. The labeled binary trees, LBT’s, are
defined recursively as follows:

Definition. Base case: if \(l\) is a label, then \(\langle l, \text{leaf} \rangle\) is an LBT, and

Constructor case: if \(B\) and \(C\) are LBT’s, then \(\langle l, B, C \rangle\) is an LBT.

The leaf-labels and internal-labels of an LBT are defined recursively in the obvious way:

Definition. Base case: The set of leaf-labels of the LBT \(\langle l, \text{leaf} \rangle\) is \(\{l\}\), and its set of internal-labels is the empty set.

Constructor case: The set of leaf labels of the LBT \(\langle l, B, C \rangle\) is the union of the leaf-labels of \(B\) and of \(C\);
the set of internal-labels is the union of \(\{l\}\) and the sets of internal-labels of \(B\) and of \(C\).

The set of labels of an LBT is the union of its leaf- and internal-labels.

The LBT’s with unique labels are also defined recursively:

Definition. Base case: The LBT \(\langle l, \text{leaf} \rangle\) has unique labels.

Constructor case: If \(B\) and \(C\) are LBT’s with unique labels, no label of \(B\) is a label \(C\) and vice-versa, and
\(l\) is not a label of \(B\) or \(C\), then \(\langle l, B, C \rangle\) has unique labels.

If \(B\) is an LBT, let \(n_B\) be the number of distinct internal-labels appearing in \(B\) and \(f_B\) be the number of
distinct leaf labels of \(B\). Prove by structural induction that

\[f_B = n_B + 1\]  \hspace{1cm} (1)

for all LBT’s \(B\) with unique labels. This equation can obviously fail if labels are not unique, so your proof
had better use uniqueness of labels at some point; be sure to indicate where.

Solution. Base case: If \(B = \langle l, \text{leaf} \rangle\), then \(f_B = 1\) and \(n_B = 0\) so (1) holds.

Constructor case: If \(A := \langle l, B, C \rangle\) has unique labels, then by definition no label appears in both \(B\) and
\(C\), and \(l\) and the internal labels of \(B\) and \(C\) are disjoint, so:
\[ f_A = |\text{leaf-labels}(B) \cup \text{leaf-labels}(C)| \quad \text{(by def. of leaf labels)} \]
\[ = f_B + f_C \quad \text{(no label appears in both } B \text{ and } C) \]
\[ = (n_B + 1) + (n_C + 1) \quad \text{(by structural induction hypothesis)} \]
\[ = (n_B + n_C + 1) + 1 \]
\[ = |\{l\} \cup \text{internal-labels}(B) \cup \text{internal labels}(C)| + 1 \quad \text{(uniqueness of labels)} \]
\[ = n_A + 1 \quad \text{(by def. of } n_A). \]

This proves (1) holds for A, completing the proof of the Constructor case. It follows by structural induction that (1) holds for all LBT’s with unique labels.

Problem 3.
In this problem you will prove a fact that may surprise you—or make you even more convinced that set theory is nonsense: the half-open unit interval is actually the “same size” as the nonnegative quadrant of the real plane!\(^1\) Namely, there is a bijection from \((0, 1]\) to \([0, \infty) \times [0, \infty).

(a) Describe a bijection from \((0, 1]\) to \([0, \infty).

\text{Hint: } 1/x \text{ almost works.}

\text{Solution. } f(x) := 1/x \text{ defines a bijection from } (0, 1] \text{ to } [1, \infty), \text{ so } g(x) := f(x) - 1 \text{ does the job.}

(b) An infinite sequence of the decimal digits \(\{0, 1, \ldots, 9\}\) will be called \textit{long} if it does not end with all 0’s. An equivalent way to say this is that a long sequence is one that has infinitely many occurrences of nonzero digits. Let \(L\) be the set of all such long sequences. Describe a bijection from \(L\) to the half-open real interval \((0, 1]\).

\text{Hint: } \text{Put a decimal point at the beginning of the sequence.}

\text{Solution. } \text{Putting a decimal point in front of a long sequence defines a bijection from } L \text{ to } (0, 1]. \text{ This follows because every real number in } (0, 1] \text{ has a unique long decimal expansion. Note that if we didn’t exclude the non-long sequences, namely, those sequences ending with all zeroes, this wouldn’t be a bijection. For example, putting a decimal point in front of the sequences 1000\ldots and 099999\ldots maps both sequences to the same real number, namely, 1/10.}

(c) Describe a surjective function from \(L\) to \(L^2\) that involves alternating digits from two long sequences.

\text{Hint: } \text{The surjection need not be total.}

\text{Solution. } \text{Given any long sequence } s = x_0, x_1, x_2, \ldots, \text{ let}
\[ h_0(s) := x_0, x_2, x_4, \ldots \]
\text{be the sequence of digits in even positions. Similarly, let}
\[ h_1(s) := x_1, x_3, x_5, \ldots \]
\text{be the sequence of digits in odd positions. Then } h \text{ is a surjective function from } L \text{ to } L^2, \text{ where}
\[ h(s) := \begin{cases} (h_1(s), h_2(s)), & \text{if } h_1(s) \in L \text{ and } h_2(s) \in L, \\ \text{undefined}, & \text{otherwise}. \end{cases} \quad (2)\]
Notice that any element \((y, z)\) of \(L^2\) can be merged, alternating between picking digits of \(y\) and \(z\), to produce some element of \(L\) that maps to \((y, z)\), hence \(h\) is surjective. One potential worry is whether the merge is indeed long, but that property actually holds even if only one of \(y\) and \(z\) is long, and we know that both are.

(d) Prove the following lemma and use it to conclude that there is a bijection from \(L^2\) to \((0, 1]^2\).

**Lemma 3.1.** Let \(A\) and \(B\) be nonempty sets. If there is a bijection from \(A\) to \(B\), then there is also a bijection from \(A \times A\) to \(B \times B\).

**Solution.** *Proof.* Suppose \(f : A \to B\) is a bijection. Let \(g : A^2 \to B^2\) be the function defined by the rule \(g(x, y) = (f(x), f(y))\). It is easy to show that \(g\) is a bijection:

- **\(g\) is total:** Since \(f\) is total, \(f(a_1)\) and \(f(a_2)\) exist \(\forall a_1, a_2 \in A\) and so \(g(a_1, a_2) = (f(a_1), f(a_2))\) also exists.
- **\(g\) is surjective:** Since \(f\) is surjective, for any \(b_1 \in B\) there exists \(a_i \in A\) such that \(b_i = f(a_i)\). So for any \((b_1, b_2)\) in \(B^2\), there is a pair \((a_1, a_2)\) in \(A^2\) such that \(g(a_1, a_2) := (f(a_1), f(a_2)) = (b_1, b_2)\). This shows that \(g\) is a surjection.
- **\(g\) is injective:**

\[
g(a_1, a_2) = g(a_3, a_4) \iff (f(a_1), f(a_2)) = (f(a_3), f(a_4)) \quad \text{(by def of } g)\]
\[
\quad \text{iff } f(a_1) = f(a_3) \text{ AND } f(a_2) = f(a_4) \quad \text{iff } a_1 = a_3 \text{ AND } a_2 = a_4 \quad \text{since } f \text{ is injective} \]
\[
\quad \text{iff } (a_1, a_2) = (a_3, a_4). \]

which confirms that \(g\) is injective.

Since it was shown in part (b) that there is a bijection from \(L\) to \((0, 1]\), an immediate corollary of the Lemma is that there is a bijection from \(L^2\) to \((0, 1]^2\).

(e) Conclude from the previous parts that there is a surjection from \((0, 1]\) to \((0, 1]^2\). Then appeal to the Schröder-Bernstein Theorem to show that there is actually a bijection from \((0, 1]\) to \((0, 1]^2\).

**Solution.** There is a bijection between \((0, 1]\) and \(L\) by part (b), a surjective function from \(L\) to \(L^2\) by part (c), and a bijection from \(L^2\) to \((0, 1]^2\) by part (d). These surjections compose to yield a surjection from \((0, 1]\) to \((0, 1]^2\).

Conversely, there is obviously a surjective function \(f : (0, 1]^2 \to (0, 1]\), namely

\[
f((x, y)) := x. \]

The Schröder-Bernstein Theorem now implies that there is a bijection from \((0, 1]\) to \((0, 1]^2\).

(f) Complete the proof that there is a bijection from \((0, 1]\) to \([0, \infty)^2\).

**Solution.** There is a bijection from \((0, 1]\) to \((0, 1]^2\) by part (e), and there is a bijection from \((0, 1]^2\) to \([0, \infty)^2\) by part (a) and the Lemma. These bijections compose to yield a bijection from \((0, 1]\) to \([0, \infty)^2\).