HW5 Proofs and Induction

CSE20F21

Sample Solutions

Assigned questions

1. Recall the definitions from class about factoring and divisibility: when a and b are integers and a is nonzero, a **divides** b means there is an integer c such that $b = ac$. In this case, we say a is a **factor** of b, a is a **divisor** of b, b is a **multiple** of a, $a|b$. We define the function $PosFactors : \mathbb{Z}^+ \to \mathcal{P}(\mathbb{Z}^+)$ by

$$PosFactors(n) = \{x \in \mathbb{Z}^+ \mid x \text{ is a factor of } n\}$$

Sample calculation that can be used as reference for the detail expected in your answer when working with this function:

The function application $PosFactors(4)$ evaluates to

$$PosFactors(4) = \{1, 2, 4\}$$

because the only possible positive factors of 4 are $1, 2, 3, 4$ (the positive integers less than or equal to 4) and when we divide we get:

$$\begin{aligned} 4 &= 4 \cdot 1 + 0 && \text{so 4 is a factor of 4} \\ 4 &= 3 \cdot 1 + 1 && \text{so 3 is not a factor of 4} \\ 4 &= 2 \cdot 2 + 0 && \text{so 2 is a factor of 4} \\ 4 &= 1 \cdot 4 + 0 && \text{so 1 is a factor of 4} \end{aligned}$$

(a) (*Graded for correctness*[1]) Give a witness that proves the statement

$$\exists x \in \mathbb{Z}^+ \ \forall y \in \mathbb{Z}^+ \ (\ x \in PosFactors(y) \ )$$

[1]Graded for correctness means your solution will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should explain how you arrived at your conclusions, using mathematically sound reasoning. Whether you use formal proof techniques or write a more informal argument for why something is true, your answers should always be well-supported. Your goal should be to convince the reader that your results and methods are sound.

Justify your choice of witness by explanations that include references to the relevant definitions.

Solution: A witness is $x = 1$. This is in the domain because it is a positive integer. It remains to prove that $\forall y \in \mathbb{Z}^+ \ (1 \in PosFactors(y) \ )$. Let y be an arbitrary positive integer. We need to show that $1 \in PosFactors(y)$, namely (by definition of $PosFactors$), that 1 is a factor of y. By definition of factors, this means finding a witness $c \in \mathbb{Z}$ such that $y = 1 \cdot c$. Consider $c = y$, an integer (because y is a positive integer), and $1 \cdot c = 1 \cdot y = y$, as required. $\square$

(b) (*Graded for correctness*) Give a counterexample that disproves the statement

$$\forall n \in \mathbb{Z}^+ \ (\ PosFactors(n) \subseteq PosFactors(n+1) \ )$$

Justify your choice of counterexample by explanations that include references to the relevant definitions.

Solution: A counterexample is $n = 3$, a positive integer so in the domain. We need to show that $\neg(PosFactors(3) \subseteq PosFactors(4))$ so we start by calculating the two sets using the definition of $PosFactors$:

- $PosFactors(3) = \{1, 3\}$
- $PosFactors(4) = \{1, 2, 4\}$

Since $3 \in PosFactors(3)$ but $3 \notin PosFactors(4)$, 3 serves as a counterexample to show that $\forall x(x \in PosFactors(3) \rightarrow x \in PosFactors(4))$ is false, and hence (by the definition of subset inclusion) that $PosFactors(3) \subseteq PosFactors(4)$ is false, as required.

(c) (*Graded for fair effort completeness*) Consider the following attempted proof.

> **Attempted proof**: For arbitrary integers a, b, c, assume towards a direct proof that $(a + b)|c$. We need to show that $a|c$ and $b|c$. Let n be the integer c **div** $(a+b)$. Since $(a+b)|c$, by definition of divides, $n|c$ and n is an integer. Since $c = 1 \cdot n \cdot (a+b)$, $(n \cdot (a+b))|c$. Rewriting by distributing multiplication over addition, we have $na|c$ and $nb|c$. Since $a|na$ and $na|c$, we have $a|c$. Similarly, since $b|nb$ and $nb|c$, we have $b|c$. Thus, we have proved both conjuncts and the proof is complete.

Select the statement below that the attempted proof is trying to prove.

(i) $\forall a \in \mathbb{Z}^+ \ \forall b^+ \in \mathbb{Z} \ \forall c \in \mathbb{Z} \ (\ (\ a|c \ \vee \ b|c \ ) \ \rightarrow \ (a + b)|c \ )$

(ii) $\forall a \in \mathbb{Z}^+ \ \forall b^+ \in \mathbb{Z} \ \forall c \in \mathbb{Z} \ (\ (\ a|b \ \wedge \ a|c \ ) \ \rightarrow \ a|(b + c) \ )$

(iii) $\forall a \in \mathbb{Z}^+ \ \forall b^+ \in \mathbb{Z} \ \forall c \in \mathbb{Z} \ (\ (a + b)|c \ \rightarrow \ (\ a|c \ \wedge \ b|c \ ) \ )$

Identify the first major error in the attempted proof and explain why it is incorrect.

Next, disprove the statement the attempted proof was attempting to prove.

Extra practice; not for credit: prove or disprove the other two statements.

Solution: This proof is attempting to prove (iii) because we take arbitrary integers a, b, c so it's working to prove a universal claim about integers, and the direct proof strategy suggests that the predicate being universally claimed is $(\ (a + b)|c \ ) \rightarrow (\ a|c \wedge b|c \ )$.

The first major error is distributing multiplication over addition to try to rewrite the divisibility relations, which has not been proved.

To disprove (iii), take $a = 2$, $b = 7$, $c = 9$ as a candidate counterexample, noticing that $2, 7$ are positive integers and 9 is an integer. We will show that $(a + b)|c$ but it is not the case that $(a|c \wedge b|c)$. Calculating:

$$\begin{aligned} c = 9 = 1 \cdot 9 = 1 \cdot (2 + 7) = 1 \cdot (a + b) \qquad & \text{so } (a + b)|c \\ c = 9 = 4 \cdot 2 + 1 = 4 \cdot a + 1 \qquad & \text{so } c \textbf{ mod } a \neq 0 \\ c = 9 = 1 \cdot 7 + 2 = 1 \cdot b + 2 \qquad & \text{so } c \textbf{ mod } b \neq 0 \end{aligned}$$

Thus, the choices of a, b, c give a counterexample to the universal claim and we have disproved statement (iii).

2. In this question, we'll consider the function which calculates the sum of the first n positive integers.

   (a) (*Graded for fair effort completeness*[2])
   Give a recursive definition of this function, including domain, codomain and both the basis step and recursive step of the rule. That is, fill in the blanks

   $$sumOfFirst : \underline{\quad domain \quad} \to \underline{\quad codomain}$$

   given by

   Basis step : fill in basis step
   Recursive step : fill in recursive step

   Notation: Using summation, this function can be written $sumOfFirst(n) = \sum_{i=1}^{n} i$.

   Solution: $sumOfFirst : \mathbb{Z}^+ \to \mathbb{Z}^+$ is given by

   Basis step : $sumOfFirst(1) = 1$
   Recursive step : $sumOfFirst(n + 1) = sumOfFirst(n) + n + 1$

   (b) (*Graded for fair effort completeness*) It turns out that the value of this function can also be calculated explicitly (without recursion)[3]. You will prove this by completing the proof of the identity

   $$\forall n \in \mathbb{Z}^+ \left(sumOfFirst(n) = \frac{n(n+1)}{2} \right)$$

[2]Graded for fair effort completeness means you will get full credit so long as your submission demonstrates honest effort to answer the question. You will not be penalized for incorrect answers.

[3]When the value of a function that is recursively defined can also be calculated without recursion, we call the formula that we can use to calculate the value without recursion the "closed form formula" for the function.

Fill in the missing parts of the proof of this statement:

Proof: We proceed by mathematical induction on the set of positive integers.
Basis Step: Choose $n = 1$ as the basis step. Using the Basis Step in the recursive definiton of $sumOfFirst$,

$$sumOfFirst(1) = 1$$

Plugging n into the RHS of the desired formula, $\frac{1(1+1)}{2} = \frac{2}{2} = 1$. Since LHS=RHS, the Basis step is complete.
Recursive Step: Consider an arbitrary $k \geq 1$. We assume (as the induction hypothesis) that $\underline{sumOfFirst(k) = \frac{k(k+1)}{2}}$.
We want to show that $sumOfFirst(k+1) = \frac{(k+1)\cdot((k+1)+1)}{2}$.
Solution:

$$\begin{aligned}
sumOfFirst(k+1) &= sumOfFirst(k) + k + 1 && \text{using recursive definition of } sumOfFirst \\
&= \frac{k(k+1)}{2} + k + 1 && \text{using induction hypothesis} \\
&= \frac{k(k+1)}{2} + \frac{2(k+1)}{2} && \text{putting the terms over a common denominator} \\
&= \frac{(k(k+1) + 2(k+1))}{2} \\
&= \frac{(k+2)(k+1)}{2} && \text{factoring the } (k+1) \text{ term in the numerator} \\
&= \frac{(k+1)((k+1)+1)}{2}
\end{aligned}$$

which completes the recursive step.

(c) (*Graded for fair effort completeness*) When calculating the runtime of an algorithm, nested for loops sometimes lead to program runtimes that involve the sum of the first n positive integers. To estimate the rate of growth of this runtime, it is useful to find an upper bound for this function in terms of a simpler function. Use the explicit formula from the earlier parts of this question and mathematical induction to prove

$$\forall n \in \mathbb{Z}^+ \ \left(\ sumOfFirst(n) \leq n^2 \right)$$

Solution: We proceed by mathematical induction on the set of positive integers.
Basis step: Using the basis step of the recursive definition of the functiton,

$$sumOfFirst(1) = 1$$

Plugging 1 into the RHS, we get $LHS = 1$, $RHS = 1$, so $LHS \leq RHS$ and the basis step is complete.

Recursive step: Let k be an arbitrary positive integer. Suppose as our inductive hypothesis that the claim holds for this k. We want to show the predicate holds for k+1 By our IH, $sumOfFirst(k) \leq k^2$, and we want to show that $sumOfFirst(k+1) \leq (k+1)^2$. Calculating:

$$\begin{aligned} sumOfFirst(k+1) &= sumOfFirst(k) + (k+1) && \text{using definition of } sumOfFirst \\ &\leq k^2 + (k+1) && \text{by induction hypothesis} \\ &\leq k^2 + k + k + 1 && \text{because } 1 \leq k \\ &= k^2 + 2k + 1 = (k+1)^2 && \text{as required} \end{aligned}$$

The inductive step is proven.

3. Recall the definition of linked lists that we discussed in class.

   Define the function *count* which returns the number of occurrences of a datum in the list. Formally, $count : L \times \mathbb{N} \to \mathbb{N}$, where

$$\begin{aligned} \textbf{Basis Step:} \quad & \text{If } m \in \mathbb{N}, \quad count(\ ([], m)\ ) = 0 \\ \textbf{Recursive Step:} \quad & \text{If } l \in L \text{ and } n \in \mathbb{N} \text{ and } m \in \mathbb{N}, \text{ then} \\ & count(\ (\ (n, l), m\ )\ ) = \begin{cases} 1 + count(\ (l, m)\ ) & \text{if } n = m \\ count(\ (l, m)\ ) & \text{otherwise} \end{cases} \end{aligned}$$

   A mystery function is defined by $mystery : L \times \mathbb{N} \to L$, where

$$\begin{aligned} \textbf{Basis Step:} \quad & \text{If } m \in \mathbb{N}, \quad mystery(\ ([], m)\ ) = [] \\ \textbf{Recursive Step:} \quad & \text{If } l \in L \text{ and } n \in \mathbb{N} \text{ and } m \in \mathbb{N}, \text{ then} \\ & mystery(\ (\ (n, l), m\ )\ ) = \begin{cases} l & \text{if } n = m \\ mystery(\ (l, m)\ ) & \text{otherwise} \end{cases} \end{aligned}$$

   (a) (*Graded for correctness*) Prove that

$$\forall m \in \mathbb{N}\ \exists l \in L\ \ (count(\ (l, 20)\ ) = m\ )$$

   Solution: We proceed by mathematical induction on $\mathbb{N}$

   Basis step: We need to prove that $\exists l \in L\ \ (count(\ (l, 20)\ ) = 0\ )$. Consider the witness $l = (1, [])$, a linked list by definition of L. To confirm that this candidate witness satisfies the predicate:

$$count(\ ((1, []), 20)\ ) = count(\ ([], 20)\ ) = 0$$

   where the first equality is by the second case in the recursive step in the definition of *count* and the second is by the basis step in the definition of count.

Note: there were lots of different lists we could have used for the the witness l. Can you think of some others?

Recursive step: Consider an arbitrary natural number m. We need to show that

$$\exists l \in L \ (count(\ (l, 20)\ ) = m\ ) \rightarrow \exists l \in L \ (count(\ (l, 20)\ ) = m + 1\ )$$

Assume that $\exists l \in L \ (count(\ (l, 20)\ ) = m\ )$ (for this specific m). Let l_0 be the witness list for this existential statement. We define $l_{next} = (20, l_0)$. This is a list because its head node datum is a natural number and its tail is a list. Moreover, applying the function definition

$$count(\ (l_{next}, 20)\ ) = count(\ ((20, l_0), 20)\ ) = 1 + count(\ (l_0, 20)\ ) = 1 + m = m + 1$$

where the first equality is by definition of l_{next}, the next equality is by the first case in the recursive step in the definition of *count* and the next equality is by the induction hypothesis.

(b) (*Graded for correctness*) Give an example input x to the function such that

$$mystery(\ (\ x\ ,\ 2\ )\ ) = []$$

For full credit, include all intermediate steps of the function application that justifies your choice of x, with brief justifications for each.

Solution: Consider $x = []$. Then, by the basis step in the definition of *mystery*

$$mystery(\ (\ [], 2,\ )\ ) = []$$

as required.

A different example that also works would be $(9, (8, (7, (6, []))))$:

$$\begin{aligned}
&mystery(\ (\ (9, (8, (7, (6, []))))\ , 2)\ ) \\
&= mystery(\ (\ (8, (7, (6, []))), 2\ )\ ) \qquad \text{where } n = 9,\ m = 2 \\
&= mystery(\ (\ (7, (6, [])), 2\ )\ ) \qquad \text{where } n = 8,\ m = 2 \\
&= mystery(\ (\ (6, []), 2\ )\ ) \qquad \text{where } n = 7,\ m = 2 \\
&= mystery(\ (\ [], 2\ )\ ) \qquad \text{where } n = 6,\ m = 2 \\
&= [] \qquad \text{using basis step}
\end{aligned}$$

where the intermediate steps use the recursive step in the definition of *mystery*.

(c) (*Graded for correctness*) Evaluate the function application

$$mystery(\ (\ (2, (0, (2, [])))\ ,\ 2\ )\ )$$

For full credit, include all intermediate steps of the function application, with brief justifications for each.

Solution: The input to the function application is the ordered pair $(\ (n, l), m\ )$ with $(n, l) = (2, (0, (2, [])))$ and $m = 2$. Notice that $n = 2 = m$. Applying the first case in the recursive step of the function definition:

$$mystery(\ (\ (2, (0, (2, [])))\ ,\ 2\ )\ ) = (0, (2, []))$$

because $n = 2 = m$ and $l = (0, (2, []))$

(d) (*Graded for fair effort completeness for English statements and correctness in use of syntax for symbolic statements*) Describe the rule of the function *mystery* in English. Then, write a true statement that describes an invariant using both the functions *mystery* and *count*. Express this invariant both symbolically and in English.

Solution: The rule for *mystery* amounts to returning the tail of the list after the first occurrence of m, if there is an occurrence of m in the list, and returning the empty list otherwise.

Therefore, an invariant about this function is

$$\forall l \in L \forall m \in \mathbb{N}\ (\ count(\ (l, m)\ ) \geq count(\ (mystery(\ (l, m)\ ), m\ )\ )\ )$$

In English, this statement is: The portion of a linked list after the first occurrence of a natural number has no more occurrencs of that number than the original list.