

**Q1.** On pages 55 and 56 of the textbook there is a proof for the correctness of the program `avg`. The author used the invariant

$$Inv(i, \text{sum}): 0 \leq i < \text{len}(A) \wedge \text{sum} = \sum_{k=0}^{i-1} A[k].$$

As you can see, the invariant is a predicate in two variables: `i`, `sum`. These two variables are used in the program `avg`, but neither is the variable on which the induction proof is based. The author is using simple induction, but it is not very clear what the induction variable is (also the induction predicate itself is ambiguous). We want to make sure you understand what is going on there by having you re-write the proof by yourself in a style similar to the one we use in lectures. Here is the predicate you should be proving:

$$Q(j) : \text{At the beginning of the } j^{\text{th}} \text{ iteration, } \text{sum} = \sum_{k=0}^{i-1} A[k].$$

Remark 1: By the program's design, the variables `i`, `sum` may change with each iteration (in other words, both are functions of  $j$ ). This is why it might be more appropriate to write

$$Q(j) : \text{sum}_j = \sum_{k=0}^{i_j-1} A[k] \quad \text{or} \quad Q(j) : \text{sum}(j) = \sum_{k=0}^{i(j)-1} A[k].$$

where  $i_j$  (or  $i(j)$ ) means the value of program variable `i` at the beginning of the  $j^{\text{th}}$  iteration (the same with `sum`). That said, we believe that the version above Remark 1 is the best to work with as long as one understands that `i`, `sum` are iteration dependent.

Remark 2: Using  $j$  as an index in Remark 1 has a different meaning from that which is intended by the author. The author is using indices to differentiate between the values of `i`, `sum` at the beginning of an (arbitrary) iteration and their values after the iteration is run. The end goal is to prove  $Q(\text{len}(A))$ . A proof by induction will show that

$$\forall j \in \{1, \dots, \text{len}(A)\}, Q(j).$$

Your proof must follow the style used in lectures.

*Proof.* It will be shown using simple induction that  $\forall j \in \{1, \dots, \text{len}(A)\}, Q(j)$ .

**Base Case.** Let  $j = 1$ . At the beginning of the first iteration of the loop,  $i = 0$  and  $\text{sum} = 0$ . Notice that in the expression  $\sum_{k=0}^{-1} A[k]$ , the lower bound is greater than the upper bound. This results in an empty sum that evaluates to 0. Therefore the base case holds.

**Induction Hypothesis.** Suppose that for some  $j \in \{1, \dots, \text{len}(A)\}$ ,  $Q(j)$  is true.

**Induction Step.**  $i$  increases by 1 at the end of each iteration, so at the beginning of the  $j$ th iteration,  $i = j - 1 < \text{len}(A)$ , which means that the loop body enters. From the induction hypothesis,

$$\text{sum} = \sum_{k=0}^{j-2} A[k]$$

Running through the loop, on line 9,  $A[j - 1]$  is added to  $sum$ . After this line is executed, the new value of  $sum$  becomes

$$sum = \sum_{k=0}^{j-2} A[k] + A[j - 1] = \sum_{k=0}^{j-1} A[k]$$

Then, line 10 executes and the value of  $i$  is incremented and becomes  $j$ . After the line is executed, the program returns to the top of loop, signifying the beginning of the  $j + 1$ th iteration.  $sum$  remains unchanged since the execution of line 9, and is equal to

$$\sum_{k=0}^{j-1} A[k] = \sum_{k=0}^{i-1} A[k]$$

as desired.

By the principle of simple induction,  $Q(j)$  holds for any  $j \in \{1, 2, \dots, \text{len}(A)\}$ . □