Exercise answers - Chapter 2 - Getting Started

Christian Duffau-Rasmussen

September 15, 2018

Exercise 2.1-3 Linear Search

procedure LINEAR-SEARCH(A, v)
  for $i \leftarrow 1$ **to** length$[A]$ **do**
    if $A[i] == v$ **then**
      return i
  return NILL

Proof of correctness To prove linear search is correct, I formulate a *loop invariant* that need to hold true at *initialization*, during *maintainance* and at *termination*.

Loop invariant At the start of loop iteration i, the value v is not in the sub-array $A[1..i-1]$. In formal terms, $v \notin \{A[j] | j \in \{1, ..., i-1\}\}$

Initialization At initialization $i = 0$, so the sub-array $A[1..i-1]$ is empty. So v is trivially not a member of that array, and the loop invariant holds.

Maintenance Assuming that v is not in the sub-array $A[1..i-1]$ at the start of iteration i, then two outcomes are possible during the iteration. Either, $A[i] == v$ and the loop terminates, or $A[i] \neq v$ and v is not an element of the array $A[1..i]$ which is the loop invariance condition for iteration $i+1$.

Termination The algoritm terminates in two different ways. The first happens when $A[i] == v$, assuming $v \notin A[1..i-1]$ before termination, then the loop invariant still holds. Again, assuming $v \notin A[1..i-1]$ before termination. In the second case we have that $i = n+1$, in that case $v \notin A[1..i-1] = A[1..n]$, which is the entire array, and the procedure returns NILL.

2.2-3 Running time of Linear search

Assuming the index of correct element is uniformlly distributed from 1 to n, the expected index is $E[i] = n/2$. So the average case running time must be $c_1 \frac{n}{2} + c_2$, where c_1 is the running time of each comparison $A[i] == v$ and the loop increment, and c_2 is the running time of the return statement. In the worst case the algorithm needs to run through all n elements as well as returning NILL, so the running time is $c_1 n + c_2$. Both the average and worst case running times of linear search are $\Theta(n)$.

Exercise 2.3-3 Linear Reccurence

To show that the reccursion

$$T(n) = \begin{cases} 2 & \text{if } n = 2 \\ 2T(n/2) + n & \text{if } n = 2^k, \text{ for } k > 1 \end{cases}$$

is $T(n) = n \lg n$. Assuming that n is a power of two we have $n = 2^k$ for $k \in 1, 2, \ldots$. For $k = 1$ we have that $T(n) = T(2) = 2$ by definition of T.

$$2 \lg(2) = 2 = T(2).$$

Hence the claim holds for $k = 1$.

Next, assuming that $T(2^{k-1}) = 2^{k-1} \lg 2^{k-1} = (k-1)2^{k-1}$ we have that,

$$\begin{aligned} T(n) &= 2T(n/2) + n \\ &= 2T(2^k/2) + 2^k \\ &= 2T(2^{k-1}) + 2^k \\ &= 2((k-1)2^{k-1}) + 2^k \\ &= k2^k \\ &= 2^k \lg(2^k) = n \lg n \end{aligned}$$

so the claim also holds for k given it holds for $k-1$, at since it holds true for $k = 1$ it holds for any $k \geq 1$. $\square$

Exercise 2.3-5 Binary search

procedure BINARY-SEARCH(A, p, r, v)
  if $p \leq r$ **then**
    $m = p + \left\lfloor \frac{r-p}{2} \right\rfloor$
    if $A[m] == v$ **then**
      return m
    else if $A[m] > v$ **then**
      return BINARY-SEARCH(A, p, m - 1, v)
    else if $A[m] < v$ **then**
      return BINARY-SEARCH(A, m +1, r, v)
  return NILL

In the worst case, binary search needs to split the array in two, untill the last split goes from two elements to a single one. A way of assessing the running time is to understand how many possible splits can be made. Assuming the length of the array is $n = 2^k$, the length after the first split is simply $n/2$. Finding the length of the consequent halfed arrays, amounts to raising the *half* to the nummmber of splits and multiplying by the initial size, n. So, to find the number of splits which result in the length of 1, we need to solve the following,

$$1 = \frac{n}{2^x} \quad \Leftrightarrow \quad 2^x = n \quad \Leftrightarrow \quad x = \lg n.$$

So we can conclude, that the recursion calls itself at most $\lg n$ times, this is an indication of the worst case running time being $\Theta(\lg n)$.

Problem 2-3 Horner's rule

a. Running time The running time of each line in Horner's algoritm is given by,

$y \leftarrow 0$ ▷ Running time: c_1
$i \leftarrow n$ ▷ Running time: c_1
while $i \geq 0$ **do** ▷ Running time: c_2 repeated $n+2$ times.
    $y \leftarrow a_i + x \cdot y$ ▷ Running time: $c_1 + c_+ + c_\bullet$ repeated $n+1$ times.
    $i \leftarrow i - 1$ ▷ Running time: $c_1 + c_-$ repeated $n+1$ times.

The total running time is $T(n) = 2c_1 + c_2(n+2) + (2c_1 + c_+ + c_\bullet + c_-)(n+1)$, that is

$$T(n) = C_1 + C_2 n$$

with $C_1 = 4c_1 + 2c_2 + c_+ + c_\bullet + c_-$ and $C_2 = 2c_1 + c_2 + c_+ + c_\bullet + c_-$. So, it's clear that $T(n)$ is a linear function of n so $T(n) = \Theta(n)$.

b. Naive polynomial evaluation

$y \leftarrow 0$
$z \leftarrow 1$
$i \leftarrow 0$
$j \leftarrow 0$
while $i \leq n$ **do**
    while $j < i$ **do** ▷ This inner loop is evaluated $\sum_{k=1}^{n+1} k$ times.
        $z \leftarrow z \cdot x$
        $j \leftarrow j + 1$
    $y \leftarrow y + a_i \cdot z$
    $i \leftarrow i + 1$
    $z \leftarrow 1$
    $j \leftarrow 0$

The inner loop is called $\sum_{k=1}^{n+1} k = (n+1)(n+2)/2 = n^2 + 3n + 2$ times, and the evaluation in the outer loop are called n times, so the running time is no more than,

$$T(n) = C_0 + C_1 n + C_2 n^2$$

for the appropriate constants. So for the naive implementation $T(n) = \Theta(n^2)$.

c. Proof of correctness Consdering the loop invariant,

$$y = \sum_{k=0}^{n-(i+1)} a_{k+i+1} x^k.$$

Initialization We have $i = n$ so the above expression is $y = \sum_{k=0}^{-1} a_{k+n+1} x^k = 0$, which is in line with the assigned value to y of 0 in line 1 of the algorithm. So the loop invariant is true at initialization.

Maintainance During maintenance assuming $i = j+1$ with $0 \leq j+1 < n$ and sssuming y is given by

$$y = \sum_{k=0}^{n-(j+2)} a_{k+j+2} x^k = a_{j+2} + a_{j+3}x + a_{j+4}x^2 + \ldots + a_n x^{n-(j+2)}$$

at the start of the loop, then at the start of the next iteration, $i = j$, y is given by

$$\begin{aligned} y \leftarrow & a_{j+1} + x \cdot y \\ = & a_{j+1} + x \left(a_{i+2} + a_{j+3}x + a_{j+4}x^2 + \ldots + a_n x^{n-(j+2)} \right) \\ = & a_{j+1} + a_{i+2}x + a_{j+3}x^2 + a_{j+4}x^3 + \ldots + a_n x^{n-(j+1)} \end{aligned}$$

which corresponds to the loop invariant at iteration $i = j$. So, the loop invariant holds in the maintainance step as well.

Termination At termination $i = -1$ so the loop invariant becomes the full polynomial,

$$y = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k.$$

Assuming the invariant holds true at the start of the iteration, the algorithm evaluates y to the correct value.

d. Conclusion In conclusion, the loop invariant holds true at initialization, through maintainance all the way to termination by induction, and ends with a correct evaluation of the polynomial, so we can conclude that the algorithm is correct.

Problem 2-4 Inversions

a. The array $A = \langle 2, 3, 8, 6, 1 \rangle$ has the inversions $(4, 5)$, $(3, 5)$, $(2, 5)$, $(1, 5)$ and $(3, 4)$.

b. The array with all the elements in reverse order, $A = \langle n, n-1, \ldots, 2, 1 \rangle$, has the most *inversions*. The last element $A[n]$ appears in $n-1$ inversions, since all other elements are greater and have a smaller index. Element $A[n-1]$ appears in $n-2$ inversions, since all elements to the left are greater and have a smaller index. So, starting from the left the number of inversions for each element in the array is, $n-1, n-2, \ldots, 1, 0$. So the total number of inversions in the inverted array are,

$$\sum_{k=1}^{n}(k-1) = \sum_{k=1}^{n} k - n = \frac{n(n+1)}{2} - n = \frac{n(n-1)}{2}, \quad \text{for} \quad n > 0.$$