

CS208/204: Data Structures & Abstractions  
Mahidol University International College

## Assignment 5: Performance Characterization

Austin J. Maddison

June 19, 2023

# Contents

Contents	i
1 Task 1: Hello, Definition	1
2 Task 2: Hello, Definition	3
3 Task 3: How Long Does This Take	4
4 Task 4: Halving, Sum	7
5 Task 5: More Running Time Analysis	9
6 Task 6: Recursive Code	11
7 Task 7: Counting Dashes	14

# Task 1: Hello, Definition

## P1

To prove that  $f(n) \leq c \cdot g(n)$ , such that  $c$  and  $n$  exist in  $\mathbf{I}^+$ , and  $n \geq n_0$ , where  $f(n) = n$  and  $g(n) = n \log n$ . We need to find values for  $c$  and  $n_0$  such that the inequality holds true.

Let's choose  $c = 1$  and  $n_0 = 1$ .

For  $n \geq 1$ , we have:

$$\begin{aligned}f(n) &= n \\g(n) &= n \log n\end{aligned}$$

Substituting these values into the inequality, we get:

$$n \leq n \log n$$

Taking the logarithm of both sides, we have:

$$\log n \leq \log(n \log n)$$

Using the logarithmic property  $\log(a \cdot b) = \log a + \log b$ , we can rewrite the right side as:

$$\log(n \log n) = \log n + \log(\log n)$$

Now, the inequality becomes:

$$\log n \leq \log n + \log(\log n)$$

Since  $\log n$  is a positive value, we can subtract  $\log n$  from both sides, resulting in:

$$0 \leq \log(\log n)$$

For  $n \geq 1$ , the logarithm of the logarithm of  $n$  is always a positive value or zero. Therefore, the inequality  $0 \leq \log(\log n)$  holds true.

We have shown that for  $c = 1$  and  $n_0 = 1$ ,  $f(n) = n$  is less than or equal to  $c \cdot g(n)$  for all  $n \geq n_0$ .

Therefore, we can conclude that  $f(n) = n$  is  $O(n \log n)$ .

**P2**

Consider the proposition  $d(n)$  is  $O(f(n))$  and  $e(n)$  is  $O(g(n))$  such that  $d(n) \cdot e(n)$  is  $O(f(n) \cdot g(n))$

To prove that  $d(n) \cdot e(n)$  is  $O(f(n) \cdot g(n))$  we got to make sure that  $f(n) \cdot g(n)$  still valid under the definition of  $O()$ . Which means the equality:

$$d(n) \cdot e(n) \leq f(n) \cdot g(n) \text{ needs to be satisfied for } O(f(n) \cdot g(n)) \text{ to be valid.}$$

The equality can be proven because we know the

$$d(n) \text{ is } O(f(n)) \text{ and } e(n) \text{ is } O(g(n))$$

Which implies:

$$d(n) \leq f(n) \text{ and } e(n) \leq g(n)$$

Hence the product satisfies the equality

$$d(n) \cdot e(n) \leq f(n) \cdot g(n)$$

thus concludes the  $O(n)$  of  $d(n) \cdot e(n)$  is  $O(f(n) \cdot g(n))$ .

**P3**

```
void fnA(int S[]) {
    int n = S.length;
    for (int i=0; i<n; i++) {
        fnE(i, S[i]);
    }
}
```

The following codes snippet's complexity is  $O(n^2)$ . The for loop inside  $fnA()$  has the time complexity of  $O(n)$  as it takes  $n$  operations to iterate through the list. The time complexity of the  $fnE()$  is also  $O(n)$  as  $n$  increases the amount of operations  $fnE$  has to do is linear. Notice that since the call for  $fnE()$  is inside  $fnA()$ 's for loop multiplies the time complexity making the amount of work needed to be done quadratic.

$$\underbrace{O(n)}_{fnA()} \cdot \underbrace{O(n)}_{fnE()} \text{ is } O(n^2)$$

**P4**

The reason why  $h(n) = 16n^2 + 11n^4 + 0.1n^5$  is not  $n^4$  is because  $O(n^4)$  does not satisfy the definition of  $O()$ . Let  $f(n) = n^4$ .

$$h(n) \text{ is not } \leq f(n) \text{ for all values}$$

We can prove this by evaluating the limits using L'Hopital:

$$\begin{aligned} &= \lim_{n \rightarrow \infty} \frac{h(n)}{f(n)} \\ &= \lim_{n \rightarrow \infty} \frac{16n^2 + 11n^4 + 0.1n^5}{n^4} \\ &= \lim_{n \rightarrow \infty} \frac{32n + 44n^3 + 0.5n^4}{4n^3} \\ &= \lim_{n \rightarrow \infty} \frac{n + n^3 + n^4}{n^3} \end{aligned}$$

This shows that  $h(n)$  contains a polynomial of  $n$  with higher order than  $f(n)$ . This concludes that  $h(n)$  is greater than  $f(n)$  when  $n$  is a big number thus  $f(n)$  cannot be an upper-bound to  $h(n)$ .

## Task 2: Hello, Definition

To prove that  $f(n) \leq c \cdot g(n)$ , such that  $c$  and  $n$  exist in  $\mathbf{I}^+$ , and  $n \geq n_0$ , where  $f(n) = n$  and  $g(n) = n \log n$ . We need to find values for  $c$  and  $n_0$  such that the inequality holds true.

Let's choose  $c = 1$  and  $n_0 = 1$ .

For  $n \geq 1$ , we have:

$$\begin{aligned}f(n) &= n \\g(n) &= n \log n\end{aligned}$$

Substituting these values into the inequality, we get:

$$n \leq n \log n$$

Taking the logarithm of both sides, we have:

$$\log n \leq \log(n \log n)$$

Using the logarithmic property  $\log(a \cdot b) = \log a + \log b$ , we can rewrite the right side as:

$$\log(n \log n) = \log n + \log(\log n)$$

Now, the inequality becomes:

$$\log n \leq \log n + \log(\log n)$$

Since  $\log n$  is a positive value, we can subtract  $\log n$  from both sides, resulting in:

$$0 \leq \log(\log n)$$

For  $n \geq 1$ , the logarithm of the logarithm of  $n$  is always a positive value or zero. Therefore, the inequality  $0 \leq \log(\log n)$  holds true.

We have shown that for  $c = 1$  and  $n_0 = 1$ ,  $f(n) = n$  is less than or equal to  $c \cdot g(n)$  for all  $n \geq n_0$ .

Therefore, we can conclude that  $f(n) = n$  is  $O(n \log n)$ .

## Task 3: How Long Does This Take

```
void programA(int n) {  
    long prod = 1;  
    for (int c=n; c>0; c=c/2)  
        prod = prod * c;  
}
```

### programA()

The time complexity of the programA() is  $\Theta(\log_2 n)$ . This can be proven by translating the for loop into a recurrence relation making sure it has the same iterating and terminating behavior to ensure it shares the same time complexity. We can ignore the variable `prod` as it doesn't have significance in analyzing the time complexity of the function.

$$T(n) = T\left(\frac{n}{2}\right)$$

Now we can plug and chug to inspect the behavior of the function, specifically the behavior being how fast the function takes terminate relation to how big  $n$  is.

$$\begin{aligned} T(16) &= T\left(\frac{16}{2}\right) \\ &= T\left(\frac{8}{2}\right) \\ &= T\left(\frac{4}{2}\right) \\ &= T\left(\frac{2}{2}\right) \\ &= \dots \end{aligned}$$

We can see that for every recurrence of the function  $T(n)$  input  $n$  starts off reducing rapidly then slows down as it gets closer to zero. We can deduce that the time complexity must be some log function.

Furthermore, we might stare the numerators long enough such that we realize they are numbers of base two where their exponent is being reduced by 1 every recurrence.

$$\begin{aligned}
 T(2^4) &= T\left(\frac{2^4}{2}\right) \\
 &= T\left(\frac{2^3}{2}\right) \\
 &= T\left(\frac{2^2}{2}\right) \\
 &= T\left(\frac{2}{2}\right) \\
 &= \dots
 \end{aligned}$$

Thus we can further deduce with confidence that the time complexity must be of  $\log_2 n$ .

Since we were able to deduce the exact behavior of the function using the recurrence relation it is safe to say that there exist constants that can create a tight upper-bound and lower-bound of `programA()`. Hence, we can conclude that the time complexity of `programA()` is  $\Theta(\log_2 n)$ .

---

### **programB()**

```

void programB(int n) {
    long prod = 1;
    for (int c=1; c<n; c=c*3)
        prod = prod * c;
}

```

Similarly, the time complexity of the `programB()` is  $\Theta(\log_3 n)$ . This can be proven by again translating the for loop into a recurrence relation making sure it has the same iterating and terminating behavior to ensure it shares the same time complexity.

$$T(n) = T\left(\frac{n}{3}\right)$$

Now we can plug and chug to inspect the behavior of the function, specifically the behavior being how fast the function takes terminate relation to how big  $n$  is.

$$\begin{aligned}
 T(81) &= T\left(\frac{81}{3}\right) \\
 &= T\left(\frac{27}{3}\right) \\
 &= T\left(\frac{9}{3}\right) \\
 &= T\left(\frac{3}{3}\right) \\
 &= \dots
 \end{aligned}$$

We can see that for every recurrence of the function  $T(n)$  input  $n$  starts off reducing rapidly then slows down as it gets closer to zero. We can deduce that the time complexity must be some log function.

Again, we might stare the numerators long enough such that we realize they are numbers of base 3 where their exponent is being reduced by 1 every recurrence.

$$\begin{aligned}T(3^4) &= T\left(\frac{3^4}{3}\right) \\&= T\left(\frac{3^3}{3}\right) \\&= T\left(\frac{3^2}{3}\right) \\&= T\left(\frac{3}{3}\right) \\&= \dots\end{aligned}$$

Thus we can further deduce with confidence that the time complexity must be of  $\log_3 n$ .

Since we were able to deduce the exact behavior of the function using the recurrence relation it is safe to say that there exist constants that have a tight upper-bound and lower-bound. Hence, we can conclude that the time complexity of `programB()` is  $\Theta(\log_3 n)$ .



## Task 4: Halving, Sum

```
def hsum(X): # assume len(X) is a power of two
    while len(X) > 1:
        (1) allocate Y as an array of length len(X)/2
        (2) fill in Y so that Y[i] = X[2i] + X[2i+1] for i = 0, 1, ..., len(X)/2 - 1
        (3) X = Y
    return X[0]
```

---

### P1

The amount of work being done in steps 1-3 of fsum() in terms of  $k_1$  and  $k_2$  where  $k_1, k_2 \in \mathbf{R}_+$  under the assumptions provided is:

$$\underbrace{\left(\frac{1}{2} \cdot k_1\right)z}_{\text{Step 1}} + \underbrace{\left(\frac{z}{2} \cdot k_2\right)}_{\text{Step 2}} + \underbrace{(k_2)}_{\text{Step 3}}$$

**Step 1:** We have to allocate  $Y[]$  which is half the size of array  $X$ . We know that  $k_1 \cdot z$  is the cost of allocating array of size  $X$ . Hence allocating array of size  $Y$  is the same as allocating size array of size  $X$  divided by 2.

$$\left(\frac{1}{2} \cdot k_1\right)z$$

**Step 2:** We have to read, summate, then store  $X[i]$  and  $X[i+1]$  in  $Y[i]$  for  $i = 0, 1, 2, \dots, \frac{z}{2} - 1$ . In the assumption it states that this set of operation done for each value that ends up in array  $Y$  is equal to the cost of  $k_2$ . Since we are summing every subsequent element pair  $X[], Y[]$  will be half the size of  $X[]$ . Hence, we will have to use the  $k_2 \cdot \frac{z}{2}$  times.

$$\left(\frac{z}{2} \cdot k_2\right)$$

**Step 3:** Finally we will have to use 1 more operation  $k_2$  to pass the array made in  $Y[]$  to  $X[]$ .

$$k_2$$

**P2**

To analyze the time complexity of this function I will construct a recurrence relation that has the same behavior as the function `hsum()`. This will help us observe how the length of `X[]` changes as the function iterates. To do this we can create a recurrence relation that halves the input for every iteration of the recurrence.

$$T(n) = T\left(\frac{n}{2}\right)$$

Let's plug and chug.

$$\begin{aligned} T(32) &= T\left(\frac{32}{2}\right) \\ &= T\left(\frac{16}{2}\right) \\ &= T\left(\frac{8}{2}\right) \\ &= T\left(\frac{4}{2}\right) \\ &= T\left(\frac{2}{2}\right) \\ &= \dots \end{aligned}$$

From inspection, we can see that the size of `X[]`  $n$ , reduces at a rate of  $\log_2$ . Thus, we can come to the conclusion that the function `hsum()` has the time complexity of  $\Theta(n)$ ,

## Task 5: More Running Time Analysis

### P1

```
static void method1(int[] array) {
    int n = array.length;
    for (int index=0; index<n-1; index++) {
        int marker = helperMethod1(array, index, n - 1);
        swap(array, marker, index);
    }
}

static void swap(int[] array, int i, int j) {
    int temp=array[i];
    array[i]=array[j];
    array[j]=temp;
}

static int helperMethod1(int[] array, int first, int last) {
    int max = array[first];
    int indexOfMax = first;
    for (int i=last; i>first; i--) {
        if (array[i] > max) {
            max = array[i];
            indexOfMax = i;
        }
    }
    return indexOfMax;
}
```

**Time complexity:** The best and worst case time complexity of `method1()` is  $\Theta(n^2)$ . This is because the outer loop runs  $n - 1$  times and the inner loop runs  $n - 1$  times. The swap function runs constant time as the amount of operations performed don't change as  $n$  changes, so we can ignore it in the calculation of the time complexity. Hence, we can calculate the tight upper-bound and lower-bound of the time complexity as the following:

$$\begin{aligned} &= \Theta((n-1) \cdot (n-1)) \\ &= \Theta(n \cdot n) \\ &= \Theta(n^2) \end{aligned}$$

**P2**

```
static boolean method2(int[] array, int key) {
    int n = array.length;
    for (int index=0; index<n; index++) {
        if (array[index] == key) return true;
    }
    return false;
}
```

**Time complexity:** The time complexity of `method2()` is  $\Theta(n)$ . The upper-bound/worst case time complexity is  $O(n)$  and the lower-bound/best case time complexity is  $\Omega(1)$ . Combining the lower-bound and upper-bound time complexity we get  $\Theta(n)$ .

---

**P3**

```
static double method3(int[] array) {
    int n = array.length;
    double sum = 0;
    for (int pass=100; pass >= 4; pass--) {
        for (int index=0; index < 2*n; index++) {
            for (int count=4*n; count > 0; count/=2)
                sum += 1.0*array[index/2]/count;
        }
    }
    return sum;
}
```

**Time complexity:** The best and worst case time complexity of `method3()` is  $(n \cdot \log_2 n)$ . This is because the most outer loop runs constant time, so it can be ignored. The inner loop runs  $(n)$  times because its rate to termination is linear. Lastly, the most inner loop rate to termination shortens exponentially (specifically at a rate of  $n/2$  every iteration) thus we can deduce that it runs a time complexity of  $\log_2 n$ .

Combining the separate time complexities:

$$= (n) \cdot (\log_2 n)$$

$$= (n \cdot \log_2 n)$$

Since this is the time complexities of the best case and the worst case we can conclude the time complexity for `method3()` is  $\Theta(n \cdot \log_2 n)$

## Task 6: Recursive Code

### P1

```
// assume xs.length is a power of 2
int halvingSum(int[] xs) {
    if (xs.length == 1) return xs[0];
    else {
        int[] ys = new int[xs.length/2];
        for (int i=0; i<ys.length; i++)
            ys[i] = xs[2*i]+xs[2*i+1];
        return halvingSum(ys);
    }
}
```

#### 1. Method of Measuring Problem Size:

The problem size is measured using the length of the input array `xs[]`.

#### 2. Recurrence Relation:

Let  $n = \text{length of the array } xs[]$ . The recurrence relation can be written as:

$$T(n) = \underbrace{T\left(\frac{n}{2}\right)}_{\text{recurrence}} + \underbrace{\left(\frac{n}{2}\right) \cdot c_1}_{\text{allocating new array}} + \underbrace{\left(\frac{n}{2}\right) \cdot c_2}_{\text{for loop}}; T(1) = O(1)$$
$$T(n) = T\left(\frac{n}{2}\right) + O(n); T(1) = O(1)$$

#### 3. Running Time Complexity

According to a table of common recurrences assuming  $T(0)$  and  $T(1)$  are constant the recurrence relation solves to the following:

$$T(n) = T\left(\frac{n}{2}\right) + O(n) \rightarrow O(n)$$

**P2**

```

int anotherSum(int [] xs) {
    if (xs.length == 1) return xs[0];
    else {
        int [] ys = Arrays.copyOfRange(xs, 1, xs.length);
        return xs[0] + anotherSum(ys);
    }
}

```

**1. Method of Measuring Problem Size:**

The problem size is measured using the length of the input array `xs[]`.

**2. Recurrence Relation:**

Let  $n = \text{length of the array } xs[]$ . The recurrence relation can be written as:

$$T(n) = \underbrace{T(n-1)}_{\text{recurrence}} + \underbrace{(n-1) \cdot c_1}_{\text{copying array}}; T(1) = O(1)$$

$$T(n) = T(n-1) + O(n); T(1) = O(1)$$

**3. Running Time Complexity**

According to a table of common recurrences assuming  $T(0)$  and  $T(1)$  are constant the recurrence relation solves to the following:

$$T(n) = T(n-1) + O(n) \rightarrow O(n)$$


---

**P3**

```

int [] prefixSum(int [] xs) {
    if (xs.length == 1) return xs;
    else {
        int n = xs.length;
        int [] left = Arrays.copyOfRange(xs, 0, n/2);
        left = prefixSum(left);
        int [] right = Arrays.copyOfRange(xs, n/2, n);
        right = prefixSum(right);
        int [] ps = new int[xs.length];
        int halfSum = left[left.length-1];
        for (int i=0; i<n/2; i++) { ps[i] = left[i]; }
        for (int i=n/2; i<n; i++) { ps[i] = right[i - n/2] + halfSum; }
        return ps;
    }
}

```

**1. Method of Measuring Problem Size:**

The problem size is measured using the length of the input array `xs[]`.

**2. Recurrence Relation:**

Let  $n = \text{length of the array } xs[]$ . The recurrence relation can be written as assuming that:

$$T(n) = \underbrace{T(\frac{n}{2}) + T(\frac{n}{2})}_{\text{recurrence left + right}} + \underbrace{(\frac{n}{2}) \cdot c_1 + (\frac{n}{2}) \cdot c_1}_{\text{copying array left + right}}; T(1) = O(1)$$

$$T(n) = 2T(\frac{n}{2}) + O(n); T(1) = O(1)$$

**3. Running Time Complexity**

According to a table of common recurrences assuming  $T(0)$  and  $T(1)$  are constant the recurrence relation solves to the following:

$$T(n) = 2T\left(\frac{n}{2}\right) + O(n) \rightarrow O(n \cdot \log_2 n)$$

## Task 7: Counting Dashes

To prove that  $f(n) \leq c \cdot g(n)$ , such that  $c$  and  $n$  exist in  $\mathbf{I}^+$ , and  $n \geq n_0$ , where  $f(n) = n$  and  $g(n) = n \log n$ :

We need to find values for  $c$  and  $n_0$  such that the inequality holds true.

Let's choose  $c = 1$  and  $n_0 = 1$ .

For  $n \geq 1$ , we have:

$$\begin{aligned}f(n) &= n \\g(n) &= n \log n\end{aligned}$$

Substituting these values into the inequality, we get:

$$n \leq n \log n$$

Taking the logarithm of both sides, we have:

$$\log n \leq \log(n \log n)$$

Using the logarithmic property  $\log(a \cdot b) = \log a + \log b$ , we can rewrite the right side as:

$$\log(n \log n) = \log n + \log(\log n)$$

Now, the inequality becomes:

$$\log n \leq \log n + \log(\log n)$$

Since  $\log n$  is a positive value, we can subtract  $\log n$  from both sides, resulting in:

$$0 \leq \log(\log n)$$

For  $n \geq 1$ , the logarithm of the logarithm of  $n$  is always a positive value or zero. Therefore, the inequality  $0 \leq \log(\log n)$  holds true.

We have shown that for  $c = 1$  and  $n_0 = 1$ ,  $f(n) = n$  is less than or equal to  $c \cdot g(n)$  for all  $n \geq n_0$ . Therefore, we can conclude that  $f(n) = n$  is  $O(n \log n)$ .