CSC4120 Spring 2024 - Written Homework 1

Yohandi 120040025 Andrew Nathanael 120040007

January 24, 2024

Asymptotic notation

Problem 1.

Consider the following functions:

- $f_1(n) = (\log_2(n))^2$
- $f_2(n) = \log_e(2^{\log_2(n)})$
- $f_3(n) = \log_2(n!)$
- $f_4(n) = 5^{(n + \log_2(n))}$

Questions:

- 1. Compute the tight asymptotic bounds $\Theta(\cdot)$ for each of the above functions.
- 2. Rank the above functions by increasing the order of growth. Please explicitly include the comparison in your answer.
- 1. The tight asymptotic bounds for all of the above functions are computed as follows:
 - As $\exists c_0 = c_1 = \log_2^2(b), n_0 = 2$ such that $c_0 \log_b^2(n) \le \log_2^2(n) \le c_1 \log_b^2(n), \forall n \ge n_0$. Then, $f_1(n) = \Theta(\log_b^2(n))$ by definition provided that b > 1. The asymptotic behavior is the same regardless of the logarithm's base due to the proportional relationship between logarithms of different bases, implying $f_1(n) = \Theta(\log^2 n)$.
 - $f_2(n)$ can be simplified to be $\ln(2^{\log_2(n)}) = \ln(n)$. As $\exists c_0 = c_1 = \log_b(e), n_0 = 2$ such that $c_0 \ln(n) \leq \log_b(n) \leq c_1 \ln(n), \forall n \geq n_0$. Then, $f_2(n) = \Theta(\log_b(n))$ by definition provided that b > 1. The asymptotic behavior is the same regardless of the logarithm's base due to the proportional relationship between logarithms of different bases, implying $f_2(n) = \Theta(\log n)$.
 - By Stirling's approximation, $f_3(n) = \log_2(n!) \approx \log_2(\sqrt{2\pi n}(\frac{n}{e})^n) = \frac{1}{2} + \frac{\log_2(\pi)}{2} + \frac{\log_2(n)}{2} + n\log_2(n) n\log_2(e)$.

Let $g(n) = \frac{1}{2} + \frac{\log_2(\pi)}{2} + \frac{\log_2(n)}{2} - n \log_2(e)$, then:

$$\frac{dg(n)}{dn} = \frac{1 - 2n}{2\ln(2)n}$$

Notice that g(-1) < 0 and $\frac{dg(n)}{dn} < 0$ when $n \ge 1$ due to division of negative number with positive number; therefore, $g(n) < 0, \forall n \ge 1$. This implies that $\exists c_0 = 1, n_0 = 1$ such that $f_3(n) = n \log_2(n) + g(n) \le c_0 n \log_2(n), \forall n \ge n_0$. By definition, $f_3(n) = \mathcal{O}(n \log_2 n)$.

Let $h(n) = g(n) + \frac{1}{2}n\log_2(n)$, then:

$$\frac{dh(n)}{dn} = \frac{1 - 2n}{2\ln(2)n} + \frac{\ln(n) + 1}{2\ln(2)}$$

Notice that h(3) > 0 and $\frac{dh(n)}{dn} > 0$ when $n \ge 2$ due to outgrow by $n \ln(n)$ to 2n; therefore, h(n) > 0, $\forall n \ge 3$. This implies that $\exists c_0 = \frac{1}{2}, n_0 = 3$ such that $f_3(n) = h(n) + \frac{1}{2}n \log_2(n) \ge c_0 n \log_2(n), \forall n \ge 3$. By definition, $f_3(n) = \Omega(n \log_2 n)$. As $f_3(n) = \mathcal{O}(n \log_2 n)$ and $f_3(n) = \Omega(n \log_2 n), f_3(n) = \Theta(n \log_2 n) = \Theta(n \log_2 n)$.

- $f_4(n)$ can be simplified to be $5^n \cdot (5^{\log_5(2)})^{\log_2(n^{\log_2(5)})} = 5^n \cdot n^{\log_2(5)}$. $\exists c_0 = c_1 = 1, n_0 = 2$ such that $c_0 5^n \cdot n^{\log_2(5)} \le f(n) \le c_1 5^n \cdot n^{\log_2(5)}, \forall n \ge n_0$. By definition, $f_4(n) = \Theta(5^n \cdot n^{\log_2(5)})$.
- 2. The increasing order of growth for the above functions is $f_2(n)$, $f_1(n)$, $f_3(n)$, $f_4(n)$. The comparisons are included as follows:
 - $\lim_{n\to\infty} \frac{f_2(n)}{f_1(n)} = \lim_{n\to\infty} \frac{\ln n}{\log_2^2 n} = \lim_{n\to\infty} \frac{\ln(2)}{\log_2(n)} = 0$, which indicates that $f_2(n)$'s growth is slower than $f_1(n)$'s.
 - $\lim_{n\to\infty} \frac{f_1(n)}{f_3(n)} \leq \lim_{n\to\infty} \frac{f_1(n)}{\frac{1}{2}n\log_2(n)} = \lim_{n\to\infty} \frac{2\log_2(n)}{n} = 0$ as n outgrows $\log_2(n)$. As $f_1(n) > 0$ and $f_3(n) > 0$ when n > 1, $\lim_{n\to\infty} \frac{f_1(n)}{f_3(n)} = 0$, which indicates that $f_1(n)$'s growth is slower than $f_3(n)$'s.
 - $\lim_{n\to\infty} \frac{f_3(n)}{f_4(n)} \leq \lim_{n\to\infty} \frac{n\log_2(n)}{5^n n\log_2(5)} = \lim_{n\to\infty} \frac{1}{5^n} \lim_{n\to\infty} \frac{1}{n\log_2(5)-2} \lim_{n\to\infty} \frac{\log_2(n)}{n} = 0 \cdot 0 \cdot 0 = 0$ due to n outgrowing $\log_2(n)$. As $f_3(n) > 0$ and $f_4(n) > 0$ when n > 1, $\lim_{n\to\infty} \frac{f_3(n)}{f_4(n)} = 0$, which indicates that $f_3(n)$'s growth is slower than $f_4(n)$'s.

Problem 2.

Consider the following functions:

- $\bullet \ f_1(n) = (n^{\frac{1}{\log_2(n)}})$
- $f_2(n) = (\log_2(n))!$
- $f_3(n) = \log_e(\log_e(n))$
- $f_4(n) = 2^{\sqrt{2\log_2(n)}}$

Are the statements (a) - (c) true? Indicate the reason.

- (a) The asymptotic complexity of f_1 is $\Theta(n^{\frac{1}{\log_2(n)}})$
- (b) The asymptotic complexity of f_1 is $\mathcal{O}(1)$
- (c) The asymptotic complexity of f_2 is $\Theta([(\log_2(n)]^{(\log_2(n)+\frac{1}{2})} \cdot e^{-\log_2(n)})$
- (d) Rank the above functions in terms of increasing order of growth. Please include the comparison in your answer explicitly
- (a) As $\exists c_0 = 1, c_1 = 1, n_0 = 2$ such that $c_0 n^{\frac{1}{\log_2(n)}} \leq f_1(n) \leq c_1 n^{\frac{1}{\log_2(n)}}, \forall n \geq n_0$, the statement "The asymptotic complexity of f_1 is $\Theta(n^{\frac{1}{\log_2(n)}})$ " is true by definition.

- (b) As $\exists c_0 = 2, n_0 = 1$ such that $f_1(n) = n^{\frac{1}{\log_2(n)}} = n^{\log_n(2)} = 2 \le c_0, \forall n \ge n_0$, the statement "The asymptotic complexity of f_1 is $\mathcal{O}(1)$ " is true by definition.
- (c) By Stirling's approximation, $f_2(n) = (\log_2(n))! \approx \sqrt{2\pi \log_2(n)} (\frac{\log_2(n)}{e})^{\log_2(n)}$. Let $g(n) = \log_2(n)^{\log_2(n) + \frac{1}{2}} \cdot e^{-\log_2(n)}$. As $\exists c_0 = 2, c_1 = 3, n_0 = 2$ such that $c_0 g(n) \le f(n) \le c_1 g(n), \forall n \ge n_0$, the statement "The asymptotic complexity of f_2 is $\Theta([(\log_2(n)]^{(\log_2(n) + \frac{1}{2})} \cdot e^{-\log_2(n)})$ " is true.
- (d) The increasing order of growth for the above functions is $f_1(n)$, $f_3(n)$, $f_4(n)$, $f_2(n)$. The comparisons are included as follows:
 - $\lim_{n\to\infty} \frac{f_1(n)}{f_3(n)} = \lim_{n\to\infty} \frac{n^{\frac{1}{\log_2(n)}}}{\ln \ln(n)} = \lim_{n\to\infty} \frac{2}{\ln \ln(n)} = 0$, which indicates that $f_1(n)$'s growth is slower than $f_3(n)$'s.
 - $\lim_{n\to\infty} \frac{f_3(n)}{f_4(n)} = \lim_{n\to\infty} \frac{\ln\ln(n)}{2^{\sqrt{2\log_2(n)}}} = \lim_{n\to\infty} \frac{1}{\sqrt{\ln(n)\ln(2)} \cdot 2^{\sqrt{2\log_2(n)} \frac{1}{2}}} = 0,$ which indicates that $f_3(n)$'s growth is slower than $f_4(n)$'s.
 - $\lim_{n\to\infty} \frac{f_4(n)}{f_2(n)} \le \lim_{n\to\infty} \frac{n^2}{f_2(n)}$ as $n \ge \sqrt{2} \Rightarrow \sqrt{\log_2(n^2)} \le \log_2(n^2) \Rightarrow f_4(n) \le n^2$. Claim that $\forall m \ge 13$, $(m-1)! = \Gamma(m) > (2^m)^2 = 4^m$. Prove by induction:
 - ∘ Base case: As $\Gamma(x)$ is increasing when x > 1, the claim holds true for $\forall m \in [13, 14)$, $(m-1)! \ge 12! = 479001600 > 4^{14} = 268435456$.
 - o Inductive step: Assume that $\forall m \in [k, k+1), (k-1)! > 4^{k+1}$ is true. Then $\forall m \in [k+1, k+2), k! = k \cdot (k-1)! > k \cdot 4^{k+1} > 4^{k+2}$ is also true as $k \geq 13 > 4$. Hence, proven. Let $m = \log_2(n)$, then $(\log_2(n) 1)! > 4^{\log_2(n)} = n^2$. Then, $\lim_{n \to \infty} \frac{n^2}{\log_2(n)!} = \lim_{n \to \infty} \frac{1}{\log_2(n)} \frac{n^2}{(\log_2(n)-1)!} = 0 \cdot c = 0$ for some finite constant c. This is supported by the fact that $\lim_{n \to \infty} \log_2(n) = +\infty \geq 13$. The previous implies that $\lim_{n \to \infty} \frac{f_4(n)}{f_2(n)} \leq 0$. As $f_4(n) > 0$ and $f_2(n) > 0$ when n > 2, $\lim_{n \to \infty} \frac{f_4(n)}{f_2(n)} = 0$, which indicates that $f_4(n)$'s growth is slower than $f_2(n)$'s.

Problem 3.

Let f(n) and g(n) be asymptotically non-negative functions. Prove that if $h(n) = \max\{f(n), g(n)\}$, then $h(n) = \Theta(f(n) + g(n))$.

Suppose $h(n) \neq \Theta(f(n) + g(n))$, which implies that there is no $c_0, c_1 > 0, n_0$ such that $c_0(f(n) + g(n)) \leq h(n) \leq c_1(f(n) + g(n)), \forall n \geq n_0$. However, if h(n) were to be $\max\{f(n), g(n)\}$:

- If $f(n) \leq g(n)$, then $\exists c_0 = \frac{1}{2}, c_1 = 1, n_0$ that satisfies:
 - $\circ \frac{1}{2}(f(n) + g(n)) \le \frac{1}{2}2g(n) = g(n) = \max\{f(n), g(n)\} = h(n), \forall n \ge n_0$
 - $\circ \ h(n) = \max\{f(n), g(n)\} = g(n) \le f(n) + g(n), \forall n \ge n_0$

• If f(n) > g(n), then $\exists c_0 = \frac{1}{2}, c_1 = 1, n_0$ that satisfies:

$$\circ \frac{1}{2}(f(n) + g(n)) \le \frac{1}{2}2f(n) = f(n) = \max\{f(n), g(n)\} = h(n), \forall n \ge n_0$$

$$h(n) = \max\{f(n), g(n)\} = f(n) \le f(n) + g(n), \forall n \ge n_0$$

Those contradict our supposition about $h(n) \neq \Theta(f(n) + g(n))$; consequently, h(n) can't be $\max\{f(n), g(n)\}$. As it has been shown that $h(n) \neq \Theta(f(n) + g(n)) \Rightarrow h(n) \neq \max\{f(n), g(n)\}$, the equivalent claim follows $h(n) = \max\{f(n), g(n)\} \Rightarrow h(n) = \Theta(f(n) + g(n))$; hence, proven.

Problem 4.

Suppose that our list of size n is being sorted with insertion sort, and each element is at most k away from its final position (k-sorted list). What is the asymptotic complexity if

a)
$$k = O(1)$$
,

b)
$$k = \frac{n}{2}$$
, and why?

Insertion sort works as follows:

- 1. Iterates from 1 to n-1 and compare each element to its predecessor.
- 2. If the comparison shows the correct sort, then the algorithm ends.
- 3. Otherwise, shift the position of the found element until it finds the smaller element as its predecessor.
- 4. Repeat.
- (a) As each element is at most k away from its final position and $\exists c_0, n_0$ such that $k \leq c_0, \forall n \geq n_0$, we can conclude that the asymptotic complexity is $\mathcal{O}(n)$. This is true because there will be at most k comparisons for each of the n-1 elements. Let g(n) be such a function, then choose c_0, n_0 that establishes the definition of k being $\mathcal{O}(1)$, implying $g(n) \leq k(n-1) \leq kn \leq c_0 n, \forall n \geq n_0$. By definition, $g(n) = \mathcal{O}(n)$.
- (b) In a similar way, as each element is at most $k = \frac{n}{2}$ away from its final position, we can conclude that the asymptotic complexity is $\mathcal{O}(n^2)$. As $\exists c_1 = \frac{1}{2}, n_1$ such that $k \leq c_1 n, \forall n \geq n_1$, the previous conclusion is true because there will be at most k comparisons for each of the n-1 elements. Let h(n) be such a function, then choose c_1, n_1 that establishes the definition of k being $\mathcal{O}(n)$, implying $h(n) \leq k(n-1) \leq kn \leq c_1 n^2, \forall n \geq n_1$. By definition, $h(n) = \mathcal{O}(n^2)$. Notice that the asymptotic complexity uses \mathcal{O} instead of Θ as the problem uses the term "at most".

Recurrences

Problem 5.

Solve $T(n) = T(\sqrt{n}) + 1$ and give a Θ bound.

Let $n = 2^m$ and $S(m) = T(2^m)$, then:

$$T(2^m) = T(2^{\frac{m}{2}}) + 1 \Rightarrow S(m) = S\left(\frac{m}{2}\right) + 1 \underset{\text{by Master's Theorem}}{\Rightarrow} S(m) = \Theta(\log m)$$

Therefore, $T(n) = T(2^m) = S(m) = \Theta(\log m) = \Theta(\log \log n)$.

Problem 6.

Are the following statements true? Indicate the reason.

- (a) The solution of T(n) = T(n-1) + n is $\mathcal{O}(n^2)$.
- (b) The solution of T(n) = 2T(n-1) + 1 is $\mathcal{O}(n \log n)$ or $\mathcal{O}(2^n)$.
- (c) The recurrence $T(n) = T(\frac{n}{3}) + T(\frac{2n}{3}) + cn$ (c > 0) is
 - I. $\Omega(n \log n)$
 - II. $\mathcal{O}(n \log n)$
- (d) The recurrence $T(n) = 4T(\frac{n}{2}) + cn^2\sqrt{n}$ (c > 0) is
 - I. $\Theta(n^2)$
 - II. $\Theta(n^2\sqrt{n})$
- (e) The recurrence $T(n) = T(\frac{n}{2}) + T(\frac{n}{4}) + T(\frac{n}{8}) + n$ is
 - I. $\Theta(n \log n)$
 - II. $\Theta(n)$
- (a) The expansion of T(n) is as follows:

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

$$= T(n-2) + (n-1) + n$$

$$\vdots$$

$$= \Theta(1) + 2 + \dots + (n-1) + n$$

$$= \mathcal{O}(1) + 2 + \dots + (n-1) + n$$

This implies $\exists c_0, n_0$ such that:

$$T(n) \le c_0 + 2 + \dots + (n-1) + n, \forall n \ge n_0$$

= $c_0 - 1 + \frac{n(n+1)}{2}, \forall n \ge n_0$
 $\le c_0 + n^2, \forall n \ge \max(1, n_0)$

We can choose $c_1 = \max(1, c_0) + \lim_{\epsilon \to 0^+} \epsilon$ and $n_1 = \max(1, n_0, \sqrt{\frac{c_1}{c_1 - 1}})$ such that $T(n) \leq c_1 n^2, \forall n \geq n_1 \underset{\text{by definition}}{\Rightarrow} T(n) = \mathcal{O}(n^2)$. This further implies that the statement "The solution of T(n) = T(n-1) + n is $\mathcal{O}(n^2)$ " is true.

(b) The expansion of T(n) is as follows:

$$T(n) = 2T(n-1) + 1$$

$$= 4T(n-2) + 2 + 1$$

$$\vdots$$

$$= 2^{n-1}\Theta(1) + 2^{n-2} + \dots + 2 + 1$$

• $T(n) = 2^{n-1}\Theta(1) + 2^{n-2} + \ldots + 2 + 1$ implies $\exists c_0, n_0$ such that:

$$T(n) \ge c_0 2^{n-1} + 2^{n-2} + \dots + 2 + 1, \forall n \ge n_0$$

 $\ge 2^{n-2}, \forall n \ge n_0$

For a function $f(n) = 2^{n-2} - cn \log n$, it can be claimed that $\lim_{n\to\infty} f(n) = +\infty$ as $\lim_{n\to\infty} \frac{2^{n-2}}{cn \log n} = +\infty$ provided that c > 0. When $c \le 0$, f(n) is simply a positive function due to the addition of two positive values. Both cases lead to a conclusion where $2^{n-2} > cn \log n$ when $n \to \infty$; therefore, one can't find such c_1, n_1 such that $T(n) \le c_1 n \log n$, $\forall n \ge n_1$. The statement "The solution of T(n) = 2T(n-1) + 1 is $\mathcal{O}(n \log n)$ " is false.

• $T(n) = 2^{n-1}\Theta(1) + 2^{n-2} + \ldots + 2 + 1$ implies $\exists c_0, n_0$ such that:

$$T(n) \le c_0 2^{n-1} + 2^{n-2} + \dots + 2 + 1, \forall n \ge n_0$$

$$\le \max(1, c_0) (2^{n-1} + 2^{n-2} + \dots + 2 + 1), \forall n \ge n_0$$

$$\le \max(1, c_0) 2^n, \forall n \ge n_0$$

We can choose $c_1 = \max(1, c_0)$ and $n_1 = \max(1, n_0)$ such that $T(n) \le c_1 2^n, \forall n \ge n_1 \Longrightarrow_{\text{by definition}} T(n) = \mathcal{O}(2^n)$. This further implies that the statement "The solution of T(n) = 2T(n-1) + 1 is $\mathcal{O}(2^n)$ " is true.

(c) Both statements will be proved using the substitution method.

I. Claim that $T(n) \geq c_0 n \log n, \forall n \geq n_0$. Then,

$$T(n) = T(\frac{n}{3}) + T(\frac{2n}{3}) + cn$$

$$\geq c_0 \frac{n}{3} \log(\frac{n}{3}) + c_0 \frac{2n}{3} \log(\frac{2n}{3}) + cn, \forall n \geq n_0$$

$$\geq c_0 n \log(\frac{n}{3}) + cn, \forall n \geq n_0$$

$$\geq c_0 n \log n - c_0 n \log 3 + cn, \forall n \geq n_0$$

The claim is correct when $cn \geq c_0 n \log 3 \Rightarrow c_0 \leq \frac{c}{\log 3}$. The statement "The recurrence $T(n) = T(\frac{n}{3}) + T(\frac{2n}{3}) + cn$ (c > 0) is $\Omega(n \log n)$ " is true as $\exists c_0$ (such as $\frac{c}{\log 3}$) and $n_0 > 1$ that holds the claim.

II. Claim that $T(n) \leq c_0 n \log n, \forall n \geq n_0$. Then,

$$T(n) = T(\frac{n}{3}) + T(\frac{2n}{3}) + cn$$

$$\leq c_0 \frac{n}{3} \log(\frac{n}{3}) + c_0 \frac{2n}{3} \log(\frac{2n}{3}) + cn, \forall n \geq n_0$$

$$\leq c_0 n \log(\frac{2n}{3}) + cn, \forall n \geq n_0$$

$$\leq c_0 n \log n + c_0 n \log 2 - c_0 n \log 3 + cn, \forall n \geq n_0$$

The claim is correct when $c_0 n \log 2 + cn \le c_0 n \log 3 \Rightarrow c_0 \ge \frac{c}{\log 3 - \log 2}$. The statement "The recurrence $T(n) = T(\frac{n}{3}) + T(\frac{2n}{3}) + cn \ (c > 0)$ is $\mathcal{O}(n \log n)$ " is true as $\exists c_0$ (such as $\frac{c}{\log 3 - \log 2}$) and $n_0 > 1$ that holds the claim.

(d) By Master's Theorem, $T(n) = \Theta(n^2\sqrt{n})$. Therefore, the statement "The recurrence $T(n) = 4T(\frac{n}{2}) + cn^2\sqrt{n}$ (c > 0) is $\Theta(n^2\sqrt{n})$ " is true.

The previous implies that $\exists c_0, n_0$ such that $T(n) \geq c_0 n^2 \sqrt{n}, \forall n \geq n_0$. Assume that $\exists c_1, n_1$ such that $T(n) \leq c_1 n^2, \forall n \geq n_1$. For $n > \max(1, n_0, n_1, (\frac{c_1}{c_0})^2)$, the condition $c_0 n^2 \sqrt{n} > c_1 n^2$ invalidates the possibility of T(n) satisfying both $\Omega(n^2 \sqrt{n})$ and $\mathcal{O}(n^2)$ simultaneously. Consequently, the statement "The recurrence $T(n) = 4T(\frac{n}{2}) + cn^2 \sqrt{n}$ (c > 0) is $\Theta(n^2)$ " must be false.

(e) Claim that $T(n) \leq c_0 n, \forall n \geq n_0$. By substitution method,

$$T(n) = T(\frac{n}{2}) + T(\frac{n}{4}) + T(\frac{n}{8}) + n$$

$$\leq c_0 \frac{n}{2} + c_0 \frac{n}{4} + c_0 \frac{n}{8} + n$$

$$\leq \frac{7c_0}{8}n + n$$

The claim is correct when $\frac{7c_0}{8}n + n \le c_0n \Rightarrow c_0 \ge 8$. Therefore, $T(n) = \mathcal{O}(n)$ as $\exists c_0$ (such as $c_0 = 8$) and $n_0 > 1$ that holds the claim.

Assume $\exists c_1, n_1$ such that $T(n) \geq c_1 n \log n, \forall n \geq n_1$. For $n > \max(1, n_0, n_1, \frac{c_0}{c_1})$, the condition $c_1 n \log n > c_0 n$ invalidates the possibility of T(n) satisfying both $\mathcal{O}(n)$ and $\Omega(n \log n)$ simultaneously. Consequently, the statement "The recurrence $T(n) = T(\frac{n}{2}) + T(\frac{n}{4}) + T(\frac{n}{8}) + n$ is $\Theta(n \log n)$ " must be false.

Similarly, claim that $T(n) \geq c_1 n, \forall n \geq n_1$. By substitution method,

$$T(n) = T(\frac{n}{2}) + T(\frac{n}{4}) + T(\frac{n}{8}) + n$$

$$\geq c_1 \frac{n}{2} + c_1 \frac{n}{4} + c_1 \frac{n}{8} + n$$

$$\geq \frac{7c_1}{8}n + n$$

The claim is correct when $\frac{7c_1}{8}n + n \ge c_1n \Rightarrow c_1 \le 8$. Therefore, $T(n) = \Omega(n)$ as $\exists c_1$ (such as $c_1 = 8$) and $n_1 > 1$ that holds the claim. $T(n) = \mathcal{O}(n)$ and $T(n) = \Omega(n)$ is equivalent to $T(n) = \Theta(n)$. With that, the statement "The recurrence $T(n) = T(\frac{n}{2}) + T(\frac{n}{4}) + T(\frac{n}{8}) + n$ is $\Theta(n)$ " is true.