CS4413 Algorithm Analysis

Dr. Lakshmivarahan

1 Summation

1.1 $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}, \quad \int_0^n x dx = \frac{n^2}{2}$

1.2 $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}, \quad \int_0^n x^2 dx = \frac{n^3}{3}$

1.3 $\sum_{i=1}^{n} i^3 = \left[\frac{n(n+1)}{2}\right]^2, \quad \int_0^n x^3 dx = \frac{n^4}{4}$

1.4 $\sum_{i=1}^{n} x^i = \begin{cases} \frac{1-x^{n+1}}{1-x} & \text{if } |x| < 1 \\ \frac{x^{n+1}-1}{x-1} & \text{if } |x| > 1 \end{cases}$

1.5 $\sum_{i=1}^{\infty} x^i = \frac{1}{1-x} \quad \text{if } |x| < 1, \quad \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i = \frac{1}{(1-\frac{1}{2})} = 2$

1.6 $\sum_{i=1}^{n} \frac{1}{i} \approx \log n, \quad \int_1^n \frac{1}{x} dx = \log_e x$

1.7 $\sum_{i=1}^{n} \frac{1}{i^2} \approx \int_1^n \frac{dx}{x^2} = \frac{x^{-1}}{-1}\Big|_1^n = -\left[\frac{1}{n} - 1\right] = 1 - \frac{1}{n}$

$$\implies \quad \sum_{i=1}^{\infty} \frac{1}{i^2} < 1$$

1.8 $$\sum_{i=1}^{k} i2^i = \sum_{i=1}^{k} i[2^{i+1} - 2^i]$$

$$= \sum_{i=i}^{k} i2^{i+1} - \sum_{i=1}^{k} i2^i = \sum_{i=1}^{k} i2^{i+1} - \sum_{j=0}^{k-1} (j+1)2^{j+1}$$

$$= \sum_{i=i}^{k} i2^{i+1} - \sum_{j=0}^{k-1} j2^{j+1} - \sum_{j=0}^{k-1} 2^{j+1} \qquad (*)$$

But

$$\sum_{j=0}^{k-1} 2^{j+1} = 2 + 2^2 + \ldots + 2^k$$

$$= (1 + 2 + \ldots + 2^k) - 1 = \tfrac{2^{(k+1)}-1}{2-1} - 1 = 2^{k+1} - 2 \quad (**)$$

Also

$$\sum_{i=i}^{k} i2^{i+1} - \sum_{j=0}^{k-1} j2^{j+1} = k2^{k+1} \qquad (***)$$

Substituting (**) and (***) into (*)

$$\sum_{i=1}^{k} i2^i = k2^{k+1} - (2^{k+1} - 2) = (k-1)2^{k+1} + 2$$

1.9 Average complexity of binary search

$A(n) = \frac{1}{2n+1}\left[\sum_{t=1}^{k} t2^{t-1} + k(n+1)\right]$

From 1.8:

$\sum_{t=1}^{k} t2^{t-1} = \frac{1}{2}\sum_{t=1}^{k} t2^t = (k-1)2^k + 1$

Since $n = 2^k - 1$ in binary search

$$
\begin{aligned}
A(n) &= \tfrac{(k-1)2^k+1}{(2n+1)} + \tfrac{k(n+1)}{2n+1} \\
&= \tfrac{(k-1)2^k+1}{2^{k+1}-1} + \tfrac{k2^k}{2^{k+1}-1} \\
&= \tfrac{(2k-1)2^k+1}{2^{k+1}-1} = \tfrac{(2k-1)2^k}{2^{k+1}-1} + \tfrac{1}{2^{k+1}-1}
\end{aligned}
$$

But $\frac{2^k}{2^{k+1}-1} \approx \frac{2^k}{2^{k+1}} = \frac{1}{2}$ and $\frac{1}{2^{k+1}-1} \longrightarrow 0$ as $k \longrightarrow \infty$

$\therefore A(n) \approx \frac{(2k-1)}{2} = k - \frac{1}{2} \approx \log n$

2 Exponentials and logarithms

2.1 $a^0 = 1,\ a^1 = a,\ a^{-1} = \frac{1}{a}$

2.2 $(a^m)^n = a^{mn} = (a^n)^m$

2.3 $a^m \cdot a^n = a^{m+n}$

2.4 $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^k}{k!} + \ldots = \displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$

2.5 $\log_x ab = \log_x a + \log_x b,\ x \in \{2, e, 10\}$

2.6 $\log_x a^n = n \log_x a$

2.7 $\log_x a = \log_y a \cdot \log_x y = \dfrac{\log_y a}{\log_y x}$ $\boxed{\text{Change of base}}$

2.8 $\log^k n = (\log n)^k$ $\boxed{\text{Exponentiation}}$

2.9 $\log^{(2)} n = \log \log n = \log(\log(n))$ $\boxed{\text{Composition or iterated logarithm}}$

2.10 $a = b^{(\log_b a)}$

2.11 $\log_x(\frac{1}{a}) = -\log_x a$

2.12 $\log_b a = \frac{1}{\log_a b}$

2.13 $a^{\log_b n} = n^{\log_b a}$

2.14 $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5}$

2.15 $\frac{x}{1+x} \le \log(1+x) \le x, \, \forall x > -1$

2.16 Iterated logarithm:

$$\log_2^* n = min\{i \ge 0 | \log_2^{(i)} n \le 1\}$$

$$\log_2^* 2 = 1 \qquad 2 = 2$$

$$\log_2^* 4 = 2 \qquad 4 = 2^2$$

$$\log_2^* 16 = 3 \qquad 16 = 2^{2^2} = 2^4$$

$$\log_2^* 65,536 = 4 \qquad 65,536 = 2^{2^{2^2}} = 2^{2^4} = 2^{16}$$

$$\log_2^* 2^{65,536} = 5$$

3 Factorial and approximations

3.1 $n! = 1 \cdot 2 \cdot 3 \ldots n$

3.2 $n! = n(n-1)!$

3.3 $n! = 1$ if $n = 0$

3.4 $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ $\boxed{\text{Stirlings approximation}}$

3.5 $$\log_e n! = \sum_{i=1}^{n} \log_e i \ \ \ge \int_1^n \log_e x dx$$

$$= x \log_e x|_1^n - \int_1^n x d(\log_e x) = n \log_e n - n + 1$$

$$\therefore \log_e n! \ge n \log_e n - n$$

$$\log_2 n! = \log_2 e \log_e n! = n\left(\log_2 e \log_e n\right) - n \log_2 e$$

$= n \log_2 n - 1.5n$

Since $\log_2 e \approx 1.44 (e = 2.71\ldots)$

4 Fibonacci numbers

$F_n = F_{n-1} + F_{n-2}, \quad F_0 = 0, F_1 = 1$

$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots$

$F_n = \frac{1}{\sqrt{5}} \left(\phi^n - \hat{\phi}^n \right)$

$\phi = \frac{1+\sqrt{5}}{2} = 1.61803$ $\boxed{\text{Golden ratio}}$

$\hat{\phi} = \frac{1-\sqrt{5}}{2} = -0.61803$

5 Solution of recurrences

$$T(n) = aT(\tfrac{n}{b}) + cn \boxed{n = b^k, \;\; T(1) = d} \quad (1)$$

$T(\frac{n}{b}) = aT(\frac{n}{b^2}) + c(\frac{n}{b})$

$T(\frac{n}{b^2}) = aT(\frac{n}{b^3}) + c(\frac{n}{b^2})$

$T(\frac{n}{b^3}) = aT(\frac{n}{b^4}) + c(\frac{n}{b^3})$

...

Substituting back, we get

$$
\begin{aligned}
T(n) &= a^2T(\tfrac{n}{b^2}) + ac(\tfrac{n}{b}) + cn \\
&= a^3T(\tfrac{n}{b^3}) + a^2c(\tfrac{n}{b^2}) + ac(\tfrac{n}{b}) + cn \\
&= a^4T(\tfrac{n}{b^4}) + a^3c(\tfrac{n}{b^3}) + a^2c(\tfrac{n}{b^2}) + ac(\tfrac{n}{b}) + cn \\
&= a^4T(\tfrac{n}{b^4}) + \sum_{j=0}^{3} \left(\frac{a}{b}\right)^j cn \\
&= \dots \\
&= a^rT(\tfrac{n}{b^r}) + \sum_{j=0}^{r-1} \left(\frac{a}{b}\right)^j cn \qquad (2)
\end{aligned}
$$

when is $\frac{n}{b^r} = 1? \iff \frac{b^k}{b^r} = 1 \quad$ (i.e.) $r = k$

$$
\begin{aligned}
\therefore T(n) &= a^kT(\tfrac{n}{b^k}) + \sum_{j=0}^{k-1} (\frac{a}{b})^j cn \\
&= a^kd + \sum_{j=0}^{k-1} \alpha^j (cn) \ \boxed{\alpha = \tfrac{a}{b}} \quad (3)
\end{aligned}
$$

But $a^k = a^k(\frac{n}{b^k}) = \left(\frac{a^k}{b^k}\right) n = \alpha^k n$

$$
\begin{aligned}
\therefore T(n) &= \alpha^k(dn) + \sum_{j=0}^{k-1} \alpha^j (cn) \\
&= (\alpha n)(\alpha^k) + cn(\sum_{j=0}^{k-1} \alpha^j)
\end{aligned}
$$

<u>Recall:</u>

$\min\{c,d\}(nk+nm) \leq dnk + cnm \leq \max\{c,d\}(nk+nm)$

Hence

$$T(n) \quad \le \max\{c,d\}\left[n\sum_{j=0}^{k}\alpha^j\right] \qquad (4)$$

$$= c_1 n\sum_{j=0}^{k}\alpha^j, \text{where } \ c_1 = \max\{c,d\}$$

$$\alpha : \begin{cases} \text{case 1: } \ \alpha < 1 \\ \text{case 2: } \ \alpha = 1 \\ \text{case 3: } \ \alpha > 1 \end{cases}$$

Case 1: $\alpha < 1$

$$T(n) \le c_1 n\sum_{j=0}^{k}\alpha^j \quad \le c_1 n\sum_{j=0}^{\infty}\alpha^j$$

$$= \tfrac{c_1 n}{1-\alpha} = \left(\tfrac{c_1}{1-\alpha}\right)n = O(n) \qquad (5)$$

Case 2: $\alpha = 1$

$$T(n) \le c_1 n\sum_{j=0}^{k}1 = c_1(k+1)n$$

But $\quad n = b^k \Longrightarrow k = \log_b n$

$$\therefore T(n) \le c_1(1+\log_b n) = c_1 n\log_b n + c_1 n = O(n\log_b n) \quad (6)$$

Case 3: $\alpha > 1$

$$T(n) \quad \le c_1 n\sum_{j=0}^{k}\alpha^j = c_1 n\left(\frac{\alpha^{k+1}-1}{\alpha-1}\right)$$

$$= \left(\tfrac{c_1}{\alpha-1}\right)n\alpha^{k+1} - \left(\tfrac{c_1}{\alpha-1}n\right) \qquad (7)$$

But $\alpha^{k+1} = \alpha \cdot \alpha^k = \alpha \cdot (\frac{a}{b})^k = \alpha \frac{a^k}{b^k} = \alpha(\frac{a^k}{n})$

Substituting into (7)

$$
\begin{aligned}
T(n) &= (\tfrac{c_1}{\alpha-1})n \cdot \alpha \tfrac{a^k}{n} - (\tfrac{c_1}{\alpha-1})n \\
&= \left(\tfrac{\alpha c_1}{\alpha-1}\right) a^k - \left(\tfrac{c_1}{\alpha-1}\right) n \qquad (8)
\end{aligned}
$$

But $a^k = a^{\log_b n} = n^{\log_b a}$

Since $a > b \Longrightarrow \boxed{\log_b a > 1}$

$\therefore T(n) = c_2 n^{\log_b a} - c_3 n = O(n^{\log_b a})$

Where $c_2 = \left(\frac{\alpha c_1}{\alpha-1}\right)$ and $c_3 = \left(\frac{c_1}{\alpha-1}\right)$

For example, when $a = 4, b = 2, \log_b a = 2$,and $T(n) = O(n^2)$

Summary

$$
\begin{cases}
\alpha < 1, T(n) = O(n) \\
\alpha = 1, T(n) = n \log n \\
\alpha > 1, T(n) = O(n^{\log_b a})
\end{cases}
$$

<u>Example:</u>

$T(n) = 2T(\frac{n}{2}) + (n-1)$ $\boxed{n = 2^k, T(1) = 0}$

$T(\frac{n}{2}) = 2T(\frac{n}{2^2}) + (\frac{n}{2} - 1)$

$T(\frac{n}{2^2}) = 2T(\frac{n}{2^3}) + (\frac{n}{2^2} - 1)$

$$\therefore T(n) = 2^3 T(\tfrac{n}{2^3}) + \sum_{i=0}^{2} 2^i(\frac{n}{2^i} - 1) = 2^r T(\frac{n}{2^r}) + \sum_{i=0}^{r-1} 2^i(\frac{n}{2^i} - 1)$$

when is $\frac{n}{2r} = 1? \Longleftrightarrow \frac{2^k}{2^r} = 1 \Longrightarrow \boxed{\text{r = k}}$

$$\therefore T(n) = 2^k T(\tfrac{n}{2^k}) + \sum_{i=0}^{k-1} 2^i \left(\frac{n}{2^i} - 1\right)$$

Since $T(1) = 0$ $\boxed{k = \log_2 n}$, we get

$$T(n) = \sum_{i=0}^{k-1} 2^i \left(\frac{n}{2^i} - 1 \right) = \sum_{i=0}^{k-1} n - \sum_{i=0}^{k-1} 2^i$$

$$= nk - \left(\tfrac{2^k - 1}{2 - 1} \right) = n \log_2 n - (2^k - 1)$$

$$= n \log_2 n - n + 1 = O(n \log_2 n)$$

6 Solution of complete history recurrence

Let $A(n) = \frac{2}{n} \sum_{i=2}^{n-1} A(i) + (n-1) \quad (1)$

$\Longrightarrow A(n-1) = \frac{2}{n-1} \sum_{i-2}^{n-2} A(i) + (n-2) \quad (2)$

$n * (1) - (n-1) * (2) \Longrightarrow$

$nA(n) - (n-1)A(n-1)$

$= \left[2 \sum_{i=2}^{n-1} A(i) + n(n-1) \right] - \left[2 \sum_{i=2}^{n-2} A(i) + (n-1)(n-2) \right]$

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

Divide both sides by n(n+1):

$$\tfrac{A(n)}{n+1} - \tfrac{(n-1)}{n(n+1)} A(n-1) = \tfrac{2}{n(n+1)} A(n-1) + \tfrac{2n-2}{n(n+1)}$$

$$\tfrac{A(n)}{n+1} = A(n-1) \left[\tfrac{n-1}{n(n+1)} + \tfrac{2}{n(n+1)} \right] + \tfrac{2n-2}{n(n+1)}$$

$$= \tfrac{A(n-1)}{n} + \tfrac{2n-2}{n(n+1)} \qquad (3)$$

Set $\frac{A(n)}{n+1} = B(n)$. Then (3) becomes:

$B(n) = B(n-1) + \frac{2n-2}{n(n+1)}$

$$\implies B(n) = \sum_{i=2}^{n} \frac{2i-2}{i(i+1)}$$

$$= 2\sum_{i=2}^{n} \frac{1}{(i+1)} - 2\sum_{i=2}^{n} \frac{1}{i(i+1)}$$

$$\sum_{i=2}^{n} \frac{1}{(i+1)} = \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{5} + \ldots + \tfrac{1}{n}$$

$$\leq 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{5} + \ldots + \tfrac{1}{n}$$

$$= \sum_{i=1}^{n} \frac{1}{(i)} \approx \log n$$

$$\sum_{i=2}^{n} \frac{1}{i(i+1)} \leq \sum_{i=2}^{n} \frac{1}{i^2} \quad \boxed{i(i+1) > i^2,\ \tfrac{1}{i(i+1)} < \tfrac{1}{i^2}}$$

$$\leq \sum_{i=1}^{n} \frac{1}{i^2} \approx \int_{1}^{n} \frac{1}{x^2} dx$$

$$= \textstyle\int_1^n x^{-2} dx = \frac{x^{-1}}{-1}\Big|_1^n$$

$$= -\left[\tfrac{1}{n} - 1\right] = \left[1 - \tfrac{1}{n}\right] \leq 1$$

$\therefore B(n) \approx 2\log n - 2 \leq 2\log n$

$\therefore A(n) = (n+1)B(n) = 2n\log n + 2\log n = O(n\log n)$

7 Properties of binary trees

7.1 If L is the number of leaves in a binary tree of depth d (≥ 0), then L $\leq 2^d$

7.2 If L is the number of leaves in a binary tree of depth d, then d $\geq \lceil \log L \rceil$

7.3 Among 2-trees with L leaves, the EPL is minimum when all the leaves are on at most two levels.
Let k $\leq$ d-2 and d be the depth.

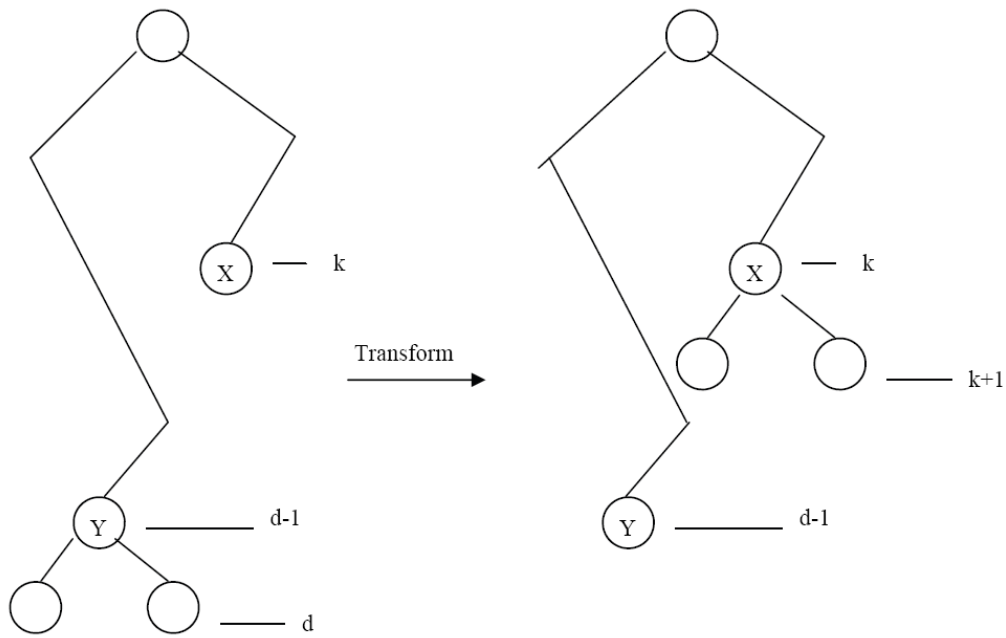

Figure 1:

In figure 1, tree with L leaves and depth d

- Decrease in EPL = 2d + k
- Increase in EPL = (d-1) + 2(k+1)
- Net decrease = (2d + k) - (d-1) - 2(k+1) = 2d + k -d + 1 - 2k -2 = d -k -1 > 0 $\boxed{\text{since } k \leq d\text{-}2}$

7.4 The minimum value of EPL in a 2-tree with L leaves is

$L\lfloor \log L \rfloor + 2[L - 2^{\lfloor \log L \rfloor}]$. In figure 2

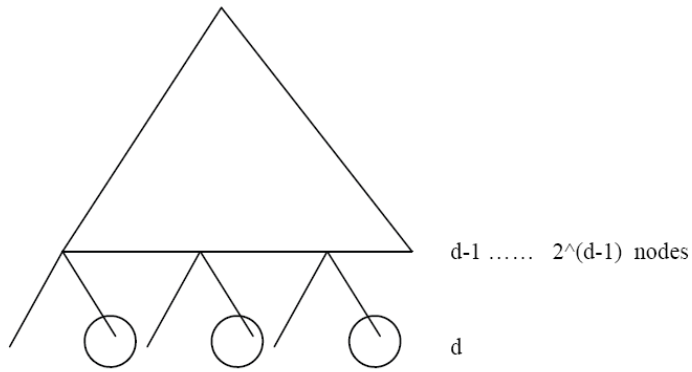

Figure 2:

- If L is a power of 2, then all the leaves are at level $\log L$. Thus, the EPL = $L \log L$
  In this case $L - 2^{\lfloor \log L \rfloor} = 0$
- If L is <u>not</u> a power of 2, then $2^{d-1} < L \leq 2^d$ and all leaves are at two adjacent leavels. Then d = $\lceil \log L \rceil$

See figure 3.
The number of leaves at level d = 2(L-2^{d-1})

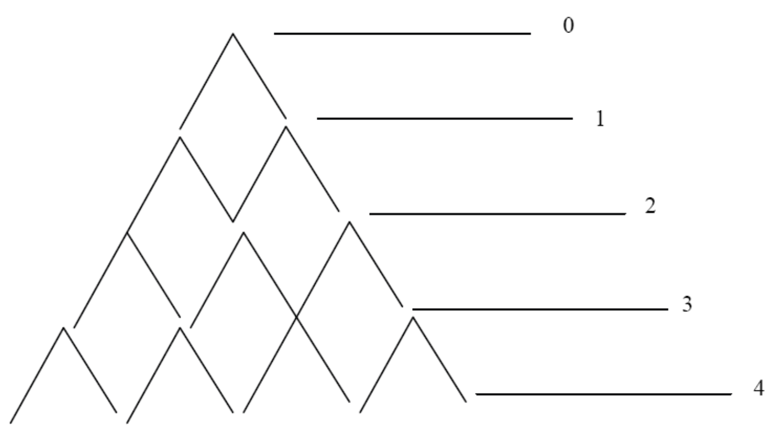

Figure 3:

Let L = 14, then

$\lceil \log L \rceil = 4, L - 2^{d-1} = 14 - 8 = 6$

$\therefore 2(L - 2^{d-1}) = 12$

$\therefore EPL = L(d-1) + 2(L - 2^{d-1})$

$= L\lfloor \log L \rfloor + 2[L - 2^{\lfloor \log L \rfloor}]$

7.5 The average EPL in a 2-tree with L-leaves is at least $\lfloor \log L \rfloor$

$\frac{EPL}{L} = \lfloor \log L \rfloor + \varepsilon$

when $\varepsilon = \frac{2[L - 2^{\lfloor \log L \rfloor}]}{L}$

But $L - 2^{\lfloor \log L \rfloor} < \frac{L}{2}$ $\boxed{L = 14, \lfloor \log L \rfloor = 3, 14 - 2^3 = 6 < 7}$

$\therefore 0 \leq \varepsilon < 1$

$\Longrightarrow$ Average EPL $\geq \lfloor \log L \rfloor$

7.6 Let L = n! Then the average EPL $\geq \log n! \approx \lfloor n \log n - 1.5n \rfloor$

Remark: No algorithm can perform better than quick sort on the average.