Nhập môn phân tích độ phức tạp thuật toán-21TN

TRẦN MINH HOÀNG-21120075

BÀI TẬP LÝ THUYẾT LẦN 1

Câu 1.1 Tại sao có thể nói hàm $f(n) = n^{\sqrt{n}}$ thì độ lớn nằm giữa "lớp các hàm đa thức n^a(với a>0)" và "lớp các hàm mũ c^n (với c>1)"?

Giải: Để chứng minh $\begin{cases} f(n) \in \omega(n^a) \\ f(n) \in o(c^n) \end{cases}$ thì ta cần chứng minh $\begin{cases} \lim\limits_{n\to\infty} \frac{f(n)}{n^a} = \infty \\ \lim\limits_{n\to\infty} \frac{f(n)}{c^n} = 0 \end{cases}$

- $\lim\limits_{n\to\infty} \frac{f(n)}{n^a} = \lim\limits_{n\to\infty} \frac{n^{\sqrt{n}}}{n^a} = \lim\limits_{n\to\infty} n^{(\sqrt{n}-a)} = \infty^\infty = \infty$. (1)
- $\lim\limits_{n\to\infty} \frac{f(n)}{c^n} = \lim\limits_{n\to\infty} \frac{n^{\sqrt{n}}}{c^n} = \lim\limits_{n\to\infty} \frac{e^{\ln(n^{\sqrt{n}})}}{e^{\ln(c^n)}} = \lim\limits_{n\to\infty} \frac{e^{\sqrt{n}\ln(n)}}{e^{n\ln(c)}} = e^{\lim\limits_{n\to\infty}(\sqrt{n}\ln(n) - n\ln(c))}$. $(*)$

  Ta có: $\lim\limits_{n\to\infty}(\sqrt{n}\ln(n) - n\ln(c)) = \lim\limits_{n\to\infty} \sqrt{n}\ln(n)(1 - \frac{n\ln(c)}{\sqrt{n}\ln(n)}) = \lim\limits_{n\to\infty} \sqrt{n}\ln(n)(1 - \frac{\sqrt{n}\ln(c)}{\ln(n)})$.

  Lại có: $\lim\limits_{n\to\infty} \frac{\sqrt{n}\ln(c)}{\ln(n)} \underset{\infty/\infty}{\overset{L'H}{=}} \lim\limits_{n\to\infty} \frac{\frac{1}{2\sqrt{n}}\ln(c)}{\frac{1}{n}} = \frac{\sqrt{n}\ln(c)}{2} = \infty$.

  Vậy $\lim\limits_{n\to\infty}(1 - \frac{\sqrt{n}\ln(c)}{\ln(n)}) = -\infty \Rightarrow \lim\limits_{n\to\infty} \sqrt{n}\ln(n)(1 - \frac{\sqrt{n}\ln(c)}{\ln(n)}) = \infty * (-\infty) = -\infty$.

  Hay $\lim\limits_{n\to\infty}(\sqrt{n}\ln(n) - n\ln(c)) = -\infty$.

  Thay vào (*) $\Rightarrow \lim\limits_{n\to\infty} \frac{f(n)}{c^n} = e^{-\infty} = 0$. (2)

  Vậy từ (1),(2) ta có: $\begin{cases} \lim\limits_{n\to\infty} \frac{f(n)}{n^a} = \infty \\ \lim\limits_{n\to\infty} \frac{f(n)}{c^n} = 0 \quad \square \end{cases}$

Câu 1.2 Đánh giá độ lớn của hàm $f(n) = \frac{\lfloor n^4\log(n) - 2n^2 + 5\rfloor}{\lceil n\sqrt{n^3+1} + 10\rceil}$dưới dạng $\Theta(n^c)$.

Giải: Đặt $g(n) = n\sqrt{n}\log(n)$

Ta có: $\lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = \lim\limits_{n\to\infty} \frac{\frac{\lfloor n^4\log(n)-2n^2+5\rfloor}{\lceil n\sqrt{n^3+1}+10\rceil}}{n\sqrt{n}\log(n)} \simeq \lim\limits_{n\to\infty} \frac{n^4\log(n)-2n^2+5}{n\sqrt{n}\log(n)(n\sqrt{n^3+1}+10)}$(Với n đủ lớn)

$= \lim\limits_{n\to\infty}(\frac{n^2\sqrt{n}}{n\sqrt{n^3+1}+10} - \frac{2\sqrt{n}}{\log(n)(n\sqrt{n^3+1}+10)} + \frac{5}{n\sqrt{n}\log(n)(n\sqrt{n^3+1}+10)})$.

- $\lim\limits_{n\to\infty} \frac{n^2\sqrt{n}}{n\sqrt{n^3+1}+10} = \lim\limits_{n\to\infty} \frac{1}{\sqrt{1+\frac{1}{n^3}} + \frac{10}{n^2\sqrt{n}}} = \frac{1}{\sqrt{1+0}+0} = 1$.
- $\lim\limits_{n\to\infty} \frac{2\sqrt{n}}{\log(n)(n\sqrt{n^3+1}+10)} = \lim\limits_{n\to\infty} \frac{2}{\log(n)(\sqrt{n}\sqrt{n^3+1} + \frac{10}{\sqrt{n}})} = 0$.
- $\lim\limits_{n\to\infty} \frac{5}{n\sqrt{n}\log(n)(n\sqrt{n^3+1}+10)} = 0$.

  $\Rightarrow \lim\limits_{n\to\infty} \frac{f(n)}{g(n)} = 1 \in \mathbb{R}^+.(*)$

Vậy $f(n) \in \Theta(g(n))$ hay $f(n) \in \theta(n\sqrt{n}\log(n))$.

Câu 1.3 Xét các hàm số $f(x), g(x), h(x)$ xác định và nhận giá trị trên tập số dương,xét tính đúng/sai của mệnh đề sau:Nếu $f(x)g(x) \in \theta(h(x)^2)$ và $f(x) \in \Omega(h(x))$thì có $g(x) \in O(h(x))$.

Giải: $f(x)g(x) \in \theta(h(x)^2) \Rightarrow \lim\limits_{n\to\infty} \frac{f(x)g(x)}{h(x)^2} \in \mathbb{R}^+ \Rightarrow \lim\limits_{n\to\infty} [\frac{f(x)}{h(x)} \frac{g(x)}{h(x)}] \in \mathbb{R}^+(*)$.

Vì $f(x) \in \Omega(h(x)) \Rightarrow \lim\limits_{x\to\infty} \frac{f(x)}{g(x)} > 0$.

- Với $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = a \in \mathbb{R}^+$:

  $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)}$ xác định trên $\mathbb{R}^+$nên thay vào (*) ta được $\lim\limits_{x\to\infty} \frac{g(x)}{h(x)} \in \mathbb{R}^+ \Rightarrow g(x) \in O(h(x))$. (1)
- Với $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = \infty$:

  Để (*) đúng thì biểu thức ở trong giới hạn phụ thuộc phải ở dạng vô định và ở đây dạng vô định phải là ở dạng $\infty.0$ do đó ta có thể suy ra được $\lim\limits_{x\to\infty} \frac{g(x)}{h(x)} = 0$.

  $\Rightarrow g(x) \in O(h(x))$. Tức là $g(x) \in O(h(x))$. (2)

Từ (1),(2) $\Rightarrow g(x) \in O(h(x)) \Rightarrow$ Mệnh đề đúng.

Câu 2.1 So sánh phí tính toán của hai đoạn code bên dưới:

```
//algorithm 1
for(int i = 1; i*i <= n; i++){
    for(int j = 1; j*j*j <= n - i*i; j++){
        res += j;
    }
}
//algorithm 2
for(int i = 1; i*i*i <= n; i++){
    for(int j = 1; j*j <= n - i*i*i; j++){
        res += j;
    }
}
```

- Xét algorithm 1:

  Xét chi phí tính toán của thuật toán bằng số lần thực hiện phép gán $res+=j$.
  Vậy từ mã giả ta có được chi phí tính toán của thuật toán trên là :$\sum_{i=1}^{\lfloor\sqrt{n}\rfloor} \sum_{j=1}^{\lfloor\sqrt[3]{n-i^2}\rfloor} 1$.

  Vậy khi đó ta cần xem xét chi phí của $f_1(n) = \sqrt[3]{n-1} + \sqrt[3]{n-4} + \sqrt[3]{n-9} + \cdots + 0$ với $\sqrt{n}$ số hạng.
  Nhận xét:
  - $f_1(n) \le \sqrt[3]{n-1} + \sqrt[3]{n-1} + \cdots + \sqrt[3]{n-1}(\sqrt{n}$ số hạng$) = \sqrt{n}\sqrt[3]{n-1} \Rightarrow f_1(n) \in O(n^{\frac{5}{6}})$.
  - $f_1(n) \ge \sqrt[3]{n-1} + \sqrt[3]{n-4} + \cdots + \sqrt[3]{n - (\frac{\sqrt{n}}{2})^2}$ $(\frac{\sqrt{n}}{2}$ số hạng$)$

    $\ge \frac{\sqrt{n}}{2}\sqrt[3]{n - \frac{n}{4}} = \frac{\sqrt[3]{6}}{4}\sqrt{n}\sqrt[3]{n} \Rightarrow f_1(n) \in \Omega(n^{\frac{5}{6}})$.

  Vậy $f_1(n) \in \theta(n^{\frac{5}{6}})$. (1)
- Xét algorithm 2:

  Xét chi phí tính toán của thuật toán bằng số lần thực hiện phép gán $res+=j$.
  Vậy từ mã giả ta có được chi phí tính toán của thuật toán trên là :$\sum_{i=1}^{\lfloor\sqrt[3]{n}\rfloor} \sum_{j=1}^{\lfloor\sqrt{n-i^3}\rfloor} 1$.

  Vậy khi đó ta cần xem xét chi phí của $f_2(n) = \sqrt{n-1} + \sqrt{n-8} + \cdots + 0$ với $\sqrt[3]{n}$ số hạng.
  Nhận xét:
  - $f_2(n) \le \sqrt{n-1} + \sqrt{n-1} + \cdots + \sqrt{n-1}(\sqrt[3]{n}$ số hạng$) = \sqrt[3]{n}\sqrt{n-1} \Rightarrow f_1(n) \in O(n^{\frac{5}{6}})$.

- $f_2(n) \ge \sqrt{n-1} + \sqrt{n-8} + \cdots + \sqrt{n-\frac{n}{8}}$ ($\frac{\sqrt[3]{n}}{2}$ số hạng)

$$> \frac{\sqrt[3]{n}}{2}\sqrt{\frac{7n}{8}} = \frac{\sqrt{14}}{8}\sqrt{n}\sqrt[3]{n} \Rightarrow f_2(n) \in \Omega(n^{\frac{5}{6}}).$$

Vậy $f_2(n) \in \theta(n^{\frac{5}{6}})$. (2)

Từ (1),(2) ta suy ra độ phức tạp hai thuật toán là như nhau.

Câu 2.2 Với n nguyên dương, đặt $T_n = (\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}})(\log(1) + \log(2) + \cdots + \log(n))$. Ước lượng độ lớn của T_n theo ký hiệu của Θ.

- Xét $f(n) = \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n}}$
  - $f(n) \ge \frac{1}{\frac{\sqrt{n}}{2}} + \cdots + \frac{1}{\sqrt{n}}(\frac{n}{2}$ số hạng$) > \frac{n}{2}\frac{1}{\sqrt{\frac{n}{2}}} = \frac{\sqrt{2}}{2}\sqrt{n}$
  - $f(n) \le n\frac{1}{\sqrt{n}} = \sqrt{n}$
- Xét $g(n) = \log(1) + \log(2) + \cdots + \log(n)$
  - $g(n) \ge \log(\frac{n}{2}) + \cdots + \log(n)(\frac{n}{2}$ số hạng$) > \frac{n}{2}\log(n/2)$
  - $g(n) \le n\log(n)$

Vậy khi đó ta có: $\begin{cases} \frac{\sqrt{2}}{2}\sqrt{n} \le f(n) \le \sqrt{n} \\ \frac{n}{2}\log(\frac{n}{2}) \le g(n) \le n\log(n) \end{cases} \Rightarrow \begin{cases} f(n)g(n) \le n\sqrt{n}\log(n) \Rightarrow T_n \in O(n\sqrt{n}\log(n)) \\ f(n)g(n) \ge \frac{\sqrt{2}}{4}n\sqrt{n}\log(\frac{n}{2}) \Rightarrow T_n \in \Omega(n\sqrt{n}\log(n)) \end{cases}$

$$\Rightarrow T_n \in \theta(n\sqrt{n}\log(n))$$

Câu 3 Cho dãy số nguyên có n số $a_1, a_2, \ldots, a_n$ mà giá trị mỗi số thuộc tập hợp 0,1,2.Cần đếm xem có tất cả mấy cách chọn ra cặp chỉ số (L, R) với $1 \le L \le R \le n$ sao cho trong dãy con của a xét từ vị trí L đến vị trí R thì có một số nào đó xuất hiện từ 3 lần trở lên.Ví dụ: nếu $a = [0, 1, 2, 2, 2]$ thì đáp số là 3, ứng với 3 cách chọn cặp chỉ số $(L, R) = (3, 5), (2, 5), (1, 5)$.

Câu 3.1 Hãy mô tả cách vét cạn naive cho bài toán trên (dùng mã giả hoặc code C++/Python) và đánh giá độ phức tạp tương ứng của thuật toán đó.

```
int res=0;
for(int i=1;i<= n-1;i++){ C[i]
    for(int j=i+1;j<= n;j++){ B[j]
        int count[]={0,0,0};
        for(int k=i;k\leqj;k++){ //A[k]
            count[a[k]]++;
        }
        if(count[0]<= 3||count[1]<= 3||count[2]<= 3) res++;
    }
}

```

Ta sẽ xét độ phức tạp trong trường hợp xấu nhất đó chính là if luôn thỏa điều kiện. Ta xét từng vòng lặp và từ vòng lặp nhỏ nhất là A(k):

- Có (j-i+1) phép gán cho giá trị k.
- Tiếp theo là gán giá trị cho mảng count.

$\Rightarrow A(k) = \sum_{k=i}^{j} 2 = 2(j-i+1)$
Ta xét vòng lặp nhỏ tiếp theo là B[j]:

- có (n-i-1+1) lần gián giá trị cho j.
- Mỗi vòng lặp gồm một phép gán cho con trỏ,1 phép gán cho res.
- Mỗi vòng lặp phụ thuộc vào i,j ở A.

$\Rightarrow B(j) = \sum_{j=i+1}^{n}(3+Q(k)) = \sum_{j=i+1}^{n}(3+2j-2i+2)$

$= \sum_{j=i+1}^{n}(5+2j-2i) = 5(n-i-1+1) - 2i(n-i) + (n-i)(n+i+1)$

$= n^2+6n-i(2n+4)+i^2$

Ta xét vòng lặp tiếp theo là C[i]:

- có (n-1) phán gán cho giá trị if
- Phép gán phụ thuộc i ở B[j]

$\Rightarrow C[i] = \sum_{i=1}^{n-1}(B[j]+1)+1 = \sum_{i=1}^{n-1}(n^2+6n-i(2n+4)+i^2+1)+1$

$= n^2(n-1)+6n(n-1)-(2n+4)(n-1)\frac{n}{2}+\frac{(n-1)(n-2)}{2}+\frac{(n-1)n(2n-1)}{6}+1$

$\lim\limits_{n\to\infty}\frac{C[i]}{n^3} = \lim\limits_{n\to\infty}[1-\frac{1}{n^2}+6(\frac{1}{n}-\frac{1}{n^2})-\frac{1}{2}(2+\frac{4}{n})(1-\frac{1}{n})+\frac{(1-\frac{1}{n})(2-\frac{1}{n})}{6}+\frac{1}{n^3}] = \frac{1}{3} \in \mathbb{R}.$

$\Rightarrow C(n) \in \theta(n^3).$

Câu 3.2 Bằng cách sử dụng thêm các dãy phụ x, y, z trong đó x_i cho biết số giá trị bằng 0 trong dãy từ a_1 đến a_i; tương tự y_i và z_i lần lượt cho biết số giá trị bằng 1 và bằng 2,hãy đề xuất cải thiến cách làm trên thành $O(n^2)$.

```
int count0[n={0}];
int count1[n={0}];
int count2[n={0}];
for(int i=1;i<= n;i++){
    count0=count0[i-1]+(a[i]==0?1:0);
    count1=count1[i-1]+(a[i]==1?1:0);
    count2=count2[i-1]+(a[i]==2?1:0);
}
int res=0;
for(int i=1;i<= n-2;i++){
    for(int j=i+2;j\<= n;j++){
        if(count0[j]-count0[i-1]<= 3||count1[j]-count1[i-1]<= 3
        ||count2[j]-count2[i-1]<= 3) res++;
    }
}
return res;

```

Ta gọi B(n) là phép gán của toàn hàm:

- có 3 phép gián cố định khai báo mảng.
- Mỗi vòng lặp khởi tạo mảng cộng dồn n lần,mỗi lần 3 phép gán tổng là 3n phép gián.
- 1 phép gán cố định cho res.
- Xét trường hợp xấu nhất của hai vòng for lòng nhau ta gọi A(n) là phép gán của hai vòng for lồng nhau.

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

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

$= n^2-2n$

Vậy $B(n) = 3+3n+1+n^2-2n = n^2+n+4 \Rightarrow \lim\limits_{n\to\infty}\frac{P(n)}{n^2} = \lim\limits_{n\to\infty}(1+\frac{1}{n}+\frac{4}{n^2}) = 1 \in \mathbb{R}.$

$\Rightarrow P(n) \in \theta(n^2)$

Câu 3.3 Bằng cách sử dụng nguyên lí Dirichlet,hãy tìm cách cải tiền cách làm trên thành $O(n)$.

Áp dụng nguyên lý Dirichlet,chỉ cần lấy ít nhất 7 phần tử sẽ có ít nhất trong các tập 0,1,2 được lặp lại 3 lần.Vậy thuật toán sẽ là:

- Đếm số cách L,R có khoảng cách ≥ 7.
- Loại các trường hợp có khoảng cách 6,5,4,3.

Số cặp [L,R] thỏa mãn bài toán bằng tổng số đoạn khoảng cách ≥ 7 + số trường hợp đếm thêm được.

- số đoạn có khoảng cách bằng 7 là n-6.
- số đoạn có khoảng cách 6 là n-5.

Kiểm tra trong 6 phần tử bằng đoạn mã :int count[3]=0,0,0

```
for(int i=start;i\leqend;i++){
    count[a[i]]++;
}
if(count[0]<= 3||count[1]<= 3||count[2]<= 3) res++;

```

Ta xét trường hợp xấu nhất sẽ là 1+6.2+1=14.

Tương tự với các trường hợp có đoạn khoảng cách là 5,4,3. Số đoạn có khoảng cách ≥ 7 là:

$(n-6)+(n-7)+(n-8)+\cdots+n-n+1=\frac{(n-6)(n-5)}{2}$.

Việc trên đồng nghĩa với việc tính toán trên phụ thuộc vào khoảng cách của 6,5,4,3.

$\Rightarrow$ Số phép gán $f(n)=(n-5)*14+(n-4)*12+(n-3)*10+(n-2)*8$.

$\lim\limits_{n\to\infty}\frac{f(n)}{n}=\lim\limits_{n\to\infty}(14(1-\frac{5}{n})+12(1-\frac{4}{n})+10(1-\frac{3}{n})+8(1-\frac{2}{n}))=4\in\mathbb{R}$

$\Rightarrow f(n)\in O(n)$.