Mathematical Analysis of Algorithm--Knuth

孙祯鹏

Introduciton

> **Q：算法分析的核心目标？**
>
> A：研究如何量化分析各个不同算法的好坏
>
> **Q：算法分析在做些什么？**
>
> A：TypeA 对一个特定算法去分析它的时间、空间复杂度；
>
> TypeB 对解决一个问题的所有算法进行分析，确定一个最优解或者确定这些算法的复杂度范围

两种分析方法各有利弊，TypeB虽然看起来很美好，一次性解决所有问题，

但TypeB有两个问题：

1）过于复杂，很难建立有效的数学模型

2）如何定义最优解？

这里作者举了一个例子来说明TypeB的局限性：

> 计算 x^{31} 需要多少次运算
>
> 不允许除法——快速幂，$x^{31} = x \times x^2 \times x^4 \times x^8 \times x^{16}$，9次乘法
>
> 允许除法——$x^{32} \div x$，5次乘法，1次除法

作者拿了两个问题作为例子来展现算法分析的大致思路。

在这次课上我会给大家讲第一个问题——重排序

In Situ Permutation

1. 问题定义

给定一个序列$x_1, x_2, \ldots, x_n$和一个1~n的排列$p(i)$，要求将原序列变为为$x_{p(1)}, x_{p(2)}, \ldots, x_{p(n)}$

要求：

1）不能修改p的存储空间

2）空间复杂度为$O(1)$

---

2. 解决算法

首先可以想到，在根据p置换x序列中元素的时候，一定会出现环的情况

即：$p = [3, 2, 1]$，$x_1 \to x_3 \to x_1$

于是我们可以对于每一个环去循环移动其中的元素

```
for (int j = 1; j <= n; ++j) {   // 1 | 枚举每一个位置寻找环的第一个元素
    int k = p(j);               // n
    while (k > j)               // n + a | 对当前枚举的j去不断找他后面的元素，直到第一个比j小的
        k = p(k)                // a
    if (k == j) {               // n | 如果k<j，则当前环已经处理过了，如果k=j那么j是环首，开始处理
        int y = x[j], l = p[j];  // b | y暂时存放x[j]的值，最后放回环尾，k是当前位置，l是下一个位置
        while (l != j) {        // b + c
            x[k] = x[l];        // c
            k = l;              // c
            l = p[l];           // c
        }
        x[k] = y;               // n | 将x[j]的原值放入环尾，当前环处理结束
    }
}
```

3.复杂度分析

我们将算法每个语句的执行次数用n,a,b,c表示，n是很显然的，

b是排列p中环的个数；

b+c=n，因为序列中的每个元素只会在所有环中出现一次；

a是元素在当前环中与环尾的距离，若环长为l，则该环的a总计$\frac{l(l+1)}{2}$，最坏情况下，$l=n$，$a=\frac{n(n+1)}{2}$。

复杂度分析的重点在于计算a和b

1）第一类斯特林数

第一类斯特林数$S(n,k)$，表示将 n 个不同元素构成m个圆（环）排列的数目

无符号第一类斯特林数$S_u(n,k)$

1. $S_u(n+1,m)=S_u(n,m-1)+n\times S_u(n,m)$

> 考虑将n+1元素构成m个环排列，考虑第n+1个元素：
>
> （1）如果n个元素构成了m-1个环排列，那么第n+1个元素独自构成一个环排列。
>
> 方案数：$S_u(n,m-1)$
>
> （2）如果n个元素构成了m个环排列，将第n+1个元素插入到任意元素的左边。
>
> 方案数：$n\times S_u(n,m)$

2. $\Sigma_{k=1}^{n}S_u(n,k)x^k=\Pi_{k=0}^{n-1}(x+k)$

> 通过数学归纳法可证：
>
> 1. 当$n=1$时，$S_u(1,1)=1$，成立
> 2. 设当n时满足上式，则需证明$\Sigma_{k=1}^{n+1}S_u(n+1,k)x^k=\Pi_{k=0}^{n}(x+k)$成立
>
> $$\begin{aligned}左式&=\Sigma_{k=1}^{n+1}[nS_u(n,k)]+S_u(n,k-1)]x^k\\&=\Sigma_{k=1}^{n+1}nS_u(n,k)x^k+\Sigma_{k=1}^{n+1}S_u(n,k-1)x^k\end{aligned}$$
>
> $$又有S_u(n,n+1)=0,S_u(n,0)=0$$
>
> $$\begin{aligned}\therefore 左式&=\Sigma_{k=1}^{n}nS_u(n,k)x^k+\Sigma_{k=1}^{n}S_u(n,k)x^{k+1}\\&=(x+n)\Sigma_{k+1}^{n}Su(n,k)x^k\\&=(x+n)\Pi_{k=0}^{n-1}(x+k)=右式\end{aligned}$$

2）求$E(b)$和$D(b)$

$$\begin{aligned}E(b)&=\frac{\Sigma_{k=1}^{n}k\times S_u(n,k)}{n!}\\&=\frac{\Sigma_{k=1}^{n}k\times S_u(n,k)x^{k-1}}{n!}|_{x=1}\\&=\frac{d}{dx}\frac{\Sigma_{k=1}^{n}Su(n,k)x^k}{n!}|_{x=1}\\&=\frac{d}{dx}\frac{\Pi_{k=0}^{n-1}(x+k)}{n!}|_{x=1}\\&=\frac{1}{n!}[\Pi_{k=0}^{n-1}(x+k)\times\Sigma_{k=0}^{n-1}\frac{1}{x+k}]|_{x=1}\end{aligned}$$

$E(b)=\frac{1}{n!}[n!\times H_n]=H_n=\Sigma_{i=1}^{n}\frac{1}{i}$

令$H_n^{(2)}=\Sigma_{i=1}^{n}\frac{1}{i^2}$

$D(b)=E(b^2)-E^2(b)=\frac{\Sigma_{k=0}^{n}k^2S_u(n,k)}{n!}-H_n^2$

同样的，我们设x=1，有
$$\begin{aligned}\Sigma_{k=0}^{n}(k^2-k)S_u(n,k)x^{k-2}=\frac{d\Sigma_{k=0}^{n}S_n(n,k)x^k}{dx^2}&=\Pi_{k=0}^{n-1}(x+i)\times\Sigma_{i=0}^{n-1}\Sigma_{j=0,j\neq i}^{n-1}\frac{1}{(x+i)(x+j)}\\&=n!(\Sigma_{i=0}^{n-1}\Sigma_{j=0}^{n-1}\frac{1}{(x+i)}\frac{1}{(x+j)}-\Sigma_{i=0}^{n-1}\frac{1}{(x+i)^2})=n!(H_n^2-H_n^{(2)})\end{aligned}$$

$\therefore D(b)=\Sigma_{k=0}^{n}(k^2-k)S_u(n,k)x^{k-2}+\Sigma_{k=0}^{n}kS_u(n,k)x^{k-2}-H_n^2=H_n-H_n^{(2)}$

3）求$E(a)$和$D(a)$

为了方便求解，我们对排列p映射到q，对于其每个环，将最小值转至环首，环与环之间按最小值从大到小排序。

如(8 2 7 1 6 9 3 4 5) -> (8 4 1) (6 9 5) (2) (73) ->(5 6 9) (3 7) (2) (1 8 4)->(5 6 9 3 7 2 1 8 4)

这个映射一定为双射，那么对排列p的考虑就转移到了排列q上。

a是指环中一个元素与环尾的距离，其路径上不可能存在比当前元素更小的数，

那么在q上，则是向后找到最后一个大于他的数，假设位置分别为i，j，它对a的贡献为$j-i$。

我们记$y_{ij}=1(\forall i<k\leq j,q_i<q_k)\quad 0(else)$，其还表示$q_i$是否为$q_i$到$q_j$的最小值

就有$a=\Sigma_{i=1}^{n}\Sigma_{j=i+1}^{n}y_{ij}=\Sigma_{1\le i<j\le n}y_{ij}$

因为期望的线性性，$E(a)=\Sigma_{1\le i<j\le n}E(y_{ij})$

$y_{i,j}$只和$q[i,j]$内相对顺序有关，假设$q[i+1,j]$的顺序已经确定，则q_i还有$j-i+1$个位置可以放，

在最小值的概率是$\frac{1}{j-i+1}$

则$E(a)=\Sigma_{1\le i<j\le n}\frac{1}{j-i+1}=\Sigma_{2\le r\le n}\frac{n+1-r}{r}=(n+1)\Sigma_{2\le r\le n}\frac{1}{r}-(n-1)=(n+1)H_n-2n$

然后是最难的方差部分

$D(a)=E(a^2)+E^2(a)$

$E(a^2)=(\Sigma_{1\le i<j\le n}y_{ij})^2$

$$\begin{aligned}\Big(\sum_{1\le i<j\le n}y_{ij}\Big)^2&=\sum_{1\le i<j\le n}y_{ij}^2+\sum_{\substack{1\le i<j\le n\\1\le k<l\le n\\(i,j)\ne(k,l)}}y_{ij}y_{kl}\\&=\sum_{1\le i<j\le n}\overline{y}_{ij}\\&\qquad+2\sum_{1\le i<j<k<l\le n}(y_{ij}y_{kl}+y_{ik}y_{jl}+y_{il}y_{jk})\\&\qquad+2\sum_{1\le i<j<k\le n}(y_{ij}y_{jk}+y_{ik}y_{jk}+y_{ij}y_{ik})\\&=\overline{a}+2(A+B+C+D+E+F)\end{aligned}$$

$$\begin{aligned}&B=\binom{n}{2}-2Z,\qquad C=Y-Z-2\binom{n}{2}+3X,\\&D=E=Z-X,\qquad F=\binom{n}{2}-2X,\end{aligned}$$

$$\begin{aligned}X&=\sum_{1\le i<j\le n}\frac{1}{j-i+1},\\Y&=\sum_{1\le i<j\le n}H_{j-i},\\Z&=\sum_{1\le i<j\le n}\frac{1}{j-i+1}H_{j-i}.\end{aligned}$$

$$\begin{aligned}X&=(n+1)H_n-2n,\\Y&=\frac{1}{2}(n^2+n)H_n-\frac{3}{4}n^2-\frac{1}{4}n,\\Z&=\frac{1}{2}(n+1)(H_n^2-H_n^{(2)})-nH_n+n.\end{aligned}$$

$$\begin{aligned}A&=\sum_{1\le i<j<k<l\le n}\frac{1}{(j-i+1)(l-k+1)}\\&=\sum_{\substack{r\ge2\\s\ge2\\r+s\le n}}\frac{1}{rs}\binom{n-r-s+2}{2}\\&=\sum_{\substack{2\le r\le l-2\\4\le t\le n}}\frac{1}{t}\left(\frac{1}{r}+\frac{1}{t-r}\right)\binom{n-t+2}{2}\\&=2\sum_{\substack{2\le r\le t-2\\4\le t\le n}}\frac{1}{rt}\binom{n-t+2}{2}\\&=\sum_{\substack{2\le r\le t-2\\4\le t\le n}}\frac{1}{rt}((n+2)(n+1)-t(2n+3)+t^2)\\&=(n+2)(n+1)U-(2n+3)V+W\end{aligned}$$

if we let $r=j-i+1,\ s=l-k+1,\ t=r+s$. Then

$$\begin{aligned}U&=\frac{1}{2}(H_n-1)^2-\frac{1}{2}H_n^{(2)}+\frac{1}{n},\\V&=(n-1)H_{n-2}-2n+4,\\W&=\frac{1}{2}((n^2+n-2)(H_{n-2}-1)-\frac{1}{2}(n-1)(n-2)+1-3(n-3)).\end{aligned}$$

1. X

> $X=\Sigma_{1\le i<j\le n}\frac{1}{j-i+1}=(n+1)H_n-2n$

2.Y

> $Y=\Sigma_{1\le i<j\le n}H_{j-i}=\Sigma_{r=1}^{n-1}(n-r)H_r=\Sigma_{r=1}^{n}(n-r)H_r=\Sigma_{i=1}^{n}\frac{1}{i}\frac{(n-r)(n-r+1)}{2}=\frac{1}{2}\Sigma_{i=1}^{n}\frac{1}{i}[(n+1)n-(2n+1)i+i^2]$
>
> $Y=\frac{1}{2}[(n+1)nH_n-n(2n+1)+\frac{n(n+1)}{2}]=\frac{1}{2}(n^2+n)H_n-\frac{3}{4}n^2-\frac{1}{4}n$

3.Z

> $Z=\Sigma_{1\le i<j\le n}\frac{H_{j-i}}{j-i+1}=\Sigma_{r=2}^{n}\frac{n+1-r}{r}H_{r-1}=(n+1)\Sigma_{r=2}^{n}\frac{H_{r-1}}{r}-\Sigma_{r=2}^{n}H_{r-1}$
>
> $Q=\Sigma_{r=2}^{n}\frac{H_{r-1}}{r}=\Sigma_{r=1}^{n-1}\frac{H_n-H_r}{r}=H_n\Sigma_{r=1}^{n-1}\frac{1}{r}-\Sigma_{r=1}^{n-1}\frac{1}{r}(H_{r-1}+\frac{1}{r})=H_n^2-\frac{1}{n}H_n-(\Sigma_{r=2}^{n}(\frac{H_{r-1}}{r}+\frac{1}{r^2})+1-(\frac{H_{n-1}}{n}+\frac{1}{n^2}))$
>
> $Q=H_n^2-H_n^{(2)}-Q$
>
> $Q=\frac{1}{2}(H_n^2-H_n^{(2)})$
>
> $P=\Sigma_{r=2}^{n}H_{r-1}=\Sigma_{i=1}^{n-1}\frac{n-i}{i}=nH_n-n$

> $Z = \frac{1}{2}(n+1)(H_n^2 - H_n^{(2)}) - nH_n + n$

4.B

> $$\begin{aligned} B = \Sigma_{1\le i<j<k<l\le n} y_{ik}y_{jl} &= \Sigma_{1\le i<j<k<l\le n} \frac{1}{l-i+1}\frac{1}{l-j+1} \\ &= \Sigma_{1\le i<j<l\le n} \frac{1}{l-i+1}\frac{l-j-1}{l-j+1} \\ &= \Sigma_{1\le i<j<l\le n} \frac{1}{l-i+1}(1-\frac{2}{l-j+1}) \\ &= \Sigma_{1\le i<j<l\le n} \frac{1}{l-i+1} - 2\Sigma_{1\le i<j<l\le n} \frac{1}{l-i+1}\frac{1}{l-j+1} \\ &= \Sigma_{1\le i<l\le n}(1-\frac{2}{l-i+1}) - 2\Sigma_{1\le i<l\le n}\frac{H_{l-i}-1}{l-i+1} \\ &= C(n,2) - 2X - 2(Z-X) = C(n,2) - 2Z \end{aligned}$$

5. C

> $$\begin{aligned} C &= \Sigma_{1\le i<j<k<l\le n} E(y_{il}y_{jk}) = \Sigma_{1\le i<j<k<l\le n}\frac{1}{l-i+1}\frac{1}{k-j+1} \\ &= \Sigma_{1\le i<l-2\le n-2}\Sigma_{i<j<k<l}\frac{1}{l-i+1}\frac{1}{k-j+1} \\ &= \Sigma_{1\le i<l-2\le n-2}\frac{1}{l-i+1}\Sigma_{r=2}^{l-i-1}\frac{l-i-r}{r} \\ &= \Sigma_{1\le i<l-2\le n-2}\frac{1}{l-i+1}\Sigma_{r=2}^{l-i-1}(\frac{l-i}{r}-1) \\ &= \Sigma_{1\le i<l-2\le n-2}\frac{1}{l-i+1}[(l-i)(H_{l-i-1}-1)-(l-i-2)] \\ &= \Sigma_{1\le i<l-2\le n-2}[(H_{l-i-1}-1)-\frac{H_{l-i-1}-1}{l-i+1}-1+\frac{3}{l-i+1}] \\ &= \Sigma_{1\le i<l-2\le n-2}[H_{l-i-1}-\frac{H_{l-i-1}}{l-i+1}-2+\frac{4}{l-i+1}] \\ &= \Sigma_{1\le i<l-2\le n-2}[H_{l-i}-\frac{1}{l-i}-\frac{H_{l-i}}{l-i+1}+\frac{1}{(l-i)(l-i+1)}-2+\frac{4}{l-i+1}] \\ &= \Sigma_{1\le i<l-2\le n-2}(H_{l-i}-\frac{1}{l-i}) - \Sigma_{1\le i<l-2\le n-2}(\frac{H_{l-i}}{l-i+1}+\frac{1}{l-i+1}-\frac{1}{l-i}) - 2[C(n,2)-(n-1)-(n-2)] + 4\Sigma_{1\le i<l-2\le n-2}\frac{1}{l-i+1}] \\ C &= \Sigma_{1\le i<l\le n}(H_{l-i}-\frac{1}{l-i}) - (n-2) - (\Sigma_{1\le i<l\le n}(\frac{H_{l-i}}{l-i+1}+\frac{1}{l-i+1}-\frac{1}{l-i}) - \frac{n-2}{3}) - 2[C(n,2)-(2n-3)] + 4(\Sigma_{1\le i<l\le n}\frac{1}{l-i+1} - \frac{n-1}{2} \\ &= Y - Z + 3X - 2C(n,2) \end{aligned}$$

6. D

> $D = \Sigma_{1\le i<j<k\le n} y_{ij}y_{jk} = \Sigma_{1\le i<j<k\le n}\frac{1}{k-i+1}\frac{1}{k-j+1} = \Sigma_{1\le i<k\le n}\frac{H_{k-i}-1}{k-i+1} = Z - X$

7.E

> 同D

8.F

> $F = \Sigma_{1\le i<j<k\le n} y_{ij}y_{ik} = \Sigma_{1\le i<j<k\le n}\frac{1}{k-i+1} = \Sigma_{1\le i<k\le n}\frac{k-i-1}{k-i+1} = \Sigma_{1\le i<k\le n}(1-\frac{2}{k-i+1}) = C(n,2) - 2X$

9.A

> 同上图

10. U

> $$\begin{aligned} U &= \Sigma_{t=4}^{n}\Sigma_{r=2}^{t-2}\frac{1}{rt} = \Sigma_{t=4}^{n}\frac{1}{t}(H_{t-1}-1-\frac{1}{t-1}) = \Sigma_{t=4}\frac{H_{t-1}}{t} - \Sigma_{t=4}^{n}\frac{1}{t-1} \\ &= Q - \frac{H_1}{2} - \frac{H_2}{3} - (\Sigma_{t=1}^{n}\frac{1}{t} - 1 - \frac{1}{2} - \frac{1}{n}) \\ &= \frac{1}{2}(H_n^2 - H_n^{(2)}) - \frac{1}{2} - \frac{1}{3} - \frac{1}{6} - H_n + 1 + \frac{1}{2} + \frac{1}{n} = \frac{1}{2}(H_n-1)^2 - \frac{1}{2}H_n^{(2)} + \frac{1}{n} \end{aligned}$$
>
> Q同Z中计算

11.V

> $$\begin{aligned} V &= \Sigma_{t=4}^{n}\Sigma_{r=2}^{t-2}\frac{1}{r} = \Sigma_{t=4}^{n}(H_{t-1}-1-\frac{1}{t-1}) = \Sigma_{t=2}H_{t-1} - H_1 - H_2 - (n-3) - \Sigma_{t=4}^{n}\frac{1}{t-1} \\ &= P - 2 - \frac{1}{2} - (n-3) - H_n + 1 + \frac{1}{2} + \frac{1}{n} = (n-1)H_{n-2} - 2n + 4 \end{aligned}$$

12.W

> $$\begin{aligned} W &= \Sigma_{t=4}^{n} t\Sigma_{r=2}^{t-2}\frac{1}{r} \\ &= \Sigma_{i=2}^{n-2}\frac{1}{i}\Sigma_{t=i+2}^{n} t = \Sigma_{i=2}^{n-2}\frac{(n-i-1)(n+i+2)}{2i} \\ &= \frac{1}{2}[(n-1)(n+2)(\Sigma_{i=1}^{n-2}\frac{1}{i}-1) - 3(n-3) - \frac{(n-1)(n-2)}{2} + 1] \\ &= \frac{1}{2}[(n^2+n-2)(H_{n-2}-1) - 3(n-3) - \frac{1}{2}(n-1)(n-2) + 1] \end{aligned}$$

最后可以得出方差$D(a) = 2n^2 - (n+1)^2 H_n^{(2)} - (n+1)H_n + 4n$

$$a = \left(\min 0,\ \ \text{ave}\ n\ln n + O(n),\ \ \max \tfrac{1}{2}(n^2 - n),\right.$$
$$\left.\text{dev}\ \sqrt{2 - \pi^2/6}\ n + O(\log n)\right);$$
$$b = \left(\min 1,\ \ \text{ave}\ \ln n + O(1),\ \ \max n,\ \ \text{dev}\ \sqrt{\ln n} + O(1)\right).$$

这个算法的平均时间复杂度为 $O(nlogn)$，在极少数情况下可能达到$O(n^2)$。

4. 算法的改进

有一个神奇的优化可以使得最坏时间复杂度降至$O(nlogn)$

在每次判定当前是否为环首时，同时向着原有置换的反方向查找比他小的值，

那么对于最坏情况即长度为n的环，我们从最小元素开始枚举，设$f(n)$为长度为n的区间最小值在1号位，区间内所有元素判定完是否是环首的最坏情况下的总时间。

假设当前最小值在位置k，则有：

$f(n) = max_{1 \leq k < n}(min(k, n-k) + f(k) + f(n-k))$

$f(1) = 0$

复杂度计算略~~（不会）~~