一种时间复杂度计算方法

嘟嘟

在理论计算科学中，评估算法的时间复杂度繁琐且费时，目前还没有统一的方法。本文提出了一种评估方法，以简化时间复杂度的计算。

快速排序

设输入长度为$N = 2^n$。

（1）理想情况下，两路快排需要迭代的深度$\mathrm{p(N) = \ lgN}$，每一次迭代需要处理的输入长度为：

$$d(i) = N - (2^{i-1} - 1)\ , i \in [\lg(N)]$$

计算总量为：

$$TC = \sum_{i=1}^{lgN}(N + 1 - 2^{i-1}) = NlgN + lgN - \frac{1}{2}\sum_{i=1}^{lgN}2^i = NlgN + lgN - N + 1$$

$$\lim_{N\to\infty}\frac{\sum_{i=1}^{lgN}\left(N + 1 - 2^{i-1}\right)}{NlgN} = \lim_{N\to\infty}\frac{NlgN + lgN - N + 1}{NlgN} = 1$$

时间复杂度$O(N) = O(NlgN)$。

(2)最差的情况下，两路快排需要迭代的深度$\mathrm{p(N) = N}$，每一次迭代需要处理的输入长度为：

$$d(i) = N - i + 1\ , i \in N$$

计算总量为：

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

$$\lim_{N\to\infty}\frac{\sum_{i=1}^{N}N + 1 - i}{N^2} = \lim_{N\to\infty}\frac{\frac{1}{2}(N + 1)N}{N^2} = \frac{1}{2}$$

时间复杂度为$O(N) = O(N^2)$ 。

快速排序算法在理想情况下，需要处理的输入长度序列为：$\mathrm{N, N-1, N-3, N-7, \ldots, \frac{N}{2} - 1}$，等于依次需要处理的操作单元数量。同一输入长度，不同的操作可能有不同的操作单元数量。本文把需要处理的操作单元数量称为计算数，多个计算数组成的序列被称为计算序列。上面的计算只考虑每一步需要处理的计算数，它等于输入长度，代表需要的计算量，$\mathrm{p(N) * h(1) = O(N)}$。

时间复杂度定理

设初始输入长度为 N；迭代深度为$\mathrm{p(N)}$，单调递增；$\mathrm{i} \in [\mathrm{p(N)}] = \{1,2,\dots,\mathrm{p[N]}\}$；$\mathrm{d(i)} \in R_{>0}$是迭代深度为 i 时的输入长度；特定数据结构计算$\mathrm{d(i)}$需要的操作单元数量为$\mathrm{f(d(i))} \in R_{>0}$；$\mathrm{h(i)} \in R_{>0}$为某种操作的计算数即操作单元数量。

$H_i = \{h(j): i \in [p(N)], j \in [i]\} = \{h(1), h(2), \dots, h(i)\}$；$H = H_{p(N)}$；

$max(H_i) \in \{h(j): h(j) \in H_i, \forall h \in H_i: h(j) \geq h\}$；$max(H) = \max(H_{p(N)})$；

$max(|H_i|) = \max(|H_1|, |H_2|, \dots, |H_i|)$；$max(|H|) = \max(|H_{p(N)}|)$。

H_i的子集$H_i' = \left\{h(j): \frac{h(j)}{max(H_i)} \geq k, h(j) \in H_i\right\}, k \in (0,1]$，$H' = H_{p(N)}'$；

$max(H_i') \in \{h(j): h(j) \in H_i', \forall h \in H_i': h(j) \geq h\}$；$max(H') = \max(H_{p(N)}')$；

$max(|H_i'|) = \max(|H_1'|, |H_2'|, \dots, |H_i'|)$；$max(|H'|) = \max(|H_{p(N)}'|)$。

$$\frac{h(j)}{max(H_i)} \geq k \Longrightarrow j \geq h^{-}(k * \max(H_i)) \Longrightarrow |H_i'| = h^{-}(k * \max(H_i))$$

$$\max(|H_i'|) = \max\big(h^{-}(k * \max(H_1)), h^{-}(k * \max(H_2)), \dots, h^{-}(k * \max(H_i))\big)$$

$$\max(|H'|) = \max\Big(h^{-}(k * \max(H_1)), h^{-}(k * \max(H_2)), \dots, h^{-}\big(k * \max(H_{p(N)})\big)\Big)$$

如果$\exists c0, c1 \in R_{>0}, \forall i \in [p(N)]: c0 \leq \frac{f(d(i))}{h(i)} \leq c1$，则称$\mathrm{f(d(i))}$的时间复杂度等于$\mathrm{h(i)}$的时间复杂度，或 f 不改变 h 的时间复杂度；如果$\frac{\mathrm{f(d(i))}}{\mathrm{h(i)}} = 0$，则称$\mathrm{f(d(i))}$的时间复杂度小于$\mathrm{h(i)}$的时间复杂度；否则称$\mathrm{f(d(i))}$的时间复杂度大于$\mathrm{h(i)}$的时间复杂度。

公理一（有界公理）：$\mathrm{h(i)}$有界，即 $\forall i \in [\mathrm{p(N)}], \exists V \in R_{>0}: h(i) \leq V$。

定理一（等价定理）：如果 f 不改变 h 的时间复杂度，则算法的时间复杂度为：

$$O\left(\sum_{i=1}^{p(N)} h(i)\right) \qquad (1)$$

证明：$\exists c0, c1 \in R_{>0}, \forall i \in [p(N)]: c0 \leq \frac{f(d(i))}{h(i)} \leq c1$，有：

$$\begin{cases} O\left(\sum_{i=1}^{p(N)} \mathrm{f}\big(\mathrm{d(i)}\big)\right) \leq O\left(\sum_{i=1}^{p(N)} c1h(i)\right) = O\left(\sum_{i=1}^{p(N)} h(i)\right) \\ O\left(\sum_{i=1}^{p(N)} \mathrm{f}\big(\mathrm{d(i)}\big)\right) \geq O\left(\sum_{i=1}^{p(N)} c0h(i)\right) = O\left(\sum_{i=1}^{p(N)} h(i)\right) \end{cases}$$

$$\Rightarrow \mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right) \le \mathrm{O}\left(\sum_{i=1}^{p(N)} \mathrm{f}\big(\mathrm{d(i)}\big)\right) \le \mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right)$$

$$\Rightarrow \mathrm{O}\left(\sum_{i=1}^{p(N)} \mathrm{f}\big(\mathrm{d(i)}\big)\right) = \mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right)$$

定理二（线性定理）：如果$\exists c \in (0,1]: \frac{|H'|}{|H|} \ge c$，则：

$$\mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right) = O\big(p(N) * max(H)\big) \quad （2）$$

$$\forall h(i) \in H': O\big(p(N) * h(i)\big) = \mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right) \quad （3）$$

证明：因为$\exists k \in (0,1], \forall h(i) \in H': \frac{h(i)}{max(H)} \ge k$，$|H| = p(N)$，有：

$$\sum_{h(i) \in H'} h(i) \ge cp(N) * kmax(H)$$

又：

$$\sum_{i=1}^{p(N)} h(i) \ \ge \sum_{i \in H'} h(i)$$

$$\Rightarrow \sum_{i=1}^{p(N)} h(i) \ \ge cp(N) * kmax(H)$$

$$\Rightarrow \frac{\sum_{i=1}^{p(N)} h(i)}{\max(H) * p(N)} \ \ge ck$$

又：

$$\frac{\sum_{i=1}^{p(N)} h(i)}{max(H) * p(N)} \le \frac{\sum_{i=1}^{p(N)} \max(H)}{\max(H) * p(N)} = 1$$

$$\Rightarrow ck \le \frac{\sum_{i=1}^{p(N)} h(i)}{max(H) * p(N)} \le 1$$

$$\Rightarrow O\big(ck * max(H) * p(N)\big) \le O\left(\sum_{i=1}^{p(N)} h(i)\right) \le O(max(H) * p(N))$$

又：

$$O\big(ck * max(H) * p(N)\big) = O\big(max(H) * p(N)\big)$$

$$\Rightarrow O\big(max(H) * p(N)\big) \le O\left(\sum_{i=1}^{p(N)} h(i)\right) \le O(max(H) * p(N))$$

$$\Rightarrow O\left(\sum_{i=1}^{p(N)} h(i)\right) = O(max(H) * p(N))$$

因为$\forall h(i) \in H': \frac{h(i)}{max(H)} \geq k$，有：

$$\forall h(i) \in H': max(H) \geq h(i) \geq k * max(H)$$

$$\Longrightarrow \forall h(i) \in H': max(H) * p(N) \geq h(i) * p(N) \geq k * max(H) * p(N)$$

$$\Longrightarrow \forall h(i) \in H': O(max(H) * p(N)) \geq O(h(i) * p(N)) \geq O(k * max(H) * p(N))$$

$$\Longrightarrow \forall h(i) \in H': O(max(H) * p(N)) \geq O(h(i) * p(N)) \geq O(max(H) * p(N))$$

$$\Longrightarrow \forall h(i) \in H': O\big(h(i) * p(N)\big) = O(max(H) * p(N))$$

$$\Longrightarrow \forall h(i) \in H': O\big(h(i) * p(N)\big) = \mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right)$$

定理三（无穷定理）：如果$\nexists c \in (0,1]: \frac{|H'|}{|H|} \geq c$，则$\lim\limits_{N\to+\infty} \max(H) \to +\infty$。

证明：假设$\exists C1 \in R_{>0}, \lim\limits_{N\to+\infty} \max(H) \leq C1$时，$\exists c \in (0,1]: \frac{|H'|}{|H|} \geq c$。

$\mathrm{h(i)} \in N_{>0}$，有：

$$\frac{h(i)}{\max(H_i)} \geq \frac{1}{\max(H)} \geq \frac{1}{C1}$$

取$k = \frac{1}{C1}$，$|H'| = |H|$。c 可取(0,1]之间的任意实数，逆否命题成立。

定理四（单调递增定理）：存在足够大的 M 当$i \geq M$时 h(i)单调，如果$\nexists c \in (0,1]: \frac{|H'|}{|H|} \geq c$，则 h(i)单调递增。

证明：根据无穷定理，$\lim\limits_{N\to+\infty} \max(H) \to +\infty$。

假设$i \geq M$时 h(i)递减，则存在足够大的$\mathrm{V} \in \mathrm{R}_{>0}$，当$\forall i \in N_{>0}, h(i) \leq V$。
与$\lim\limits_{N\to+\infty} \max(H) \to +\infty$矛盾，假设不成立。h(i)单调递增。

定理五（缩放定理）$\mathrm{O}(|\mathrm{H}'| * \max(\mathrm{H})) \leq O(\sum_{i=1}^{p(N)} h(i)) \leq O(\max(|\mathrm{H}'|) * \max(\mathrm{H}))$

证明：$O(\sum_{i=1}^{p(N)} h(i)) \leq O(\max(|\mathrm{H}'|) * \max(\mathrm{H}))$

$\forall M \in N_{>0}$，有：$\sum_{i=1}^{p(N)} h(i) = \sum_{i=1}^{M} h(i) + \sum_{i=M+1}^{p(N)} h(i)$

$i \geq M$时升序排列h(i)。设$\mathrm{p(N)} = \mathrm{M}_n, \mathrm{M}_1 = M$，$\mathrm{H}(M_j) = \{\mathrm{h(i)}: \mathrm{i} \in [\mathrm{M}, M_j]\}$，有：

$$\begin{cases} \sum_{i=M}^{p(N)} h(i) = \sum_{i=M}^{M_1} h(i) + \sum_{i=M_1}^{M_2} h(i) +..+ \sum_{i=M_{n-1}}^{M_n} h(i) \\ \forall h(i) \in \mathrm{H}(M_j): \dfrac{h(i)}{\max\left(\mathrm{H}(M_{j+1})\right)} < k \\ \max\left(\mathrm{H}(M_j)\right) < k * \max\left(\mathrm{H}(M_{j+1})\right) \end{cases}$$

$$\Rightarrow \sum_{i=M}^{p(N)} h(i) < \max\left(\mathrm{H}(M_n)\right) \left(\sum_{i=M}^{M_1} k^{n-1} + \sum_{i=M_1}^{M_2} k^{n-2} +..+ \sum_{i=M_{n-1}}^{M_n} 1\right)$$

$$\Rightarrow \sum_{i=M}^{p(N)} h(i) < \max\left(\mathrm{H}(M_n)\right) * \max(|H'|) * (k^{n-1} + k^{n-2}+..+1)$$

$$\Rightarrow \sum_{i=M}^{p(N)} h(i) < \max(\mathrm{H}) * \max(|H'|) * \frac{1-k^n}{1-k}$$

$$\sum_{i=1}^{p(N)} h(i) < \sum_{i=1}^{M} h(i) + \max(\mathrm{H}) * \max(|H'|) * \frac{1-k^n}{1-k}$$

$$O\left(\sum_{i=1}^{p(N)} h(i)\right) \le O(\max(\mathrm{H}) * \max(|H'|))$$

证明：$O(\sum_{i=1}^{p(N)} h(i)) \ge O(|\mathrm{H}'| * \max(\mathrm{H}))$

$$\frac{h(i)}{max(H)} \ge k$$

$$\sum_{i=M}^{p(N)} h(i) > \sum_{h(i)\in H'} k * \max(\mathrm{H})$$

$$\sum_{i=M}^{p(N)} h(i) > |\mathrm{H}'| * \mathrm{k} * \max(\mathrm{H})$$

$$O(\sum_{i=M}^{p(N)} h(i)) \ge O(|\mathrm{H}'| * \max(\mathrm{H}))$$

$$\Rightarrow \mathrm{O}(|\mathrm{H}'| * \max(\mathrm{H})) \le O(\sum_{i=1}^{p(N)} h(i)) \le O(\max(|\mathrm{H}'|) * \max(\mathrm{H}))$$

定理六（常数定理）：存在足够大的 M，如果在区间$[M,i]$上$\forall k \in (0,1], \exists v \in N_{>0}: |H'| \le v$，则：

$$\mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right) = O(\max(H)) \qquad (4)$$

证明：因为$1 \le |H'| \le v$及其缩放定理，有：

$$\mathrm{O}(\max(\mathrm{H})) \le O(\sum_{i=1}^{p(N)} h(i)) \le O(\mathrm{v} * \max(\mathrm{H}))$$

$$\Rightarrow \mathrm{O}\left(\sum_{i=1}^{p(N)} h(i)\right) = O(\max(H)) \quad (5)$$

定理七（非线性定理）设$\nexists c \in (0,1]: \frac{|H'|}{|H|} \ge c$，如果存在足够大的$M \in N_{>0}$，$i \ge M$ 时 h(i)单调，有：

$$O\left(\sum_{i=1}^{p(N)} h(i)\right) = O\left(h^{-}\left(k * h\big(p(N)\big)\right) * \mathrm{h}\big(\mathrm{p(N)}\big)\right)$$

证明：$i \ge M \in N_{>0}$时 h(i)单调，且$\nexists c \in (0,1]: \frac{|H'|}{|H|} \ge c$，所以h(i)单调递增。有：

$$\begin{cases} \max(\mathrm{H}) = \mathrm{h}\big(\mathrm{p(N)}\big) \\ \max(|\mathrm{H}'|) = |\mathrm{H}'| = h^{-}\left(k * h\big(p(N)\big)\right) \end{cases}$$

根据缩放定理有：

$$\mathrm{O}(|\mathrm{H}'| * \max(\mathrm{H})) \le O(\sum_{i=1}^{p(N)} h(i)) \le O(\max(|\mathrm{H}'|) * \max(\mathrm{H}))$$

$$\Rightarrow \mathrm{O}(|\mathrm{H}'| * \max(\mathrm{H})) \le O(\sum_{i=1}^{p(N)} h(i)) \le O(|\mathrm{H}'| * \max(\mathrm{H}))$$

$$\Rightarrow O\left(\sum_{i=1}^{p(N)} h(i)\right) = O(h^{-}\left(k * h\big(p(N)\big)\right) * \mathrm{h}\big(\mathrm{p(N)}\big)$$

定理八（组合定理）：设 t 为 d(i)的计算种类数量，一个计算对应一个 h(i)，一个计算对应多个 f。如果 t 有限则算法的时间复杂度只与最大的 h(i)有关。设$\mathrm{maxh(i)} = \max(\mathrm{h1(i)}, \mathrm{h2(i)}, \ldots, \mathrm{ht(i)})$。

证明：

$$\sum_{i=1}^{p(N)} \mathrm{maxh}(i) < \sum_{i=1}^{p(N)} h1(i) + \sum_{i=1}^{p(N)} h2(i) + \cdots + \sum_{i=1}^{p(N)} ht(i) < t \sum_{i=1}^{p(N)} \mathrm{maxh}(i)$$

$$\begin{aligned} &\Rightarrow O\left(\sum_{i=1}^{p(N)} \mathrm{maxh}(i)\right) \le O\left(\sum_{i=1}^{p(N)} h1(i) + \sum_{i=1}^{p(N)} h2(i) + \cdots + \sum_{i=1}^{p(N)} ht(i)\right) \\ &\qquad \le O\left(t \sum_{i=1}^{p(N)} \mathrm{maxh}(i)\right) \\ &\Rightarrow O\left(\sum_{i=1}^{p(N)} h1(i) + \sum_{i=1}^{p(N)} h2(i) + \cdots + \sum_{i=1}^{p(N)} ht(i)\right) = O\left(\sum_{i=1}^{p(N)} \mathrm{maxh}(i)\right) \end{aligned}$$

Karatsuba 乘法

Algorithm 0.4 — Karatsuba multiplication.
Input: non-negative integers x, y each of at most n digits
Output: $x \cdot y$

Procedure Karatsuba(x, y)
    if $n \leqslant 4$ then return $x \cdot y$;
    Let $m = \lfloor n/2 \rfloor$
    Write $x = 10mx + x$ and $y = 10my + y$
    $A \leftarrow$ Karatsuba(x, y)
    $B \leftarrow$ Karatsuba$(x + x,\ y + y)$
    $C \leftarrow$ Karatsuba(x, y)
    return $(10n - 10m) \cdot A + 10m \cdot B + (1 - 10m) \cdot C$
endprocedure

设 lgN 是正整数，有：

$\mathrm{p(N) = lgN - 1}$；$\mathrm{d(i)} = \frac{N}{2^{i-1}}$；$\mathrm{h(i)} = 3^{i-1}$，$\mathrm{max(H)} = 3^{lgN-2}$。

设存在子集$|H'| = ratio|H|$，$\lim\limits_{n\to\infty} \frac{3^{\mathrm{ratio*(lgN-2)}}}{3^{lgN-2}} = \lim\limits_{n\to\infty} 3^{\mathrm{(ratio-1)(lgN-2)}} = 0 \geq c$。

$\mathrm{h(i)}$单调，O(Karatsuba)=O(max(h(*)))=O(3^{lgN-2})=O(N^{lg3})。