算法及其
时间复杂度

问题及实例

- **问题**

  需要回答的一般性提问，通常含若干参数

- **问题描述**

  定义问题参数(集合,变量,函数,序列等)
  说明每个参数的取值范围及参数间的关系
  定义问题的解
  说明解满足的条件(优化目标或约束条件)

- **问题实例**

  参数的一组赋值可得到问题的一个实例

算法

- **算法**

  有限条指令的序列

  这个指令序列确定了解决某个问题的一系列运算或操作

- **算法 *A* 解问题 *P***

  把问题 *P* 的任何实例作为算法 *A* 的输入

  每步计算是确定性的

  A 能够在有限步停机

  输出该实例的正确的解

基本运算与输入规模

- 算法时间复杂度: 针对指定基本运算,计数算法所做运算次数
- 基本运算: 比较, 加法, 乘法, 置指针, 交换…
- 输入规模：输入串编码长度
  通常用下述参数度量：数组元素多少，调度问题的任务个数，图的顶点数与边数等.
- 算法基本运算次数可表为输入规模的函数
- 给定问题和基本运算就决定了一个算法类

输入规模

- 排序：数组中元素个数 n
- 检索：被检索数组的元素个数 n
- 整数乘法：两个整数的位数 m, n
- 矩阵相乘：矩阵的行列数 i, j, k
- 图的遍历：图的顶点数 n, 边数 m

…

基本运算

- 排序: 元素之间的比较
- 检索: 被检索元素 x与数组元素的比较
- 整数乘法: 每位数字相乘(位乘) 1 次

  m位和n位整数相乘要做mn次位乘
- 矩阵相乘: 每对元素乘 1 次

  $i \times j$矩阵与 $j \times k$ 矩阵相乘要做 $i\,jk$ 次乘法
- 图的遍历: 置指针

…

算法的两种时间复杂度

对于相同输入规模的不同实例，算法的基本运算次数也不一样，可定义两种时间复杂度

最坏情况下的时间复杂度 $W(n)$

算法求解输入规模为 n 的实例所需要的最长时间

平均情况下的时间复杂度 $A(n)$

在给定同样规模为 n 的输入实例的概率分布下，算法求解这些实例所需要的平均时间

$A(n)$ 计算公式

平均情况下的时间复杂度 $A(n)$

设 S 是规模为 n 的实例集

实例 $I \in S$ 的概率是 P_I

算法对实例 I 执行的基本运算次数是 t_I

$$\boxed{A(n) = \sum_{I \in S} P_I t_I}$$

在某些情况下可以假定每个输入实例概率相等

例子：检索

检索问题

输入：非降顺序排列的数组 L,元素数 n，

数 x

输出：j

若 x 在 L中, j是x首次出现的下标；

否则 $j = 0$

基本运算：x 与 L 中元素的比较

顺序检索算法

$j=1$, 将x与$L[j]$比较. 如果 $x=L[j]$，则算法停止，输出 j；如果不等，则把 j 加1，继续x与$L[j]$的比较，如果 $j>n$，则停机并输出0.

实例

1	2	3	4	5

$x = 4$，需要比较 4 次

$x = 2.5$，需要比较 5 次

最坏情况的时间估计

不同的输入有 $2n+1$ 个，分别对应：

$x = L[1],\ x = L[2],\ \ldots,\ x = L[n]$

$x<L[1],\ L[1]<x<L[2],\ L[2]<x<L[3],\ \ldots,\ L[n]<x$

最坏情况下时间：$W(n) = n$

最坏的输入：x 不在 L 中或 $x = L[n]$

要做 n 次比较

平均情况的时间估计

输入实例的概率分布：

假设x在L中概率是 p,且每个位置概率相等

$$A(n)=\underline{\sum_{i=1}^{n} i\frac{p}{n}}+(1-p)n$$

等差数列求和

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

当 p=1/2时，

$$A(n)=\frac{n+1}{4}+\frac{n}{2}\approx\underline{\frac{3n}{4}}$$

改进顺序检索算法

$j=1$, 将 x与$L[j]$比较. 如果 $x=L[j]$，则算法停止，输出 j；如果 $x>L[j]$，则把 j 加1，继续 x与 $L[j]$的比较；如果 $x<L[j]$，则停机并输出0. 如果 $j>n$，则停机并输出 0.

实例

1	2	3	4	5

$x=4$，需要比较 4 次

$x=2.5$，需要比较 3 次

时间估计

最坏情况下：$W(n) = n$

平均情况下

输入实例的概率分布：假设 x 在 L 中每个位置与空隙的概率都相等

改进检索算法平均时间复杂度是多少?

小结：

- 算法最坏和平均情况下的时间复杂度定义
- 如何计算上述时间复杂度