ANALYSIS OF ALGORITHMS

Sanjay Ranka

January 22, 2003

Mathematical Preliminaries

Logarithms

Probability

Permutations

Summation formulas

Solutions of difference equations

Logarithms

Let a and y be positive numbers.

If $a^x = y$, then x is called the logarithm of y base a.

$x = \log_a \ y$.

$log_2 \ y$ is written as $\lg y$

$log_e \ y$ is written as $\ln y$

$log_{10} \ y$ is written as $\log y$

Basic Properties of Logarithms

Log is a one-to-one, monotone increasing function.

$\log_a 1 = 0$ for any a

$\log_a(y_1 y_2) = \log_a y_1 + \log_a y_2$

$\log_a(x^b) = b \log_a x$

$\log_x y = \frac{\log_z y}{\log_z x}$

Basic Probability

Consider an experiment with a collection of possible outcomes $\{s_1, s_2, \ldots, s_n\}$

A collection of nonnegative numbers $p(s_i)$'s $i = 1, 2, \ldots, n$ such that

$$0 \leq p(s_i) \leq 1 \text{ and } \sum_{i=1}^{n} p(s_i) = 1$$

is called probabilities of outcomes s_i

$$\text{mean } E(X) = \sum_{i=1}^{n} x_i \, p(x_i)$$

$$\text{variance } var(X) = \sum_{i=1}^{n} x_i^2 \, p(x_i) - E^2(X)$$

Permutations

A rearrangement of objects.

A B C D

D C A B

Given n objects, there are $n!$ possible permutations.

Summation Formula (Arithmetic Series)

Sum of $a, a+b, \ldots, a+(n-1)b$ is given by

$$na + \frac{n(n-1)}{2}\, b.$$

Summation Formula (Geometric Series)

The sum of $a, ar, \ldots, ar^{n-1}$ is given by

$$\begin{array}{rcl} s & = & a + ar + \cdots + ar^{n-1} \\ rs & = & \qquad ar + ar^2 + \cdots + ar^{n-1} + ar^n \\ \hline (r-1)s & = & ar^n - a. \end{array}$$

Therefore,

$$s = a\,\frac{r^n - 1}{r - 1}$$

If the series is **infinite**, then the sum becomes

$$\begin{cases} \frac{a}{1-r} \ \text{ for } \ -1 < r < 1 \\ na \ \text{ for } \ r = 1 \\ \pm\infty \text{ otherwise.} \end{cases}$$

Binomial Theorem:

$(1+x)^n = \Sigma_{k=0}^{n} \binom{n}{k} x^k$

Substituting $x = 1$ gives $2^n = \Sigma \binom{n}{k}$

Substituting $x = -1$, gives $0 = \Sigma \binom{n}{k}(-1)^k$.

Example: Show $\Sigma \; k\ x^k = \frac{1}{1-x}\left\{\frac{x^{n+1}-1}{x-1} - 1 - n\ x^{n+1}\right\}$.

$\Sigma_{k=0}^{n} \; x^k = \frac{x^{n+1}-1}{x-1}$ if $x \neq 1$.

Therefore,

$$\frac{d}{dx} \sum_{k=0}^{n} x^k = \sum_{k=0}^{n} k\ x^{k-1} = \frac{d}{dx}\left(\frac{x^{n+1}-1}{x-1}\right) = \frac{(n+1)x^n}{x-1} - \frac{x^{n+1}-1}{(x-1)^2}.$$

Thus,

$$\sum_{k=0}^{n} k\ x^k = \frac{(n+1)x^{n+1}}{x-1} - \frac{x^{n+2}-x}{(x-1)^2}$$

which simplifies to

$$\frac{1}{1-x}\left\{\frac{x^{n+1}-1}{x-1} - 1 - n\ x^{n+1}\right\}.$$

Example:

$$\frac{d}{dx}(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} k\ x^{k-1}.$$

Likewise, one can use integration to obtain new results.

Running Time Functions and O, Θ and Ω notations.

Definition: A function will be called a running time function if

$$f : Z^+ \longrightarrow R$$

such that $f(n) > 0$ for all $n \geq m$, where m is some positive integer. Recall Z^+ is $\{0, 1, 2, \ldots\}$.

Definition: (Use of O is made to obtain an upper bound.) Let f and g be two real time functions. We denote $f(n) = O(g(n))$, if there exist a real constant c and integer m such that

$$f(n) \leq cg(n)$$

for all $n \geq m$.

Definition: (Use of Ω is used to obtain a lower bound.) Let f and g be two real time functions. We denote $f(n) = \Omega(g(n))$, if there exist a real constant c and integer m such that

$$cg(n) \leq f(n)$$

for all $n \geq m$.

Definition: (Use of Θ is made to indicate a comparable function.) Let f and g be two real time functions and we denote $f(n) = \Theta(g(n))$. Then there exist positive real constants c_1 and c_2 such that

$$c_1 g(n) \leq f(n) \leq c_2 g(n)$$

for all $n \geq m$ for some positive integer m.

Examples

Example 1: Note that

$$\frac{x}{x-1} = 1 + \frac{1}{x} + \frac{1}{x^2} + \ldots$$

For large x

$$\begin{aligned} \frac{x}{x-1} &= 1 + \frac{1}{x} + o(\frac{1}{x}) \\ &= 1 + \frac{1}{x} + O(\frac{1}{x^2}) \end{aligned}$$

Example 2: Note that

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

Hence, for $n \longrightarrow \infty$

$$\sum_1^n i^2 = O(n^3)$$

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

Example Application: Insertion Sort

Given $A[1], \ldots, A[n]$, give a sorted sequence.

Solution: (i) If $n = 1$, the sequence $A[1]$ is sorted.

(ii) Suppose $A[1] < A[2] < \cdots < A[j-1]$ are already sorted. Pick up the next element and place it in its appropriate place.

Example: $3 < 5 < 6 < 9$; next $A[j] = A[5] = 4$. Then by "comparing" it with 9, then 6, then 5, then 4, we find its place between 3 and 5 and continue in this manner until all elements are properly arranged.

Insertion Sort

Pseudo Code:

for $j \leftarrow 2$ **to** $n =$ length $[A]$

  do key $\leftarrow A[j]$

comment: Insert $A[j]$ into the sorted sequence $A[1..j-1]$

  $i \leftarrow j - 1$

  while $i > 0$ and $A[i] >$ key

    do $A[i+1] \leftarrow A[i]$

    $i \leftarrow i - 1$

  $A[i+1] \leftarrow$ key

Analysis of Insertion Sort

Step	Cost	Times
1	c_1	n
2	c_2	$n-1$
3	—	—
4	c_4	$n-1$
5	c_5	$\Sigma_{j=2}^{n}\ t_j$
6	c_6	$\Sigma_{j=2}^{n}\ (t_j-1)$
7	c_7	$\Sigma_{j=2}^{n}\ (t_j-1)$
8	c_8	$(n-1)$

where t_j = number of times the while loop test in line 5 is executed for that value of j.

Let $T(n)$ denote the total time spent in performing the INSERTION SORT. Then

$$T(n) = c_1\ n + (c_2 + c_4)(n-1) + c_5 \sum_{j=2}^{n} t_j + (c_6 + c_7)(\sum_{j=2}^{n}(t_j - 1) + c_8(n-1),$$

Best, Average and Worst Case Analysis

Best Case Analysis: When A is sorted, then $t_j = 0$ for all j and

$$T(n) \;=\; \mathrm{a}n - \mathrm{b}$$

Worst Case Analysis: A is in reverse sorted order; $t_j = j$ for $j = 2, \ldots, n$.

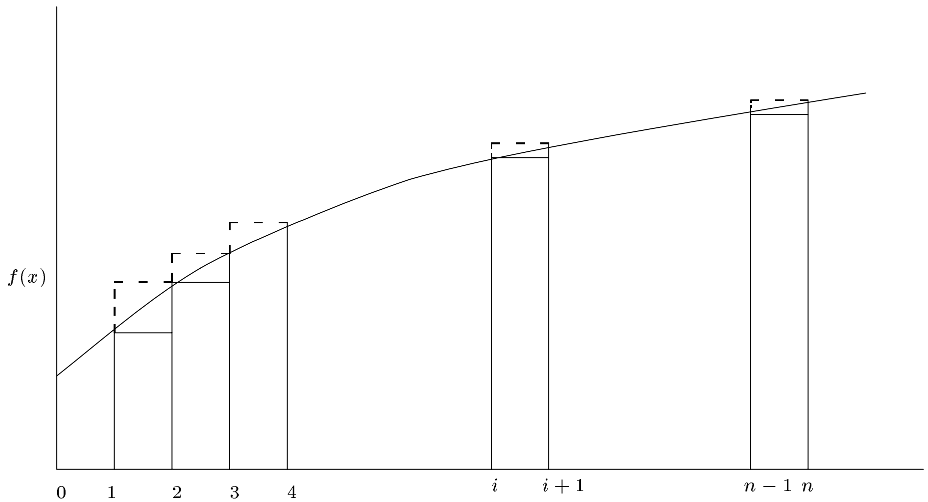
$$\sum_{j=1}^{n} t_j = \sum_{1}^{n} j = \frac{n(n+1)}{2},$$

$$T(n) \;=\; \mathrm{a}n^2 + \mathrm{b}n - \mathrm{c}$$

Average Case Analysis: Choose n numbers "randomly." Expected value of $t_j = \frac{j}{2}$.

Therefore, $\sum_{1}^{n} t_j = \dfrac{1}{4}\, n(n+1)$ (quadratic).

Comparing Sums and Integrals.



Let $f(x)$ be a monotone integrable function.

$$f(x_0)(x_1 - x_0) + f(x_1)(x_2 - x_1) + \cdots + f(x_{n-1})(x_n - x_{n-1})$$

$$\leq \int_{x_0}^{x_n} f(x)dx \leq f(x_1)(x_1 - x_0) + \cdots + f(x_n)(x_n - x_{n-1})$$

In particular, let $x_i = (i+1)$. Then

$$f(1) + f(2) + \cdots + f(n) \leq \int_1^{n+1} f(x)\ dx$$

$$\leq f(2) + \cdots + f(n+1) + [f(1) - f(1)].$$

Therefore,

$$\sum_{k=1}^{n} f(k) \leq \int_1^{n+1} \ f(x)dx$$

and

$$\int_1^{n+1} f(x)dx + f(1) - f(n+1) \leq \sum_{k=1}^{n} f(k).$$

Thus $\int_1^{n+1} \ f(x) + f(1) - f(n+1) \leq \sum_{k=1}^{n} \ f(k) \leq \int_1^{n+1} f(x)dx$, implying that an error in approximation is no more than $f(n+1) - f(1)$.

Examples

Example 1. Let $f(x) = x^\ell$.

$$\int_1^{n+1} x^\ell dx - (n+1)^\ell + 1 \leq \sum_{k=1}^{n} k^\ell \leq \int_1^{n+1} x^\ell dx,$$

$$\left\{\frac{(n+1)^{\ell+1}}{\ell+1} - \frac{1}{\ell+1} - (n+1)^\ell + 1\right\} \leq \sum_{k=1}^{n} k^\ell \leq \left\{\frac{(n+1)^{\ell+1}}{\ell+1} - \frac{1}{\ell+1}\right\}.$$

Lhs

$$= \frac{1}{\ell+1}\left[n^{\ell+1} + \binom{\ell+1}{1}n^{\ell} + \binom{\ell+1}{2}n^{\ell-1} + \cdots + \binom{\ell+1}{\ell+1}\right]$$

$$- \left[n^{\ell} + \binom{\ell}{1}n^{\ell-1} + \cdots + \binom{\ell}{\ell}\right] + 1 - \frac{1}{\ell+1}$$

$$= \frac{n^{\ell+1}}{\ell+1} + \left[\frac{\binom{\ell+1}{1}}{\ell+1} - 1\right] n^{\ell} + \left[\frac{\binom{\ell+1}{2}}{\ell+1} - \binom{\ell}{1}\right] n^{\ell-1} + \cdots$$

$$= \frac{n^{\ell+1}}{\ell+1} + O(n^{\ell}).$$

Likewise, we can show that the leading term in the right hand side expression is also $\frac{n^{\ell+1}}{\ell+1}$. Consequently, $\Sigma_{k=1}^{n} k^{\ell} = \frac{n^{\ell+1}}{\ell+1} + O(n^{\ell})$

Example 2. $f(x) = \frac{1}{x}$.

Note that this is a monotone decreasing function. Hence

$$f(1) + f(x) + \cdots + f(n) \geq \int_1^{n+1} f(x)dx \geq f(2) + \cdots + f(n+1).$$

$$H_n = \sum_{i=1}^{n} \frac{1}{i} \geq \int_1^{n+1} \frac{1}{x}\, dx \geq \sum_{i=1}^{n} \frac{1}{i} + \left\{ \frac{1}{n+1} - 1 \right\}.$$

$$H_n \geq \ln(n+1) - \ln 1 \geq H_n - \frac{n}{n+1}.$$

$$\ln(n+1) \leq H_n \leq \ln(n+1) + \frac{n}{n+1}.$$

$$\begin{aligned} \ln(n+1) &= \ln\left[n\tfrac{n+1}{n}\right] = \ln\left[n(1+\tfrac{1}{n})\right] \\ &= \ln n + \ln(1+\tfrac{1}{n}) \\ &= \ln n + \tfrac{1}{n} - \tfrac{1}{2n^2} + \tfrac{1}{3n^3} \cdots \end{aligned}$$

$$H_n = \ln n + O(1).$$