Department of Mathematical Sciences
The Johns Hopkins University

SEMINAR

Jim Fill
Department of Mathematical Sciences
The Johns Hopkins University

April 24, 2003
304 Whitehead Hall
Refreshments: 3:30 p.m.
Seminar: 4:00 p.m.

THE NUMBER OF BIT COMPARISONS USED BY QUICKSORT

Abstract

The efficiency of digital (or radix) sorting can be measured by the number of bit inspections required to sort a file. The efficiency of a sorting algorithm such as Quicksort is usually measured by the number of key comparisons required. Many years ago, Bob Sedgewick asked for a fair comparison between digital and non-digital methods. In particular, how many *bit* comparisons are required to sort a file of n distinct keys using Quicksort?

We suppose throughout that the keys are independent and identically distributed draws from $(0, 1)$ according to density f. Suppose first that f is uniform. We obtain exact formulas and an asymptotic expansion for μ_n, the expected number of bit comparisons required; in particular,

$$\mu_n = n(\ln n)(\lg n) + c_1 n \ln n + c_2 n + \pi_n n + O(\log n),$$

where ln and lg denote natural and binary logarithm, respectively; c_1 and c_2 are certain (explicit but somewhat messy) constants; and π_n is periodic in $\ln n$ with magnitude bounded by 5×10^{-9}.

More generally, if f is any density with

$$\int f(\log f)^{4+\epsilon} < \infty$$

for some $\epsilon > 0$, we show that ν_n, the expected number of bit comparisons required, satisfies

$$\nu_n = \mu_n + 2n(\ln n) \int f \lg f + o(n \log n),$$

with μ_n as above.

(This is joint work with Svante Janson of Uppsala University in Sweden.)