Analysis of Randomized Quicksort

Marcus Schaefer
Department of Computer Science
DePaul University
Chicago, Illinois 60604, USA
mschaefer@cs.depaul.edu

September 18, 2006

1 Analysis of Randomized Quicksort

We analyze the running time of randomized quicksort as presented on page 146 of CLRS. Let $T(n)$ be the time taken on average by randomized quicksort on n-element arrays. Note that if we partition around the pth element, the algorithm will take time

$$T(p-1) + T(n-p) + cn.$$

Since we picked the partition element at random, every value of p from 1 to n is equally likely, so

$$\begin{aligned} T(n) &= 1/n \sum_{p=1}^{n} (T(p-1) + T(n-p) + cn) \\ &= cn + 1/n \sum_{p=1}^{n} (T(p-1) + T(n-p)) \end{aligned}$$

Now

$$\begin{aligned} \sum_{p=1}^{n} (T(p-1) + T(n-p)) &= (T(0) + T(n-1)) + (T(1) + T(n-2)) + \ldots + (T(n-1) + T(0)) \\ &= 2 \sum_{p=0}^{n-1} (T(p) \end{aligned}$$

by reordering terms, so

$$
\begin{aligned}
T(n) &= cn + 1/n \sum_{p=1}^{n} (T(p-1) + T(n-p)) \\
&= cn + 2/n \sum_{p=0}^{n-1} T(p)
\end{aligned}
$$

Multiply both sides by n to obtain

$$nT(n) = cn^2 + 2 \sum_{p=0}^{n-1} T(p).$$

Compare this to the same equation for $n-1$ in place of n:

$$(n-1)T(n-1) = c(n-1)^2 + 2 \sum_{p=0}^{n-2} (T(p).$$

Subtract the second equation from the first, and you get

$$nT(n) - (n-1)T(n-1) = 2nc - c + 2T(n-1),$$

since $\sum_{p=0}^{n-1}(T(p) - \sum_{p=0}^{n-2}(T(p) = T(n-1)$. We reorder the terms to get

$$nT(n) = 2nc - c + (n+1)T(n-1).$$

To simplify we drop the c:

$$nT(n) \leq 2nc + (n+1)T(n-1).$$

Dividing both sides by $n(n+1)$ gets us

$$T(n)/(n+1) \leq 2c/(n+1) + T(n-1)/n.$$

Since then

$$T(n-1)/n \leq 2c/n + T(n-2)/(n-1),$$

we can conclude that

$$
\begin{aligned}
T(n)/(n+1) &\leq 2c/(n+1) + T(n-1)/n \\
&\leq 2c/(n+1) + 2c/n + T(n-2)/(n-1).
\end{aligned}
$$

Continuing like this gives us

$$T(n)/(n+1) \leq 2c \sum_{p=1}^{n+1} 1/p.$$

We now use the fact that

$$\sum_{p=1}^{n+1} 1/p = O(\log n),$$

(see equation A.7 in CLRS [Appendix A]), to conclude that

$$T(n)/(n+1) = O(\log n),$$

or, in other words,

$$T(n) = O(n \log n).$$