Mergesort analysis

Max Berger

June 8, 2004

In this analysis we are counting data comparisons only. We found out that

$$datacomp(mergesort(n)) \leq \begin{cases} 0 & \text{if } n \leq 1 \\ 0 + 2 * datacomp(mergesort(\frac{n}{2})) + n & \text{if } n > 1 \end{cases}$$

Since we are interested in analyzing the worst case, we can assume $=$ instead of $\leq$. Rewritten, the equation looks like this:

$$\begin{aligned} D(1) &= 0 \\ D(n) &= 2D(\frac{n}{2}) + n \end{aligned}$$

to find the values to plug in, we need to find $D(\frac{n}{2})$, $D(\frac{n}{3})$, $D(\frac{n}{4})$, ...

$$\begin{aligned} D(\frac{n}{2}) &= 2D(\frac{n}{4}) + \frac{n}{2} \\ D(\frac{n}{4}) &= 2D(\frac{n}{8}) + \frac{n}{4} \\ D(\frac{n}{8}) &= 2D(\frac{n}{16}) + \frac{n}{8} \\ D(\frac{n}{16}) &= 2D(\frac{n}{32}) + \frac{n}{16} \\ D(\frac{n}{32}) &= 2D(\frac{n}{64}) + \frac{n}{32} \\ &\cdots \end{aligned}$$

If we plug these in recursively into the original equation we get:

$$\begin{aligned} D(n) &= 2D(\frac{n}{2}) + n \\ D(n) &= 2(2D(\frac{n}{4}) + \frac{n}{2}) + n \\ &= 4D(\frac{n}{4}) + n + n \\ D(n) &= 4(2D(\frac{n}{8}) + \frac{n}{4}) + n + n \\ &= 8D(\frac{n}{8}) + n + n + n \\ D(n) &= 8(2D(\frac{16}{8}) + \frac{n}{8}) + n + n + n \\ &= 16D(\frac{n}{16}) + n + n + n + n \\ D(n) &= 16(2D(\frac{n}{32}) + \frac{n}{16}) + n + n + n + n \\ &= 32D(\frac{n}{32}) + n + n + n + n + n \end{aligned}$$

This seems to develop a pattern. The pattern is:

$$\begin{aligned} D(n) &= nD(1) + \sum_{i=1}^{\lg_n} n \\ &= 0n + n \sum_{i=1}^{\lg n} 1 \\ &= n \lg n \end{aligned}$$

So mergesort uses $n \lg n$ comparisons in the worst case.