Selection Algorithm Proof, size 3 rather than 5

Elita Danilyuk

The following is an examination of the SELECTION algorithm within section 9.3, but with groups of 3 elements, rather than 5, and ignoring floors and ceilings.

Determining a lower bound, with the help of a visual:

At least half of the $\frac{n}{3}$ groups are greater than or equal to x, the median-of-medians. Thus, at least half of the $\frac{n}{3}$ groups contribute at least 2 elements that are greater than x (ignoring the 2 edge cases). By disregarding the 2 edge cases, it follows that the number of elements greater than x is at least:

$$2\left(\frac{1}{2}\left(\frac{n}{3}\right)\right) \geq \frac{n}{3}$$

$\frac{n}{3} \cdot \frac{1}{2} \times 2 = \frac{n}{3} \approx 33.\overline{3}\%$

Similarly, at least $\frac{n}{3}$ elements are less than x. Thus, in the worst case step 5 calls SELECT recursively, $T_5(n)$, on at most $\frac{2n}{3}$.

Developing a recurrance for the running time, $T(n)$ for SELECTION:

$T(n) = T_1 + T_2 + T_3 + T_4 + T_5$

$$\left.\begin{array}{l} T_{1+2+4}(n) = \Theta(n) \\ T_3(n) = T\left(\frac{n}{3}\right) \\ T_5(n) = T\left(\frac{2n}{3}\right) \leftarrow 66.\overline{6}\% \end{array}\right\} \quad T(n) = \Theta(n) + T\left(\frac{n}{3}\right) + T\left(\frac{2n}{3}\right)$$

Thus the recurrance is:

$$T(n) \leq \begin{cases} O(n \log n) \\ T\left(\frac{n}{3}\right) + T\left(\frac{2n}{3}\right) + O(n) \end{cases}$$

The running time is shown to be $O(n \log n)$ by substitution, specifically $T(n) \leq cn\log n$ for a constant c and all $n>0$. A constant a such that the function described by the $O(n)$ term above is bounded by an for all $n>0$. Through substitution the recurrance becomes:

$$\begin{aligned} T(n) &\leq c\left(\frac{n}{3}\right) + c\left(\frac{2n}{3}\right) + an \\ &\leq c\left(\frac{n}{3}\right)\log\left(\frac{n}{3}\right) + c\left(\frac{2n}{3}\right)\log\left(\frac{2n}{3}\right) + an \\ &= cn\log n + an \end{aligned}$$

Therefore, when groups of 3 are used, the algorithm is no longer linear in its worst case (like groups of 5), but it grows faster at a rate of $O(n \log n)$.