Homework: Algorithms — Order Statistics

If you use the median of median trick for order statistics on group of size 7, the argument shows that SELECT is called on at most

$$4\left(\left\lceil \frac{1}{2}\left\lceil \frac{n}{7}\right\rceil \right\rceil - 2\right) \geq \frac{4}{14}n - 8$$

elements are larger than the pivot and at least that many smaller than the pivot. Therefore, SELECT is called on at most $\frac{5}{7}n + 8$ elements. We can assume that any input smaller than 500 elements requires $O(1)$ time. This gives us the recurrence for the runtime of SELECT on n elements input as

$$T(n) \leq \begin{cases} O(1) & \text{if } n < 500 \\ T(\lceil \frac{n}{7} \rceil) + T(\frac{5}{7}n + 8) + an & \text{otherwise} \end{cases}.$$

In this equation, a represents the costs of grouping the array into groups of seven.

Problem: Show that there exists a constant $c > 0$ such that $T(n) \leq cn$.

Solution

For $n < 500$, we can find c large enough such that $T(n) < cn$. We now assume that for all $m < n$, we have $T(m) \leq cm$. We then have for $n > 500$:

$$\begin{aligned} T(n) &\leq c\lceil \frac{n}{7} \rceil + c\left(\frac{5}{7}n + 8\right) + an \\ &\leq \frac{n}{7}c + c + c(\frac{5}{7}n + 8) + an \\ &= \frac{6}{7}cn + 9c + an \end{aligned}$$

We want to conclude that this is smaller than cn. Since $n > 500$, we have $\frac{1}{7}n > 18$, which implies $\frac{1}{7}n - 9 > 2$. We pick $c > 7a\frac{500}{500 - 9 \cdot 7}$. Then

$$
\begin{aligned}
c > 7a\frac{500}{500-63} \quad &\implies \quad \frac{500-63}{500}c > 7a \\
&\implies \quad c > \frac{63}{500}c + 7a \\
&\implies \quad c > \frac{63}{n}c + 7a \\
&\implies \quad cn > 63c + 7an \\
&\implies \quad \frac{1}{7}cn > 9c + an
\end{aligned}
$$

Therefore, $\frac{6}{7}cn + 9c + an < \frac{6}{7}cn + \frac{1}{7}cn = cn$. Since we know that $T(n) < \frac{6}{7}cn + 9c + an$, the induction hypothesis follows.