Homework: Algorithms — Order Statistics

The median of median SELECT algorithm divides all elements in groups of 5. If instead we use a group size of 7, the same argument shows that at least

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

elements are larger than the pivot. Similarly, $(4/14)n - 8$ are smaller than the pivot. Therefore, SELECT is called on at most $(5/7)n + 8$ elements. We can assume that any input smaller than 500 elements (to pick just a large number), requires $O(1)$ time. This gives us the recurrence for the runtime of SELECT on n elements as

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

Here, the first element of the sum is the selection of the median of medians to determine the pivot. The second one comes from the recursive call on, which is applied to at most

$$(5/7)n + 8$$

elements. The third addend reflects the division of all arrays into groups of 7, the selection of the median of each group, and the partition around the median of medians.

Following the procedure of the book, end of Chapter 9, prove that the existence of 6a constant $c > 0$ such that $T(n) \leq cn$.