PS1-Q1

Dylan Mavrides

February 14, 2017

5/5

How many subarrays of size 2 or less are encountered, on the average, when sorting a random file of size N with quicksort? When we have an n-element subarray called, the recurrence representing this situation is

$$A_N = \frac{2}{N} \sum_{1 \le j \le N} A_{j-1}$$

since it's the same as the recurrences we've solved in class/question 1.14, but without "adding anything" (as with the "1 +" in the second one) until we hit the base case (when size = n). Thus we have

$$NA_N = 2 \sum_{1 \le j \le N} A_{j-1}$$

$$NA_N - (N-1)A_{N-1} = 2A_{N-1}$$

$$A_N = \frac{N+1}{N} A_{N-1}.$$

Now, since we are looking for subarrays of size 2 or less, we will look for subarrays of size 2, 1, and 0 separately and add the expected numbers.
Size 2:
Telescoping, we see that

$$A_N = \frac{N+1}{4} A_3.$$

We note that for A_0 and A_1 we have 0 calls with subarrays size 2, for A_2 we have 1 (the whole thing), and for A_3 we have 2/3 expected (choosing the first or last element gives a subarray of size 0 and size 2 called, one on either size, thus 1*1/3 + 1*1/3, but the middle element gives two 1-element subarrays called = 0*1/3, thus 2/3 total). (Note that since earlier in the problem we subtract the N - 1 case, we must use the upper of two legitimate base cases. Normally N=2 would work, but we must use 3 for it to work here for the stated reason.) Thus

$$A_N = \frac{2(N+1)}{3*4} = \frac{N+1}{6}.$$

Size 1:
Telescoping, we also see that

$$A_N = \frac{N+1}{3} A_2.$$

We note that for A_0 we have 0 calls with hi=lo, for A_1 we have 1, for A_2 we have 1 (one element is choosen as the pivot). (Note that since earlier in the problem we subtract the N - 1 case, we must use the upper of two legitimate base cases. Normally N=1 would work, but we must use 2 for it to work here for the stated reason.) Thus

$$A_N = \frac{N+1}{3}.$$

Size 0:
Telescoping, we also see that

$$A_N = \frac{N+1}{4} A_3.$$

We note that for A_0 we have 1 size-0 subarray called. For A_1 we have 0 as well, since Quick terminates after calling either a 1 or a 0 element array. For 2 we have 1, since either element it chooses will result in one 0-element subarray called on one side, and a 1-element subarray on the other. For 3 we have 4/3, because we have 2/3 for either end element (1 for the size 0 subarray, 1 for the size 2 subarray), and 0 for the middle element. Thus we have

$$A_N = \frac{4(N+1)}{3*4} = \frac{N+1}{3}.$$

Thus (except for some very small values of N, where our base cases were above them, but in general) the expected number of calls to arrays less than or equal to size 2 is $\frac{N+1}{3} + \frac{N+1}{6} + \frac{N+1}{3} = \frac{5(N+1)}{6}$.

In the future please list out the small cases explicitly i.e. 1, 1, 3