CS 344: Design and Analysis of Computer Algorithms **Rutgers: Fall 2019**

Lecture 4

September 16, 2019

Instructor: Sepehr Assadi

1 Sorting via Divide and Conquer: Merge Sort

In this lecture, we analyze an efficient algorithm for sorting called the *merge sort*. The technique behind designing this algorithm is typically referred to as *divide and conquer*. The overall idea is as follows:

1. Divide the given instance of the problem into several smaller instances of the same problem.
2. Recursively solve the problem on each smaller instance.
3. Combine the solutions for the smaller instances into the final solution.

If the size of any instance falls below some constant threshold, we abandon recursion and solve the problem directly, by brute force, in constant time.

The proof of correctness of a divide and conquer algorithm, similar to other recursive algorithms, almost always requires induction. Analyzing the running time requires setting up and solving a recurrence as we already saw for binary search in the previous lecture.

The Merge Algorithm

Before getting to merge sort, let us consider the *merge* algorithm, used as the third step in the above divide and conquer approach for designing merge sort.

The merge algorithm solves the following problem. Suppose we are given an array $A[1:n]$ where $A[1:\lfloor n/2 \rfloor]$ and $A[\lfloor n/2 \rfloor + 1 : n]$ are *sorted separately*. The goal of this algorithm is to sort the entire array A.

Example 1: For instance, the input to merge can be the following array $A = [5, 6, 7, 8, 1, 2, 3, 4]$. The first half of the array is $[5, 6, 7, 8]$ and the second one is $[1, 2, 3, 4]$ – each half is sorted separately but together they are not sorted anymore. The goal is then to return the array $[1, 2, 3, 4, 5, 6, 7, 8]$.

Example 2: Another example is the input $A = [1, 2, 7, 8, 3, 4, 5, 6]$ – again the first half and the second half are sorted separately but the whole array is not. The goal is then to return the array $[1, 2, 3, 4, 5, 6, 7, 8]$.

Merging an array that each of its half are sorted seems like an easier task than sorting the array from the scratch (after all, some part of the work is already done!). Intuitively, we can simply do as follows: We pick another empty array B; then we go over the elements in $A[1:\lfloor n/2 \rfloor]$ and $A[\lfloor n/2 \rfloor + 1 : n]$ one by one simultaneously and place the smaller one in array B and move the pointer in the array where this element come from one step further and continue like this.

We formalize the algorithm as follows:

Merge Algorithm: The input is an array A where $A[1:\lfloor n/2 \rfloor]$ and $A[\lfloor n/2 \rfloor + 1 : n]$ are sorted separately.

1. Create an empty array B and two pointers $p_1 = 1$ and $p_2 = \lfloor n/2 \rfloor + 1$.

2. For $i = 1$ to n do as follows:
   (a) If ($p_1 \neq \lfloor n/2 \rfloor + 1$ AND $A[p_1] < A[p_2]$) OR $p_2 = n + 1$, let $B[i] = A[p_1]$ and set $p_1 \leftarrow p_1 + 1$;
   (b) Otherwise, let $B[i] = A[p_2]$ and set $p_2 \leftarrow p_2 + 1$.
3. Return the array B as the answer.

Example: Let us consider running this algorithm on the input $A = [1, 2, 7, 8, 3, 4, 5, 6]$. In the following, the underlined numbers denote the target of the pointers p_1 and p_2.

- At the beginning:

$$A[1:4] = [\underline{1}, 2, 7, 8] \qquad A[5:8] = [\underline{3}, 4, 5, 6] \qquad B = [-, -, -, -, -, -, -, -];$$

- After iteration $i = 1$ (since $A[1] = 1 < 3 = A[5]$):

$$A[1:4] = [1, \underline{2}, 7, 8] \qquad A[5:8] = [\underline{3}, 4, 5, 6] \qquad B = [1, -, -, -, -, -, -, -];$$

- After iteration $i = 2$ (since $A[2] = 2 < 3 = A[5]$):

$$A[1:4] = [1, 2, \underline{7}, 8] \qquad A[5:8] = [\underline{3}, 4, 5, 6] \qquad B = [1, 2, -, -, -, -, -, -];$$

- After iteration $i = 3$ (since $A[3] = 7 > 3 = A[5]$):

$$A[1:4] = [1, 2, \underline{7}, 8] \qquad A[5:8] = [3, \underline{4}, 5, 6] \qquad B = [1, 2, 3, -, -, -, -, -];$$

- After iteration $i = 4$ (since $A[3] = 7 > 4 = A[6]$):

$$A[1:4] = [1, 2, \underline{7}, 8] \qquad A[5:8] = [3, 4, \underline{5}, 6] \qquad B = [1, 2, 3, 4, -, -, -, -];$$

- After iteration $i = 5$ (since $A[3] = 7 > 5 = A[7]$):

$$A[1:4] = [1, 2, \underline{7}, 8] \qquad A[5:8] = [3, 4, 5, \underline{6}] \qquad B = [1, 2, 3, 4, 5, -, -, -];$$

- After iteration $i = 5$ (since $A[3] = 7 > 6 = A[8]$):

$$A[1:4] = [1, 2, \underline{7}, 8] \qquad A[5:8] = [3, 4, 5, 6], \underline{*} \qquad B = [1, 2, 3, 4, 5, 6, -, -];$$

- After iteration $i = 6$ (since $p_2 = 9$):

$$A[1:4] = [1, 2, 7, \underline{8}] \qquad A[5:8] = [3, 4, 5, 6], \underline{*} \qquad B = [1, 2, 3, 4, 5, 6, 7, -];$$

- After iteration $i = 6$ (since $p_2 = 9$):

$$A[1:4] = [1, 2, 7, 8], \underline{*} \qquad A[5:8] = [3, 4, 5, 6], \underline{*} \qquad B = [1, 2, 3, 4, 5, 6, 7, 8].$$

Proof of Correctness: We now prove the correctness of this algorithm. Even though the algorithm is not recursive, we again prove this by *induction* over the index i of the for-loop. Our induction hypothesis is that for every iteration $i = 1$ to n, $B[1:i]$ contains the smallest first i elements of A in the sorted array (after the iteration)[1]. By proving this hypothesis we will be done as it means that after iteration $i = n$, $B[1:n]$ contains the sorted version of the array A. We now prove the induction hypothesis.

[1]Strictly speaking this argument is not an induction if we want to be really formal because instead of proving the claim for all integers i (where i can go to infinity as in an induction) we are only proving it for integers up to n; however, for the purpose of this course, it is completely fine to call this also an induction.

For $i = 1$, the algorithm places $\min(A[1], A[\lfloor n/2 \rfloor + 1])$ in $B[1]$: since both $A[1 : \lfloor n/2 \rfloor]$ and $A[\lfloor n/2 \rfloor + 1 : n]$ are sorted, $B[1]$ is the minimum of the array and hence the induction hypothesis holds. Suppose now that the induction hypothesis holds till iteration i and we prove it for iteration $i + 1$ (the induction step).

Consider $A[p_1 : \lfloor n/2 \rfloor]$ and $A[p_2 : n]$ at the beginning of iteration $i+1$. Since by induction hypothesis $B[1 : i]$ already contains the first smallest i integers in A, and since all elements in $A[1 : p_1 - 1]$ and $A[\lfloor n/2 \rfloor : p_2 - 1]$ are already placed in the array B (by definition of how we update the pointers p_1 and p_2), we only need to ensure that in this iteration, $B[i+1]$ becomes the minimum of the elements in $A[p_1 : \lfloor n/2 \rfloor] \cup A[p_2 : n]$. This happens since both $A[p_1 : \lfloor n/2 \rfloor]$ and $A[p_2 : n]$ are sorted and thus minimum of the union is either $A[p_1]$ or $A[p_2]$: considering the algorithm checks for out of range values of p_1 and p_2 and then places $\min(A[p_1], A[p_2])$ in $B[i+1]$, we obtain that $B[1 : i+1]$ also consists of the first $(i+1)$ elements of A in the sorted order. This proves the induction and the correctness of the merge algorithm.

Runtime Analysis: This is quite easy: the algorithm simply has a for-loop of n iterations and each iteration takes $O(1)$ time, so the runtime is $O(n)$.

The Merge Sort Algorithm

We now use the merge algorithm we already designed in the divide and conquer framework to obtain an algorithm for sorting. Recall that the merge algorithm could sort an array A in $O(n)$ time *assuming each half of the array were sorted already.* But what if we want to sort an arbitrary array? Can we first sort each of its half separately and then provide this new array to the merge algorithm to finish the problem? But of course we can do that! How? By using *recursion.*

The entire merge sort algorithm is simply as follows: if the size of the array is a small constant simply solve the problem by brute force. Otherwise, recursively merge sort the first half and the second half separately. Finally, run the merge algorithm on each part to obtain the final sorted array. Formally,

Merge Sort Algorithm: The input is an array A.

1. If the size of the array is $n = 0$ or 1 return the array as it is.
2. Recursively run merge sort on $A[1 : \lfloor n/2 \rfloor]$ and $A[\lfloor n/2 \rfloor + 1 : n]$.
3. Run the merge algorithm on A, copy the returned array of this algorithm back into A, and return A.

Example: Let us consider running this algorithm on the input $A = [5, 3, 7, 1, 8, 2, 4, 6]$.

- The splitting (recursive call) steps are as follows:

$[5, 3, 7, 1, 8, 2, 4, 6]$							
$[5, 3, 7, 1]$	$[8, 2, 4, 6]$						
$[5, 3]$	$[7, 1]$	$[8, 2]$	$[4, 6]$				
$[5]$	$[3]$	$[7]$	$[1]$	$[8]$	$[2]$	$[4]$	$[6]$

- Then these elements will start to get merged together in the following manner:

$[5]$	$[3]$	$[7]$	$[1]$	$[8]$	$[2]$	$[4]$	$[6]$
$[3, 5]$	$[1, 7]$	$[2, 8]$	$[4, 6]$				
$[1, 3, 5, 7]$	$[2, 4, 6, 8]$						
$[1, 2, 3, 4, 5, 6, 7, 8]$							

Proof of Correctness: As we proved the correctness of the merge algorithm already, this step is quite easy. We prove, by induction, that for every n, the merge sort algorithm sorts any given array A of n numbers. The base case for $n \in \{0,1\}$ is true since any array of size 0 or 1 is already sorted and hence the correct answer is just to return the array as it is.

We now prove the induction step: assuming the hypothesis is true for all $n \leq i$, we prove this for $n = i+1$. The algorithm firstly partitions the array into $A[1 : \lfloor n/2 \rfloor]$ and $A[\lfloor n/2 \rfloor + 1 : n]$ and recursively merge sort each one; by induction hypothesis, since size of each of these arrays is smaller than n, the recursive calls correctly sort each of them. At this point, the array A satisfies the requirement of the input for the merge algorithm and thus the output of the merge is the array A sorted entirely correctly. This implies that the output of the merge sort algorithm is correct, proving the induction step.

Runtime Analysis: We will write a *recurrence* for analyzing the runtime of this algorithm. Let $T(n)$ denote the worst-case time of running merge sort on any array of length n. We will prove that,

$$T(n) \leq T(\left\lfloor \frac{n}{2} \right\rfloor) + T(\left\lceil \frac{n}{2} \right\rceil) + O(n).$$

(note that $T(n)$ is a recursive formula and technically we should also specify what is $T(n)$ for $n = \Theta(1)$ for the above formula to make sense; however, it is *always* the case that the running time of an algorithm on a constant size input is constant, namely, $T(\Theta(1)) = \Theta(1)$. Hence, *unlike* many other base cases, say, induction or recursive algorithms, here we can actually omit the recurrence base of a running time because it is always constant and so we can just implicitly assume that).

Let us first argue that $T(n)$ correctly captures the running time of the merge sort. The merge sort algorithm involves two recursive calls on arrays $A[1 : \lfloor n/2 \rfloor]$ (with size $\lfloor n/2 \rfloor$) and $A[\lfloor n/2 \rfloor + 1 : n]$ (with size $n - \lfloor n/2 \rfloor = \lceil n/2 \rceil$); moreover, it also involves running the merge algorithm that requires $O(n)$ time, and copying the resulting array to A which again can be done in $O(n)$ time. As such, the formula above for $T(n)$ is correct.

Having computed $T(n)$ recursively, we now want to turn it into a *closed-form* formula, namely, a non-recursive and "easy to state" function of n – this is crucial for better understanding the running time of the algorithm and potentially comparing it with other algorithms. Before we get to this however, we make an important remark about floors and ceilings in computing recurrences.

Ignoring Floors and Ceilings in Recurrences: When working with recurrences for analyzing runtime of algorithms, it is quite cumbersome to take into account the exact details of floors and ceilings. Fortunately, we can safely strip out the floors and ceilings in these formulas and for instance write that the worst-case running time of merge sort follows:

$$T(n) \leq T(n/2) + T(n/2) + O(n) = 2 \cdot T(n/2) + O(n).$$

Intuitively, this is because we are anyway upper bounding this function asymptotically and one can "charge" the extra numbers coming out of floors and ceilings to the non-recursive term in the formula which itself is stated asymptotically (here the $O(n)$-term). To be sure, this requires a formal proof but we will not do that in this course (the proof can be found both in the CLRS book (Chapter 4) and the suggested reading book (Chapter 1.7)). Consequently, throughout this course, *we will always drop the floors and ceilings from recurrence relations for running times.*

Now back to analyzing runtime of merge sort. After ignoring floor and ceilings, our goal is to compute a closed-form formula for $T(n) \leq 2T(n/2) + O(n)$. This is a common recurrence and its correct answer is $T(n) = O(n \log n)$. This implies that merge sort runs in $O(n \log n)$ time (which is much faster than simple sorting algorithms such as selection sort or insertion sort). But how did we end up "solving" $T(n)$? This is the topic of the next lecture.