

1 Check in and PS3

Discuss questions, if you have any, with the tutor and the rest of the class, about the material and content so far.

2 Problems

Problem 1. QuickSort Review

- (a) Suppose that the pivot choice is the median of the first, middle and last keys, can you find a bad input for QuickSort?
- (b) Are any of the partitioning algorithms we have seen for QuickSort stable? Can you design a stable partitioning algorithm? Would it be efficient?
- (c) Consider a QuickSort implementation that uses the 3-way partitioning scheme (i.e. elements equal to the pivot are partitioned into their own segment).
 - i) If an input array of size n contains all identical keys, what is the asymptotic bound for QuickSort?
 - ii) If an input array of size n contains $k < n$ distinct keys, what is the asymptotic bound for QuickSort?

Problem 2. (A few puzzles involving duplicates or array processing)

For each problem, try to come up with the most efficient algorithm and provide the time complexity for your solution.

- (a) Given an array A, decide if there are any duplicated elements in the array.
- (b) Given an array A, output another array B with all the duplicates removed. Note the order of the elements in B does not need to follow the same order in A. That means if array A is $\{3, 2, 1, 3, 2, 1\}$, then your algorithm can output $\{1, 2, 3\}$.
- (c) Given arrays A and B, output a new array C containing all the distinct items in both A and B. You are given that array A and array B already have their duplicates removed.
- (d) Given array A and a target value, output two elements x and y in A where $(x + y)$ equals the target value.

Problem 3. Child Jumble

Your aunt and uncle recently asked you to help out with your cousin's birthday party. Alas, your cousin is three years old. That means spending several hours with twenty rambunctious three-year-olds as they race back and forth, covering the floors with paint and hitting each other with plastic beach balls. Finally, it is over. You are now left with twenty toddlers who each need to find their shoes. And you have a pile of shoes that all look about the same. The toddlers are not helpful. (Between exhaustion, too much sugar, and being hit on the head too many times, they are only semi-conscious.)

Luckily, their feet (and shoes) are all of slightly different sizes. Unfortunately, they are all very similar, and it is very hard to compare two pairs of shoes or two pairs of feet to decide which is bigger. (Have you ever tried asking a grumpy and tired toddler to line up their feet carefully with another toddler to determine who has bigger feet?) As such, you cannot compare shoes to shoes or feet to feet.

The only thing you can do is to have a toddler try on a pair of shoes. When you do this, you can figure out whether the shoes fit, or if they are too big, or too small. That is the only operation you can perform.

Come up with an efficient algorithm to match each child to their shoes. Give the time complexity of your algorithm in terms of the number of children.

Problem 4. More Pivots!

QuickSort is pretty fast. But that was with one pivot. In fact, QuickSort can also be implemented with two or more pivots! In this question, we will investigate the asymptotic running time of QuickSort when there are k pivots.

- (a) Suppose that you have a magic black box function that chooses k perfect pivots that separate the elements evenly (e.g. it picks the quartile elements when there are 3 pivots). How would a partitioning algorithm work using these pivots?
- (b) What is the asymptotic running time of your partitioning algorithm? Give your answer in terms of the number of elements, n and the number of pivots, k .
- (c) We can implement QuickSort using the partitioning algorithm you devised by recursing on each partition. Formulate a recurrence relation that represents the asymptotic running time of QuickSort with k pivots. Give your answer in terms of the number of elements, n and the number of pivots, k .
- (d) Solve the recurrence relation to obtain the asymptotic running time of your QuickSort with k pivots. Would using more pivots result in an improvement in the asymptotic running time?

Problem 5. Integer Sort

- (a) Given an array consisting of only 0's and 1's, what is the most efficient way to sort it? (Hint: Consider modifying a sorting algorithm that you have already learnt to achieve a running time of $O(n)$ regardless of the order of the elements in the input array.)

Can you do this in-place? If it is in-place, is it also stable? (You should think of the array as containing key/value pairs, where the keys are 0's and 1's, but the values are arbitrary.)

- (b) Consider an array A consisting of elements with keys being integers between 0 and M , where M is a small integer (for example, imagine an array containing key/value pairs, with all keys in the set $\{0, 1, 2, 3, 4\}$). What is the most efficient way to sort it? This time, you do not have to do it in-place; you can use extra space to record information about the input array and you can use an additional array to store the output.
- (c) Consider the following sorting algorithm for sorting integers represented in binary (each specified with the same number of bits):

- (1) First, use the in-place algorithm from part (a) to sort by the most significant bit, or MSB. That is, use the MSB of each integer as the key to sort with.
- (2) Once this is done, you have divided the array into two parts: the first part contains all the integers that begin with 0 and the second part contains all the integers that begin with 1. In other words, all the elements of the (binary) form $0xxxxxxx$ will come before all the elements of the (binary) form $1xxxxxxx$.
- (3) Now, sort the two parts using the same algorithm, but using the 2nd bit instead of the 1st. Then, sort each of the resulting parts using the 3rd bit, and so on.

Assuming that each integer is 64 bits, what is the running time of this algorithm? When do you think this sorting algorithm would be faster than QuickSort? If you want to, write some code and test it out.

- (d) Can you improve on this by using the algorithm from part (b) instead to do the partial sorting? What are the trade-offs involved?