

Module 10 - Discussion Prompt Questions

Adam Rich

April 15, 2018

Section EN.605.202.87.SP18

1. What are some key factors that differentiate the various sorts you have seen?

- a) Do they work better on random, pre-sorted, or reverse sorted data with different efficiencies?
- b) Do they require additional space?
- c) Do the data need to be in a particular structure? Or does the method work on arrays and lists?
- d) Do they work equally well on large or small data sets?

2. What would you describe as the core idea behind merge sort?

The core idea is splitting your data into a bunch of smaller files/structures and then sorting each of those individually. Once that is done, it is easier to "merge" two pieces that are already independently sorted.

3. What are some good methods for choosing a quicksort pivot, and why?

An easy one is to just take the first element of the file/array. However, if the data are sorted or reverse sorted this is not a good idea. So that is good in that it is easy. However, another method is to get the first, middle, and last data elements in the array and start with the median value. That way, you are guaranteed to have a true partitioning on the first pass.

4. How are subfiles created for shellsort?

If there are K subfiles, then elements that are spaced K spots apart are in the same subfile.

5. In what way does a natural merge exploit existing order in the dataset?

A simple merge sort creates N subfiles to sort a file of length N . However, a natural merge instead takes sequential records that are already sorted and keeps them in the same subfile. This reduces the number of initial subfiles and if there are some neighboring elements that are already sorted this might make it faster than a simple merge sort.

6. Describe in your own words the notion of a stable sort.

A stable sort is one that does not reorder elements that are already ordered. This is particularly important when there is a sort "key", i.e. there is a portion of the data record that is being used for sorting and not the whole thing. If the whole data record is being used for sorting, and there are duplicates, it really doesn't matter if the sort is stable or not.