K-way Merge: Introduction

Let's go over the K-way Merge pattern, its real-world applications, and some problems we can solve with it.



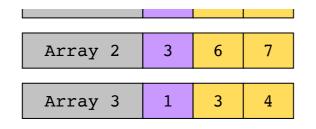
Overview

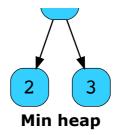
The **k-way merge pattern** helps to solve problems involving a list of sorted arrays.

Here is what the pattern looks like:

- 1. Insert the first element of each array in a min-heap.
- 2. Next, remove the smallest element from the heap and add it to the merged array.
- 3. Keep track of which array each element comes from.
- 4. Then, insert the next element of the same array into the heap.
- 5. Repeat steps 2 to 4 to fill the merged array in sorted order.

The following illustration describes how the pattern works.





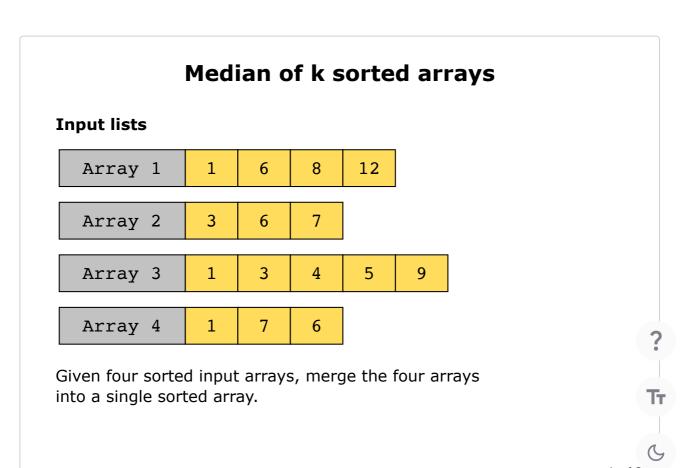
Insert the first element of each array in a Min heap.

1 of 10



Example

The following example illustrates a problem that can be solved with this approach:





Does my problem match this pattern?

- Yes, if both these conditions are fulfilled:
 - The problem involves a set of sorted arrays, or a matrix of sorted rows or sorted columns that need to be merged, either for the final solution, or as an intermediate step.
 - \circ The problem asks us to find the k^{th} smallest or largest element in a set of sorted arrays or linked lists.
- No, if either of these conditions are fulfilled:
 - The input data structures are neither arrays, nor linked lists.
 - The data is not sorted, or it's sorted but not according to the criteria relevant to solving the problem.

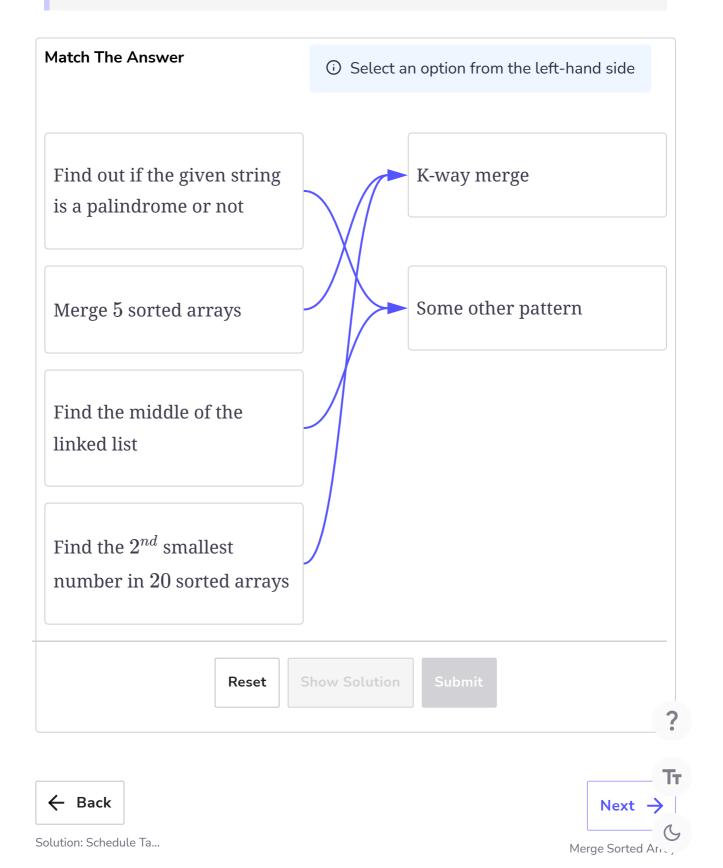
Real-world problems

Many problems in the real world use the k-way merge pattern. Let's look at some examples.

- Merge tweets in twitter feed: Sometimes we need to implement a module that adds a user's Tweets into an already populated Twitter feed in chronological order.
- Used in external sorting procedures: When an algorithm is processing huge amounts of data, it needs to repeatedly fetch it from external storage because RAM capacity is fixed. To overcome the speed limitation of external storage, k-way merges are used in external sorting. Let's consider a case where we need to perform six merges. A binary merge requires three merge passes while a 6-way merge only requires one pass. K-way merge reduces the number of accesses to external storage, which in turn greatly improves performance when dealing with large amounts of data.



Note: Select a problem in the left-hand column by clicking it, and then click one of the two options in the right-hand column.



?

Īτ

