\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Approach 3: Sorting and Merging

In this approach, I sort the lines while traversing the given D string and L vector. Lines that overlap (drawn along the same line) are merged during the process.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Approach 4: Hash Implementation

In this implementation, the longer the horizontal line, the more times we need to find indices in the v\_lines hash map. The actual vertical lines that are found are at least 100x less than the number of lines checked. For example, if a horizontal line in a given iteration has endpoints [-1000, 1000], then, on average, only 20 values among those 2000 possible values (1000 - (-1000)) are found in the v\_lines. This means, for every line, we end up checking 1980 hash values, all of which are O(n).

The larger the length of the lines, the larger is the waste. If we allow the lines to be 1,000,000,000 units long, then, a horizontal line will be, on average, 500,000,000 units long, so we will end up checking 500,000,000 hash values in the hope of finding a vertical line appropriate for the given horizontal line.

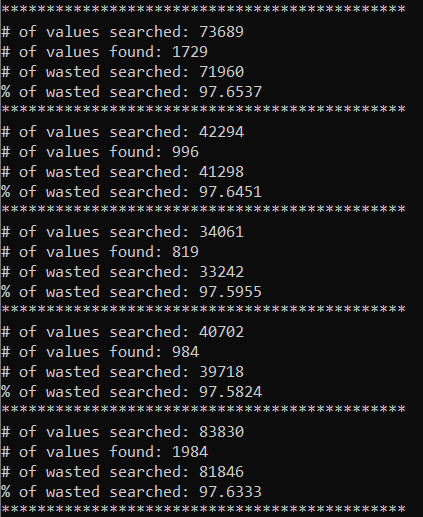
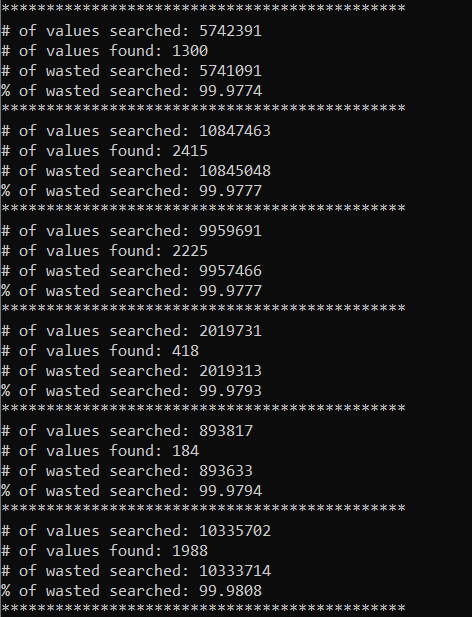
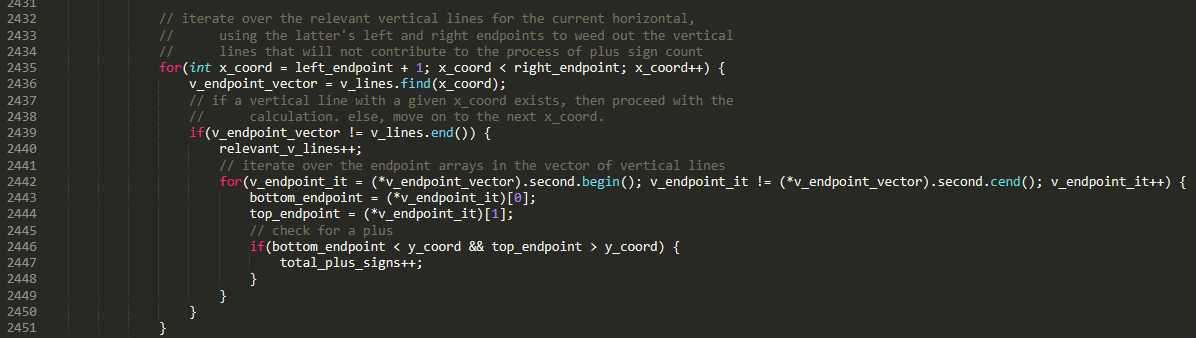
 

Figure : L = 10,000,000 Figure : L = 50,000

As we can see above, most of the time is spent on searching for vertical lines, and 99% of those searches are wasted for bigger values of L.

Below is the code that is responsible for the above behavior.



The larger the gap between the left and right endpoints, the more values are being searched in the v\_lines hash map. However, as seen in Figure 1 and Figure 2, most of these values (x-axis coordinates) do not have a vertical line with that coordinate. To improve the algorithm, we must eliminate all the unnecessary searches of vertical lines. That is, the number of searches must not be based on the length of each horizontal line.