**Lab1 Readme**

1. Write code for insertion sort.

Use typical method to complete insertion sort,you could see the code and annotation in the file named “insertion.cpp”.

1. Write code for merge sort.

Use typical method to complete merge sort,you could see the code and annotation in the file named “merge.cpp”.

3. The running time of merge sort can be improved in practice by taking advantage of the fast running time of insertion sort when its input is “nearly” sorted. When merge sort is called on a subarray with fewer than k elements, use insertion sort to sort the subarray. Argue that this sorting algorithm runs in O(f(n, k)) expected time. What is f(n, k) and how should k be picked, both in theory and in practice by experiments?

In theory,O(f(n,k)) depends on the k,because it is the k which decides when to transform from merge sort to insertion sort.In fact,maybe there is not any definite k values to lead the minnimum sorting time due to the factors like data distribution,hardware performance and so on.

In practice,we set the size of dataset to be 10w,which means we generate 10w numbers between 0-100000 randomly once,using it to test our hybrid merge sort.

First,when we set the k to be for(int k = 2,k <= 100000,k \*= 2),the time is decreasing and then increasing.Among these test data, we find that the time are both more than 1 second when k is less than 16 and more than 8192,which is too long.

So next,we set the k to be for(int k = 16,k <= 8192,k \*= 2),and for each k we get the test time accordingly,then we use these datas drawing a diagram,which is in the data.xlsx(We did it twice,which are Scond round1 and Second round2).Considering of the diagram,it is obvious that the ideal k should between 256-600.

So in the third time,we set the k to be for(int k = 256; k <= 600; k += 20),getting the datas,drawing a diagram,which is in data.xlsx(We did it three times,which are Third round1,Third round2 and Third round3).We can see that the k values are not the same each time due to the factors like data distribution,hardware performance and so on,but mainly distributing during 400-550.

In the last,we can say the ideal k value in practice is **462** coarsely.

4.Write code for improved version of sorting algorithm which combines merge sort with insertion sort.

Combine insertion sort and merge sort to complete the improved merge sort,you could see the code and annotation in the file named “merge1.cpp”.The main idear of it is that for the main body we use the merge sort solution,for the particularly small subarrays,especially when they are nearly sorted.we use insertion sort.Therefore we combine the advantages of both insertion sort and merge sort,reducing the running time effectively.