Name: Zixuan Zhao
StuID: 932 282 628

Proposal

In my project, I plan to test a sorting library based on Skip List data structure. This library can be found in Github: https://github.com/sepeth/python-skiplist.

Skip List is a data structure for searching elements faster based on an ordered sequence of elements. Generally, Skip List is a Linked List built in several layers with basic element node. The number of layers is determined randomly. This library mainly contains two sorting function: sorting dictionary and sorting set.

Generally, I will test this library from three parts:

1. Skip List Correctness
   There are several methods providing some general operations of Skip List in this library, which is creating, inserting, removing, updating, and so on. To examine the correctness of Skip List, I will test by these aspects:
   a. Data Structure Correctness
      First, I will examine if the library can create a correct order Skip List for a given data. It means the data will be in a specific order by different random numbers. Then, I will examine if the library can insert, delete, and update data correctly under circumstances. For example, check if I can insert a data with value None, delete a node from empty Skip List and update a node that does not exist in the Skip List correctly. This will examine the basic operations and reduce the bug caused by these operations for the further functions.
   b. Different Data Set
      Data set can be various on these aspects, the number of data size, the type of data, the value of data. For example, I will examine whether this library can support to build up a Skip List with a huge size of data, and how large the data set can be supported. Also, I will input a data set with the same value. This can examine how this library deal with different data set.
   c. The Random Testing
      Since operations in Skip List is randomly. I can control the randomizations in the Skip List to see what happened. In this case, I will give specific value to the random variables. Also, I will examine the affections with various range of random numbers.
2. Sorting Set Correctness
   Set is a data list which does not have order and repeated elements. To test this function in this library, I will examine one aspect, which is whether function can sort distinct input data set correctly. As I mentioned above, there are multiple ways to generate special data set. In this case, I will mainly consider the data set with several repeated values, the data set with uncommon symbols such as “\t”, “⌐”, or the data set with several types of data.
3. Sorting Dictionary Correctness
   To test the correctness of this function, except the contents mentioned above, I should concern on examining whether this function correctly sort the key instead of the value.

What’s more, since Skip List increase the performance of searching data. In this case, I would like to also test the performance to see if this library performs really better than the built-in sort function. These are my plan of project.