Program Structure and Algorithm

Assignment 4 Hash Problem

Shouting Lyu 001822094

To solve two problems in the assignment, I implemented a solution in the Birthday.java file which will read in an integer and print out the results of the experiment. In order to operate the experiment scientifically, the program will run 100 times and calculate the average value of birthday and card-collecting problems.

The following figures will demonstrate both results of the experiment and the theory relationship between the array size and the times of hashing operation.

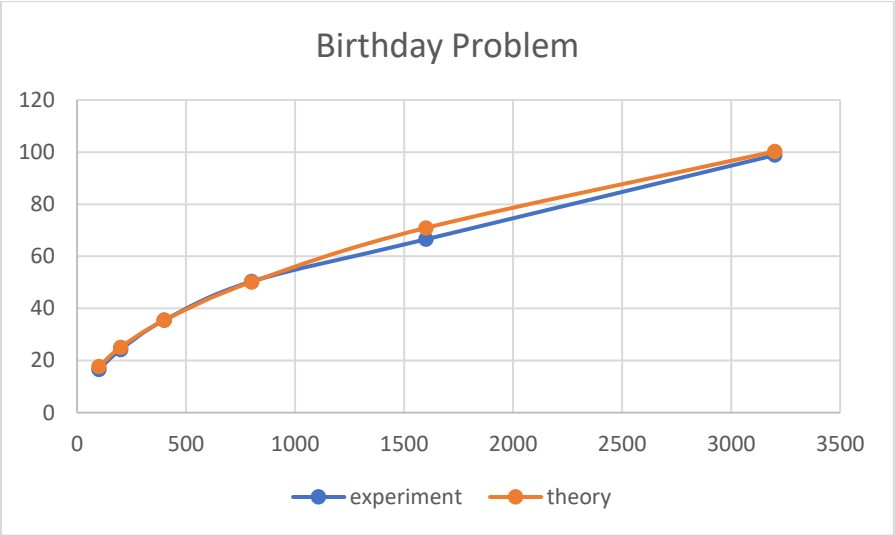

Figure 1

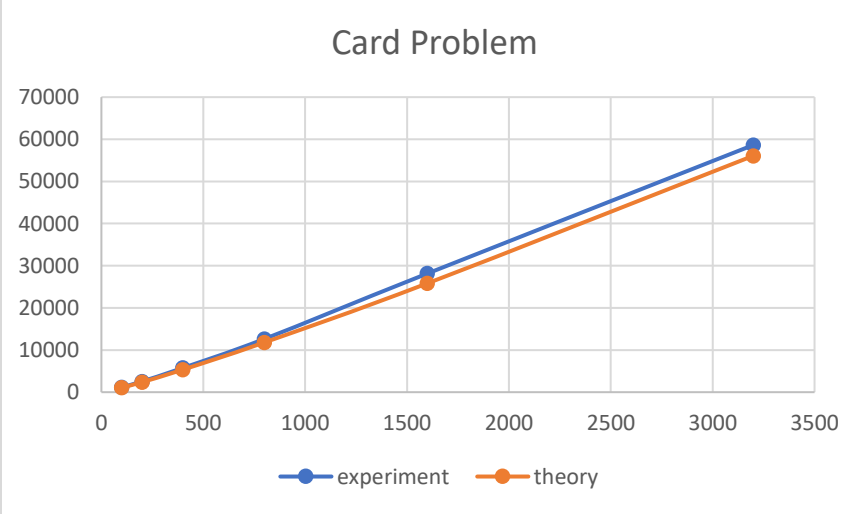

Figure 2

In the Figure 1, the orange curve represents the theory relationships which is $\sqrt{\frac{\pi M}{2}}$ and the blue curve was generated by the experiment data. Obviously, the experiment data fit the curve very well. Similarly, in the Figure 2, the curve of the experiment fit the theory relationship which is $\mathrm{M} \ln M$. Besides, the real data is slightly larger than the theory's because the theory implies the lower limit.

In addition, when I first implemented the hashing code, I made a mistake. Before the experiment, I generated a random number as a key and put it into the hash function, however, the range of the random number is just 10 times of the array size M, and the collision happens very often which is much more than the theoretical value. After I changed the range from 0 and 10*M to 0 and Integer.MAX_VALUE, everything goes well.