Explanation Task 2



In task 2, we were tasked to test how long it takes for the program to read a text file and place it in the hash table. Due to the variation of base and table sizes, the possibility of collision might occur hence, collision in task 2 was handled using linear probing. With a time limit of 2 minutes (120 seconds), different bases with different table size was tested against each other and was compared with each other.

According to the data of the first graph, when the base of the hash table is too small compared to the table size, the time taken for the program to insert all words into the hash table takes longer than 2 minutes. It occurs for all text file, english\_small.txt, english\_large.txt and french.txt. When the base is too small, the difference between the hash value will be too small, hence there will be a higher chance for collision to occur. Hence, the keys that are hashed are not random enough to help the hash function to perform optimally.

When the value of base of the hash table and the table size is too close to each other, the time taken for the program to insert all the words into the hash table also takes longer than 2 minutes. When the value of the base and the table size are too similar, the hash value obtained may be too big, hence there will be a higher chance for collision to occur. Hence, the keys that are hashed are not random enough to help the hash function to perform optimally.

In order for the hash table to perform faster and better, the value difference between the base and the table size must be bigger. According to the clearer graph above, time taken for the program to read the file is below 2 minutes when the difference between the base and table size is not too big.