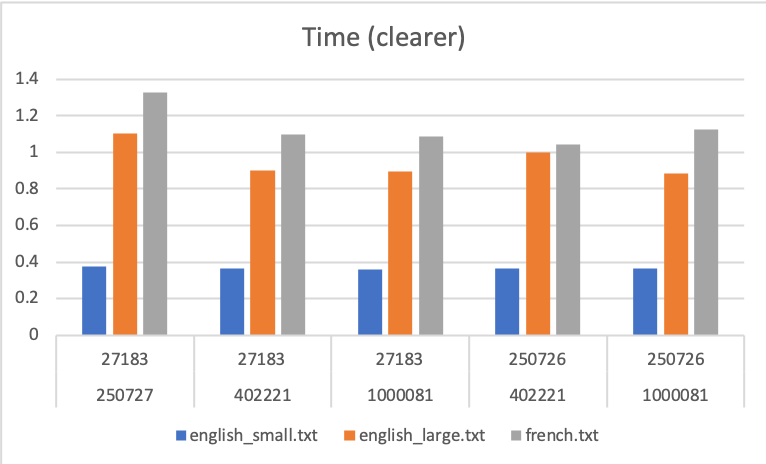
Explanation task 3



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 using linear probing, 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 using linear probing, the keys that are hashed are not random enough to help the hash function to perform optimally.

According to the graph, collision occurs less frequently when there is a big value difference and a small value difference between the base and table size. This is because collision is only calculated when the program enters the set item function again, not when it repeats the for loop. More of time is spent probing the key through the hash table compared to it entering the function again with a new key that only probes once. Hence using linear probing, collision occurs more frequently when there is a value difference between base and tables size as the chances of a key finding a new spot in the hash table after colliding is higher.

According to the graph, the length of the probe is longer when there is a big value difference and a small value difference between the base and table size. This is because once the keys collide in the hash table, it is required to probe again and again because when the value of the base is too small, the hash values are too close to each other, hence longer length of probing is required. When the value of the base and the table size is too close, the hash values are too big, and so all the keys will be inserted at the similar large value hence longer length of probing is also required.

The length of the probe is shorter when the value difference between the base and table size is not too small and not too big because the hash value of the keys are more random, hence through linear probing, they are not required a longer length of probing.

According to the graph, through linear probing, the max of the probe is bigger when there is a big value difference and a small value difference between the base and table size. This is because once the keys collide in the hash table, it is required to probe again and again because when the value of the base is too small, the hash values are too close to each other, hence longer length of probing is required. When the value of the base and the table size is too close, the hash values are too big, and so all the keys will be inserted at the similar large value hence longer length of probing is also required. When there is a long length of probing the chances of obtaining a large value of max probe is higher.

The length of the probe is shorter when the value difference between the base and table size is not too small and not too big because the hash value of the keys are more random, hence they are not required a longer length of probing. When there is a short length of probing the chances of obtaining a small value of max probe is higher.