

Problem 1

A

```
[y336@wind ~/inf503/Homework5]$ ./main problem1A /common/contrib/classroom/inf503/SARS_COV2.txt 1000 0
The number of arguments passed: 5
The first argument is: ./main
The second argument is: problem1A
The third argument is: /common/contrib/classroom/inf503/SARS_COV2.txt
The forth argument is: 1000
The fifth argument is: 0
found 1000 in the sequence
Total Nodes: 31754
[y336@wind ~/inf503/Homework5]$ ./main problem1A /common/contrib/classroom/inf503/SARS_COV2.txt 50000 0
The number of arguments passed: 5
The first argument is: ./main
The second argument is: problem1A
The third argument is: /common/contrib/classroom/inf503/SARS_COV2.txt
The forth argument is: 50000
The fifth argument is: 0
found 29868 in the sequence
Total Nodes: 867321
[y336@wind ~/inf503/Homework5]$ ./main problem1A /common/contrib/classroom/inf503/SARS_COV2.txt 100000 0
The number of arguments passed: 5
The first argument is: ./main
The second argument is: problem1A
The third argument is: /common/contrib/classroom/inf503/SARS_COV2.txt
The forth argument is: 100000
The fifth argument is: 0
found 29868 in the sequence
Total Nodes: 867321
[y336@wind ~/inf503/Homework5]$
```

The size of the tree depends on the number of 36-mers. The results for 50K, and 100K random 36-mers are the same, because of the total length of SARS-CoV2 genome is limited.

The search result for 1000 random 36-mers is 1000, because all the 36-mers are generated from the genome.

The search results for 50K and 100K also make sense. 29868 is the total number of 36-mers we got from SARS-CoV2 genome.

B

```
Total Nodes: 867321
[y336@wind ~/inf503/Homework5]$ ./main problem1B /common/contrib/classroom/inf503/SARS_COV2.txt 1000 5
The number of arguments passed: 5
The first argument is: ./main
The second argument is: problem1B
The third argument is: /common/contrib/classroom/inf503/SARS_COV2.txt
The forth argument is: 1000
The fifth argument is: 5
found 293 in the sequence
Total Nodes: 31795
[y336@wind ~/inf503/Homework5]$ ./main problem1B /common/contrib/classroom/inf503/SARS_COV2.txt 50000 5
The number of arguments passed: 5
The first argument is: ./main
The second argument is: problem1B
The third argument is: /common/contrib/classroom/inf503/SARS_COV2.txt
The forth argument is: 50000
The fifth argument is: 5
found 13889 in the sequence
Total Nodes: 878272
[y336@wind ~/inf503/Homework5]$ ./main problem1B /common/contrib/classroom/inf503/SARS_COV2.txt 100000 5
The number of arguments passed: 5
The first argument is: ./main
The second argument is: problem1B
The third argument is: /common/contrib/classroom/inf503/SARS_COV2.txt
The forth argument is: 100000
The fifth argument is: 5
found 14512 in the sequence
Total Nodes: 897820
[y336@wind ~/inf503/Homework5]$
```

The size of the tree depends on the number of 36-mers. The results for 50K, and 100K random 36-mers are the same, because of the total length of SARS-CoV2 genome is limited, and they use the same seed to generate random numbers.

The search results also make sense. The more 36-mers, the more chances to find a match. Besides, because of the limited bases in the genome, the results for 50K and 100K are similar.