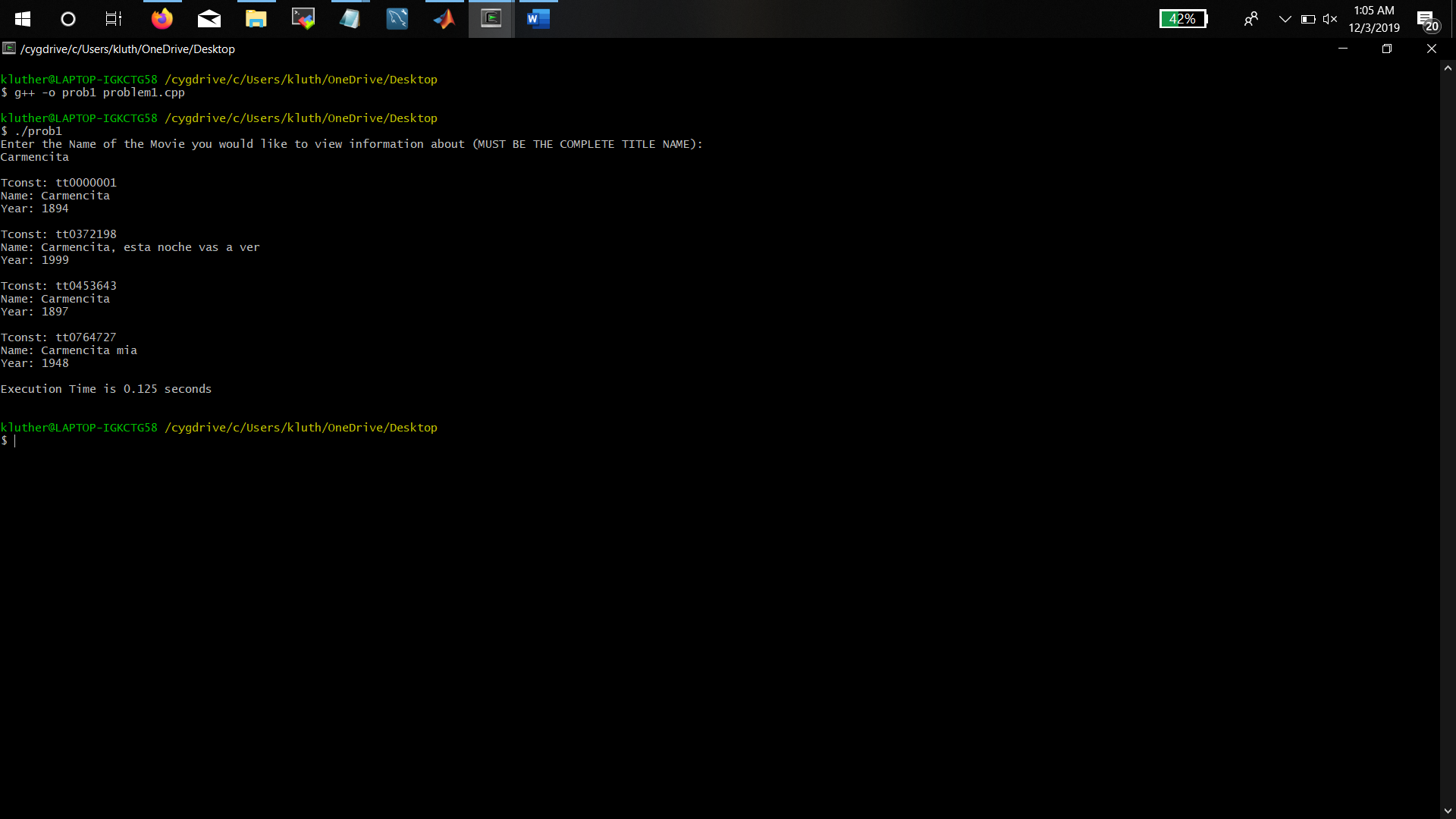
1. **Indexing - Attached data file ”out.dat” is a file which store a table of movies information. Its schema is described as follows:**

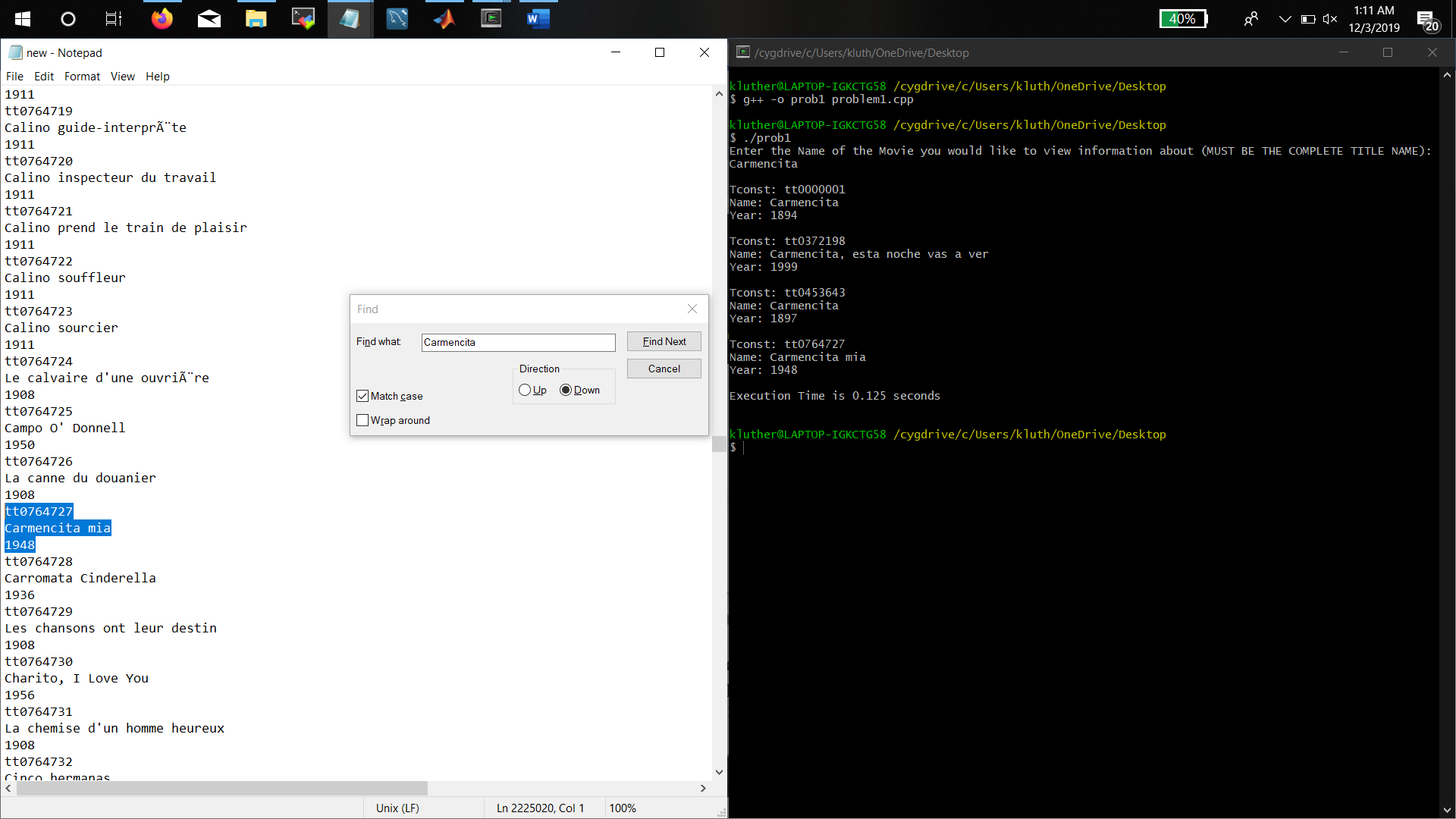
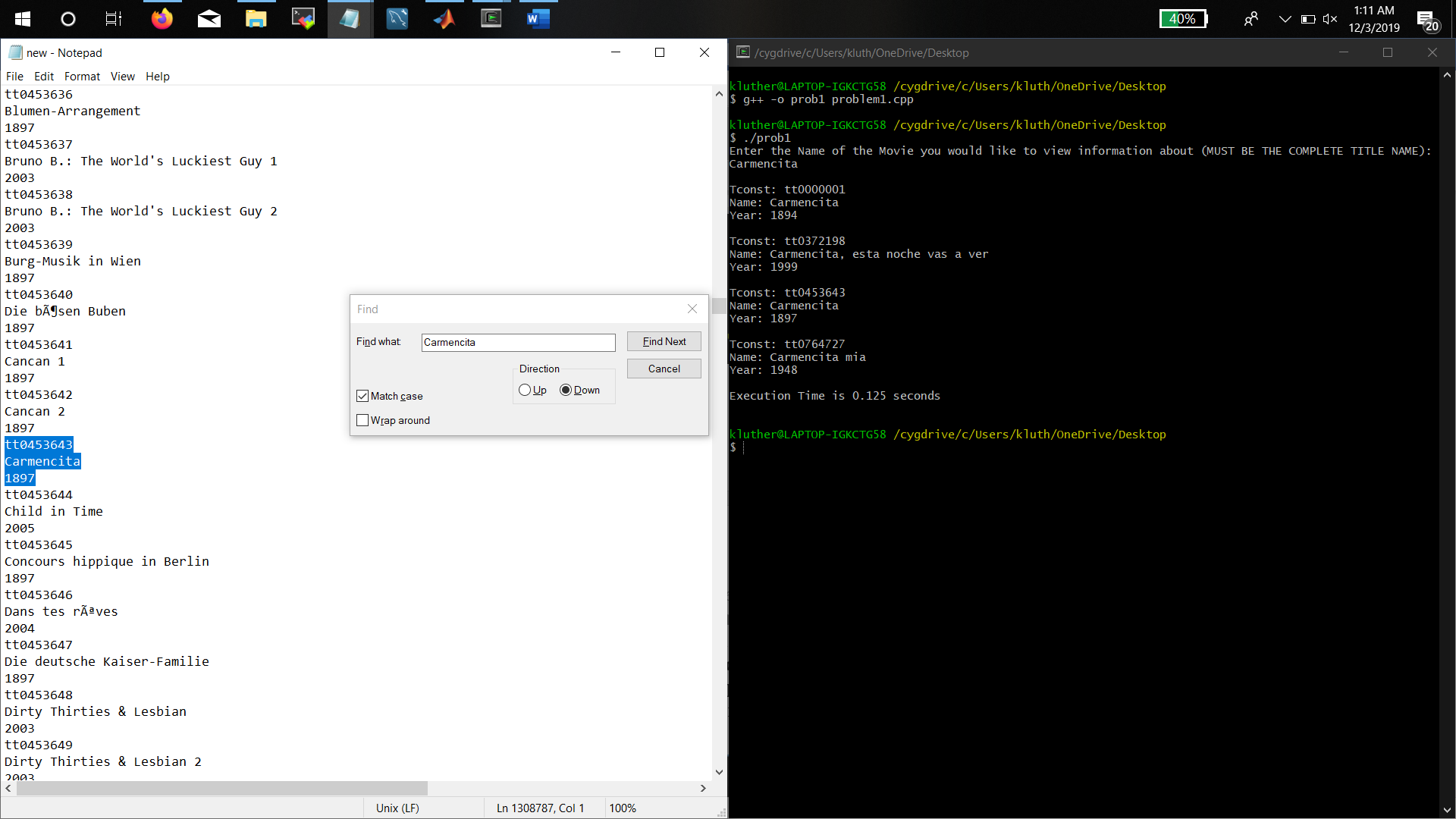
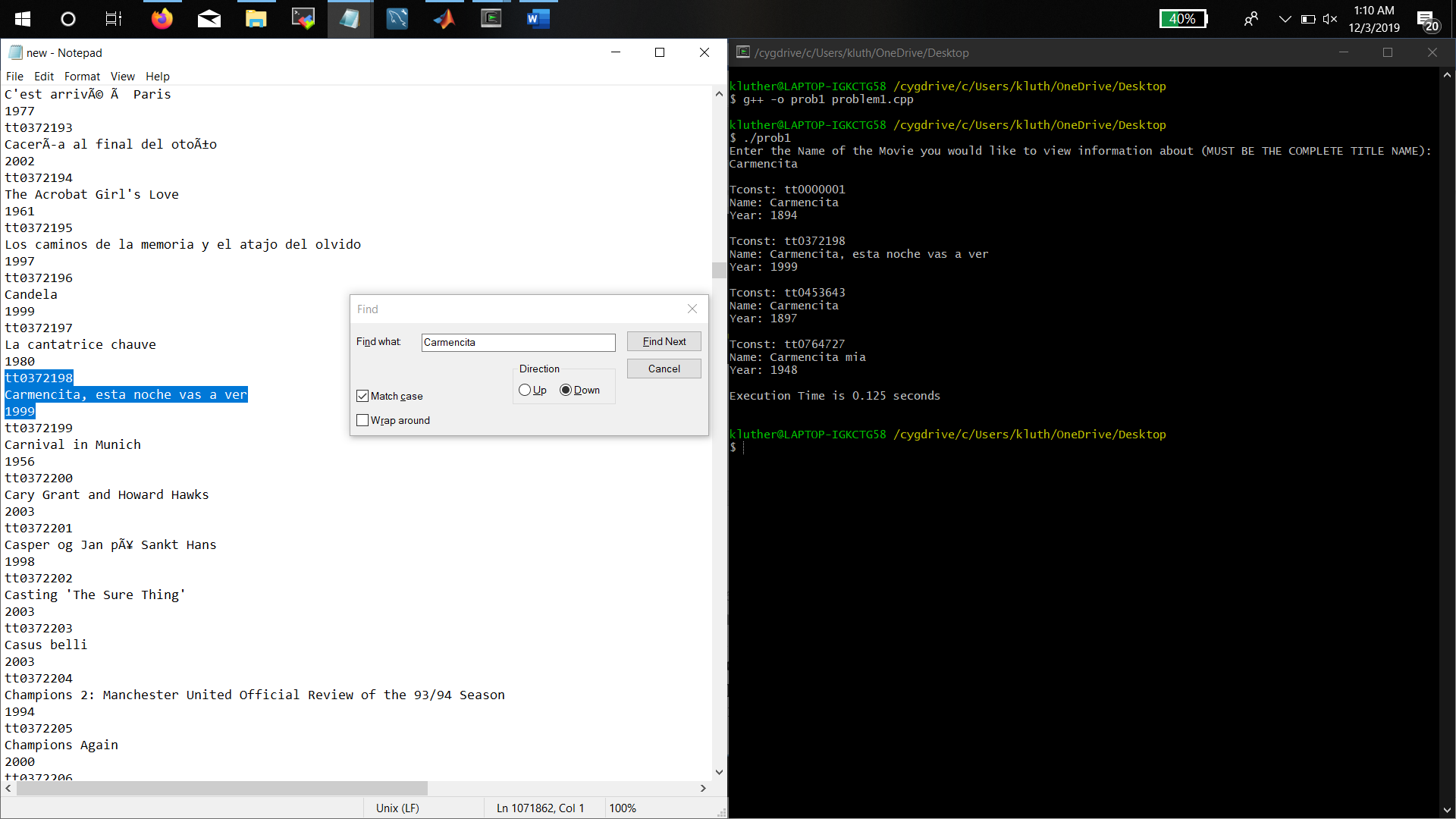
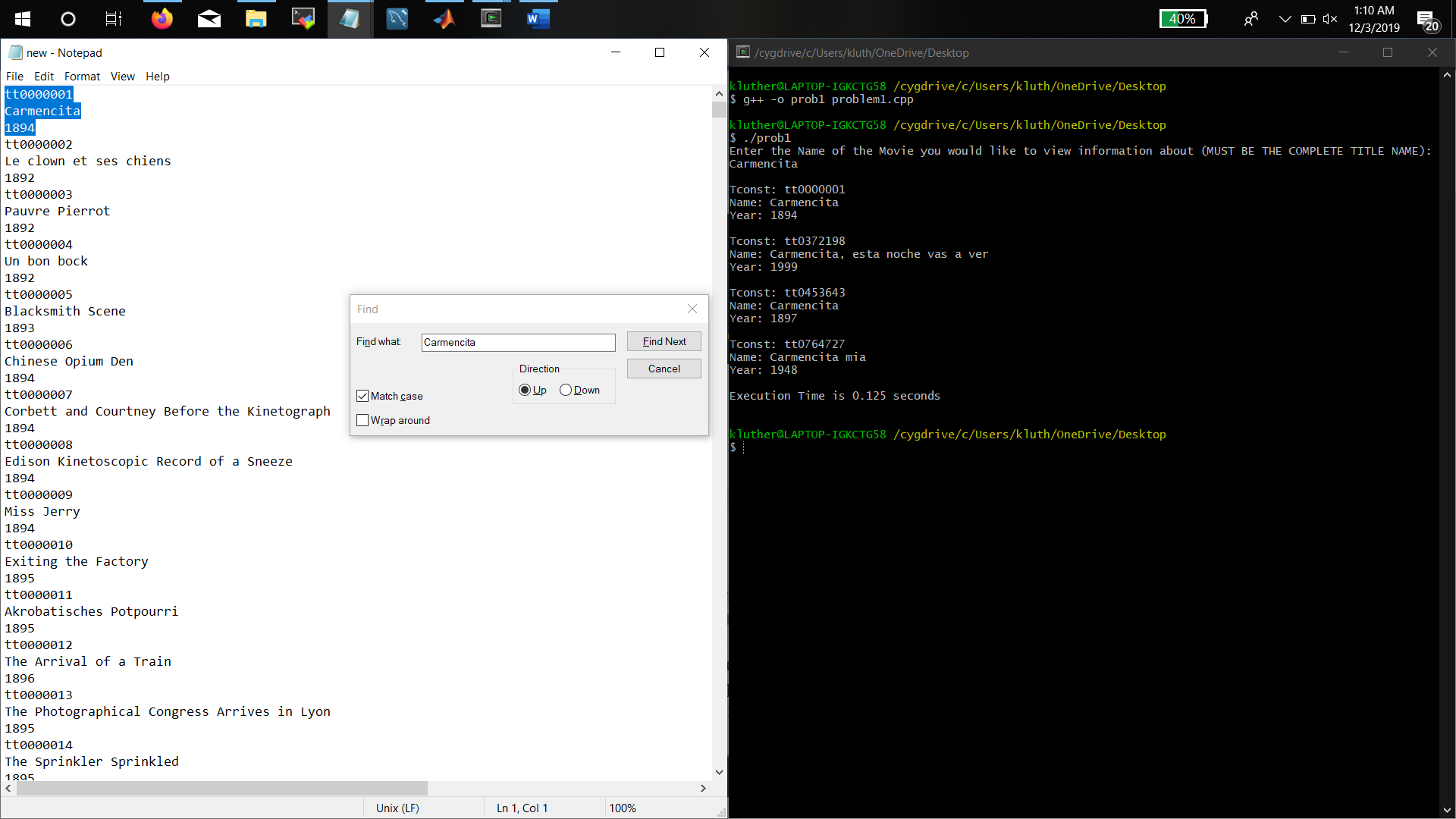
|  |  |  |
| --- | --- | --- |
| tconst | Title ID | char(12) |
| name | Movie name | char(255) |
| year | Release year | int |

|  |
| --- |
| typedef struct t i t l e {  char tconst [1 2] ;  char name[255];  int year ;  } Title ; |

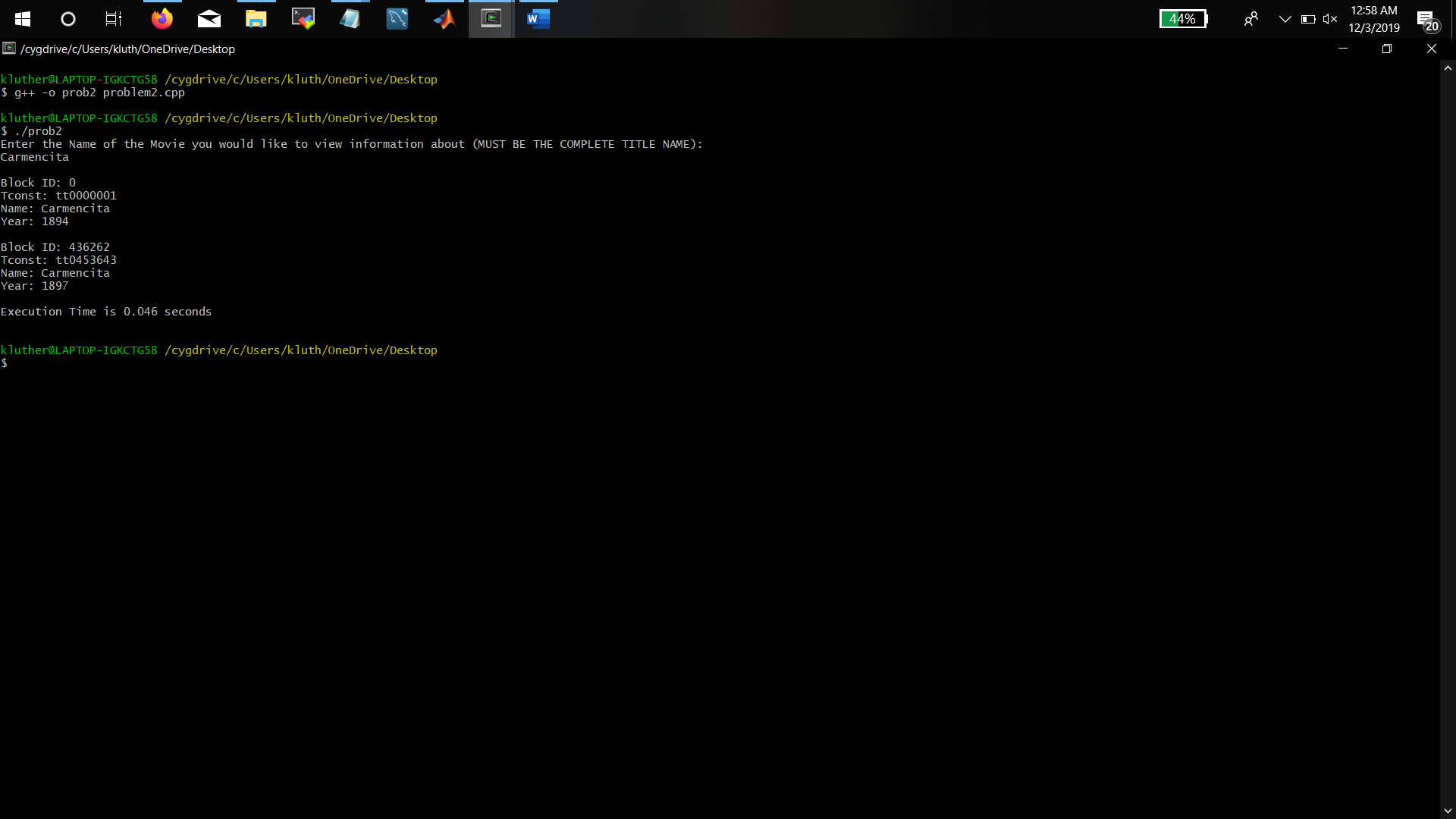
Each row is stored in a block has 12+255+4 bytes of data. You can use the following C structure to present one block/row of data:

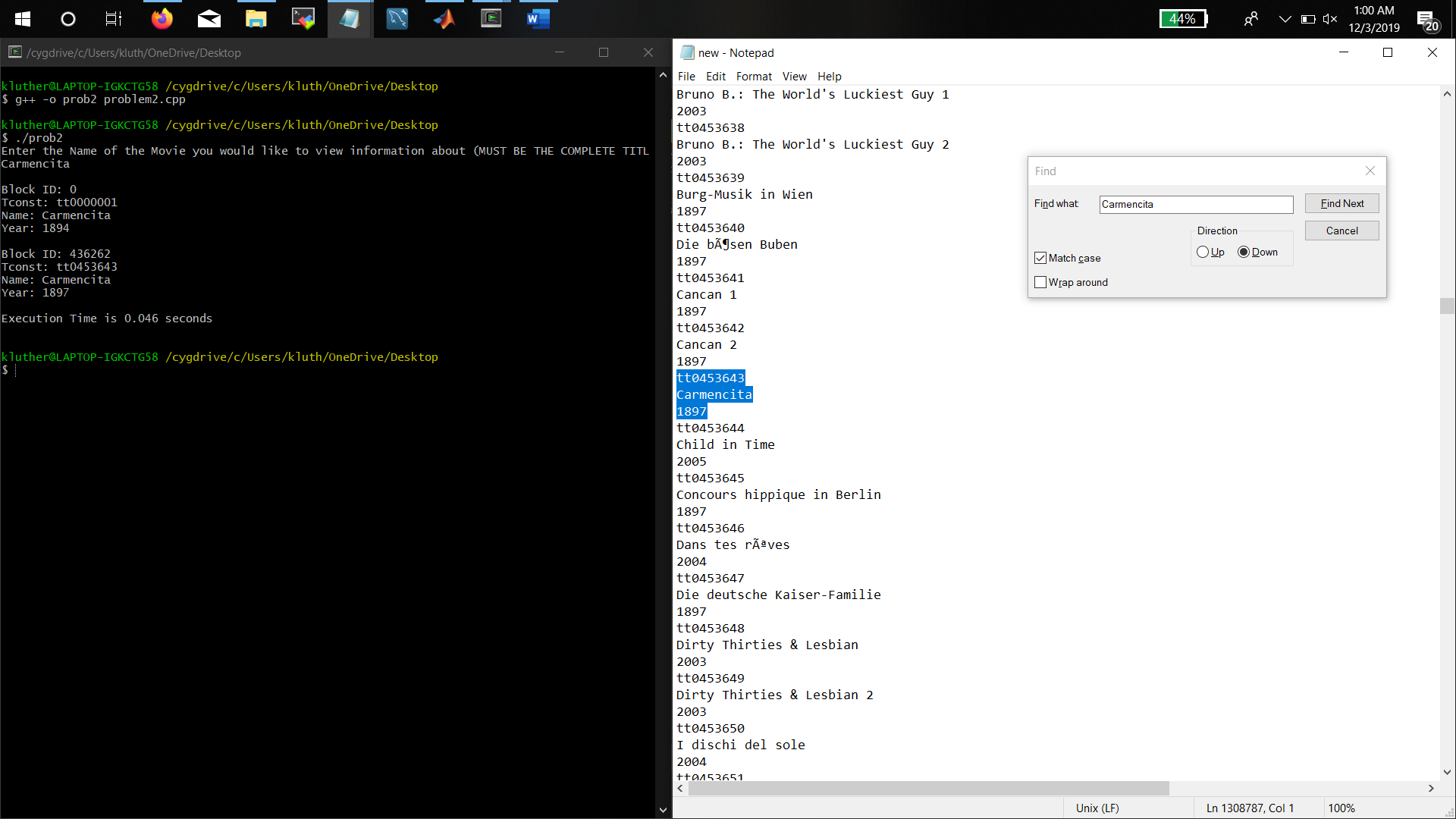
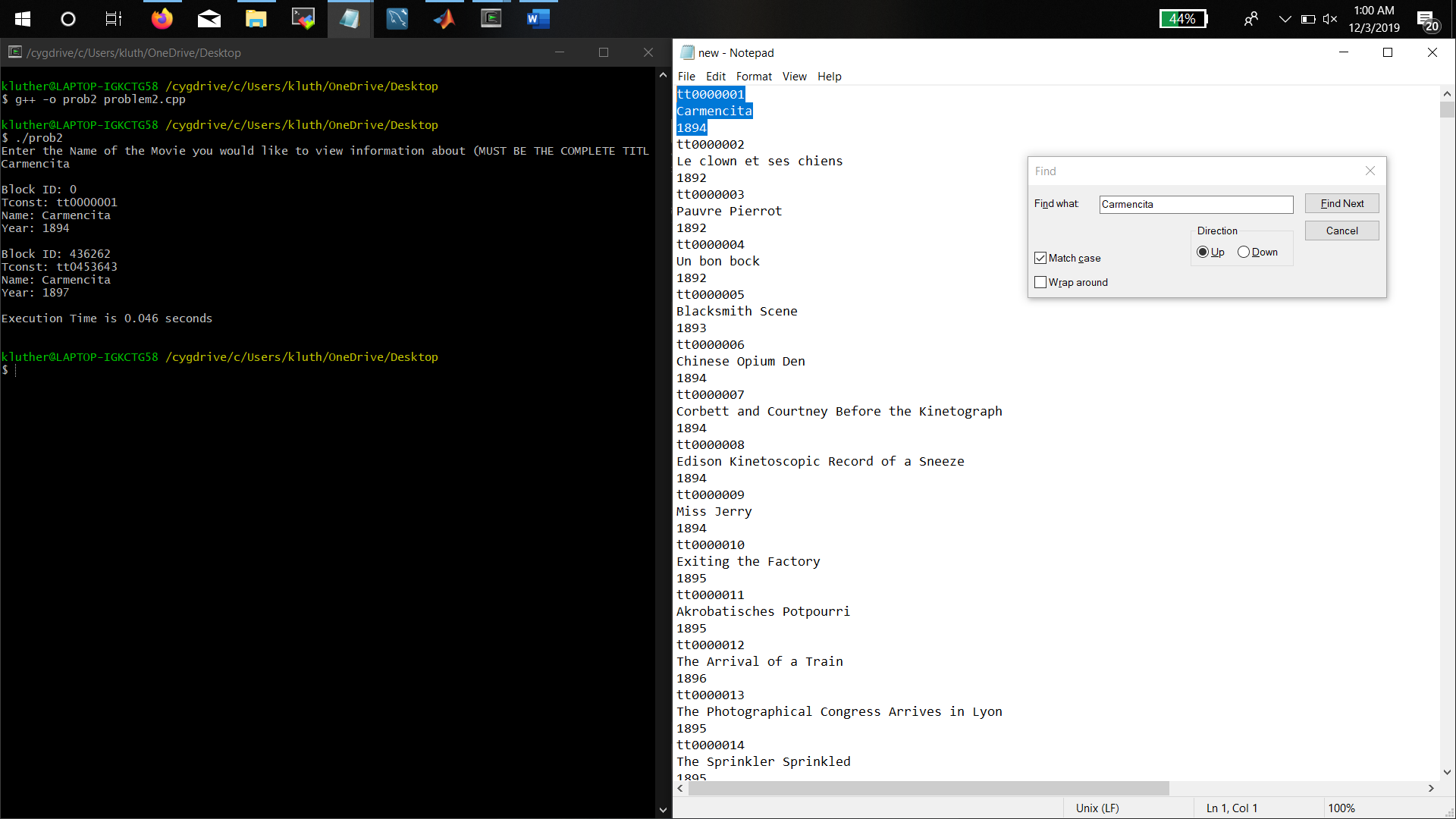
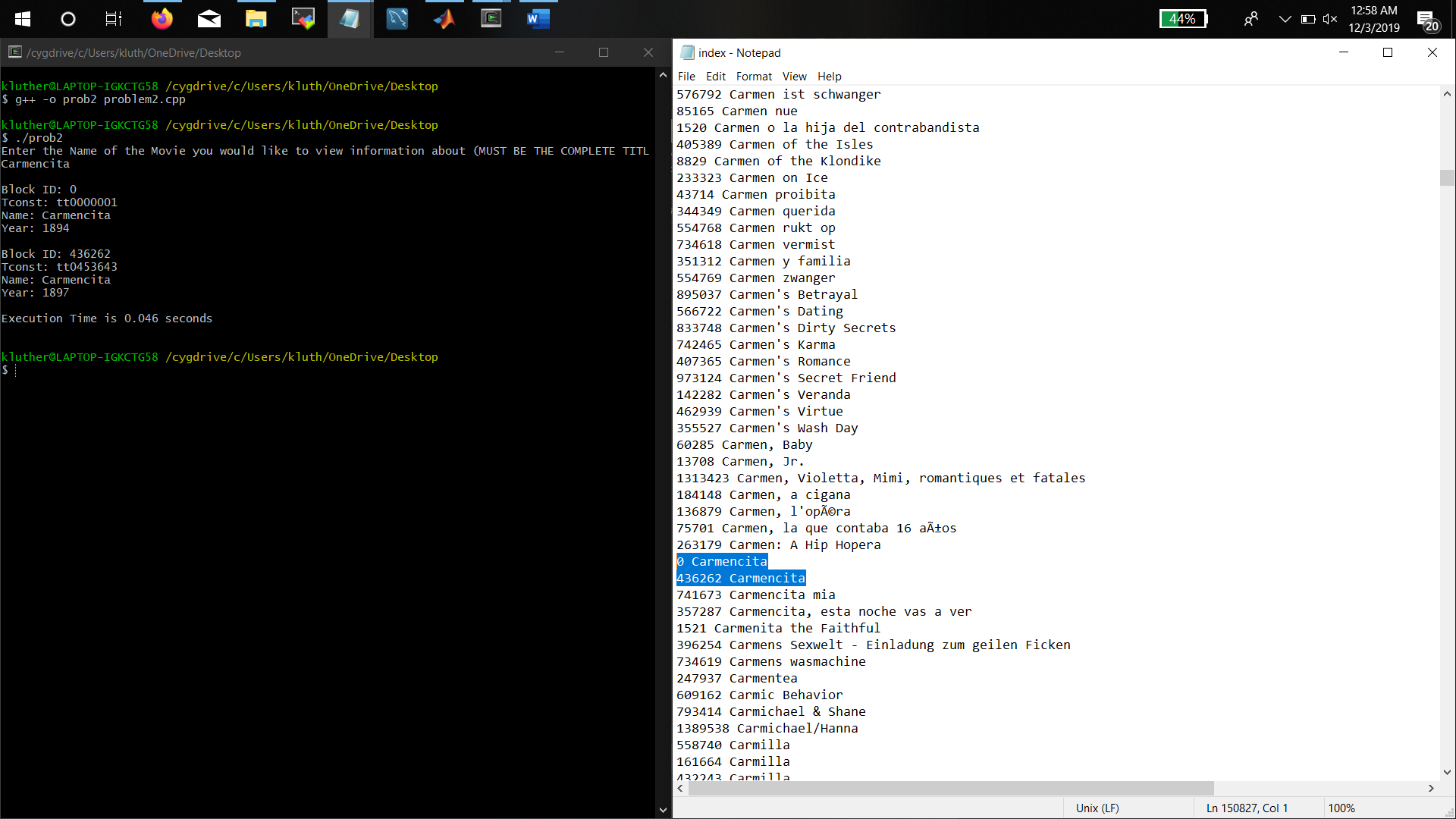
* 1. **Write a program to find information of movies by searching movie’s name with a keyword. You should mark down the execution time. (40 pts)**





**1.2 Write a program to create a index file for movie’s name. Your index file should contain a sorted index list with movie’s name and block id. Then use this index file to perform the search described in the previous question again and mark down the execution time. (40 pts)**





**1.3 Calculate the speedup after using index. (10 pts)**

**1.4 Bonus 20 pts. Instead of using a sorted index file, use a hash index file to complete the same task and compare the performance.**

2