

```

1  /*****
2  *
3  *****/
4
5      phraseFilterL <input.txt> <output.txt> <lim>"
6
7  Reads a text file, puts all phrases (between 2 consecutive full stops) in the rows of a
8  2D array. If a phrase contains more words than given by 'lim', the phrase is copied in a
9  text file. If no phrase has more than 'lim' words a message is printed.
10 */
11
12
13 # include <stdio.h>
14 # include <string.h>
15 # include <stdlib.h>
16 # include <ctype.h>
17
18 # define MAX 1000
19
20 int FtoM (FILE *, char [][][MAX]);
21 int words (char []);
22
23
24 int main (int argc, char * argv[]) {
25     if (argc != 4) {
26         printf("\nInvalid Command!\n");
27         return 0;
28     }
29
30     if (*(argv + 3)[0] == '-') {
31         printf("\nError: negative number not allowed!\n");
32         return 0;
33     }
34
35     unsigned short lim;
36     lim = atoi (*(argv + 3));
37
38     FILE *fin, *fout;
39     char m[MAX][MAX];
40     int rows, numWords, numPhrases=0;
41
42     fin = fopen (*(argv + 1), "r");
43     fout = fopen (*(argv + 2), "w");
44
45     rows = FtoM (fin, m);
46
47     fprintf(fout, "Phrases with more than %d words:\n\n", lim);
48
49     for (int i = 0; i < rows; i++) {
50         if ((numWords = words(m[i])) > lim) {
51             fprintf(fout, "Phrase n. %d:\n%s (%d words) \n\n", (i + 1), m[i], numWords);
52             numPhrases++;
53         }
54     }
55
56     if (numPhrases == 0)
57         fprintf (fout, "None.\n\n");
58

```

```
59     printf ("\nRead results in %s\n\n", *(argv + 2));
60
61     fclose(fin);
62     fclose(fout);
63
64     return 0;
65 }
66
67
68
69
70 /*****
71  *
72  *****/
73
74 Counts the number of words in a string containing a phrase.
75
76 It performs the task by counting the number of empty spaces or '\', '/', and apostrophes
77 between words; this counting incremented by 1 is the number of words.
78 In case of consecutive empty spaces (' ', '\t'), only 1 is counted.
79
80 The enumeration 'st' is used to detect whether we're inside a word (in) or not (out).
81 The count of word separators is incremented when one is read when st=in. Word separators
82 read when st=out are ignored.
83 If st=out and alphanumerical characters are read, 'st' is set to 'in'.
84 If st=in and alphanumerical characters are read, they are ignored.
85
86 NB: If there are multiple empty spaces before the end of the string, 'st' will be equal to
87     'out' and 1 more space would be counted. So there's some correction at the end.
88
89 NB: It's assumed that there is no empty space at the beginning of the string.
90
91 NB: '-' and '_' are assumed to be used only to merge 2 words together into 1 word.
92 */
93
94
95
96 int words (char s[]) {
97     unsigned short separators = 0;
98     enum st {out, in} st;
99
100     st = in;
101
102     for (int i = 0; s[i] != '\0'; i++) {
103         if ((isspace(s[i]) || s[i] == '\'' || s[i] == '\\' || s[i] == '/')
104             && st == in) {
105             separators++;
106             st = out;
107         }
108         else if ((isspace(s[i]) || s[i] == '\'' || s[i] == '\\' || s[i] == '/')
109                 && st == out)
110             continue;
111         else if (isalnum(s[i]) && st == out)
112             st = in;
113         /* Alphanumerical characters or punctuation AND st=in*/
114         else
115             continue;
116     }
```

```

117
118     if (st == in)
119         return separators + 1;
120     else
121         return separator;
122
123 }
124
125
126
127
128
129 /*****
130  *                               FtoM
131  *****/
132
133 Receives a pointer to a file and a 2D array of char.
134 Copies each sentence in a row of the array and returns the number of rows that have been
135 filled, which is the number of sentences.
136 It ignores empty spaces at the beginning.
137
138 To determine the end of a sentence it checks that an empty space, a new line or a
139 tabulation character follows a full stop character (the latter is copied too). A new line
140 character in the text is translated into an empty space in the array. If other characters
141 follow a full stop character, it continues to copy in the same sentence.
142
143 BEWARE: other types of mistakes (e.g. punctuation characters after '.') are not checked.
144 */
145
146 int FtoM (FILE *f, char m[][max]) {
147     int c, i=0, j=0;
148
149     while ((c = getc(f)) != EOF && j < max) {
150         if (isspace(c) && j == 0)
151             continue;
152         else if (c != '.' && c != '\n')
153             m[i][j++] = c;
154         else if (c == '.') {
155             m[i][j++] = c;
156
157             if ( isspace( (c = getc(f)) ) ) {
158                 m[i++][j] = '\0';
159                 j = 0;
160             }
161             else if (c == EOF) {
162                 m[i][j] = '\0';
163                 break;
164             }
165             else
166                 m[i][j++] = c;
167         }
168         else
169             m[i][j++] = ' ';
170     }
171     return i + 1;
172 }

```