

```
1  /*****
2  *
3  * countWordsN
4  *
5  * contaParole <input.txt> <output.txt> <n>
6
7  Reads a text file and prints on another text file the number of words of at least
8  n characters (n must be greater or equal to 1).
9  */
10
11
12 # include <stdio.h>
13 # include <stdlib.h>
14 # include <ctype.h>
15
16
17 unsigned readerOne (FILE *);
18 int reader (FILE *, int);
19
20
21 int main (int argc, char *argv[]) {
22     /* check command format */
23     if (argc != 4){
24         printf ("Invalid Command!\n");
25         return 0;
26     }
27
28     if (*(argv + 3)[0] == '-') {
29         printf("\nERROR.\n");
30         return 0;
31     }
32
33     unsigned short n;
34     n = atoi(*(argv + 3));
35
36     FILE *fin;
37     FILE *fout;
38
39     fin = fopen(*(argv + 1), "r");
40     fout = fopen(*(argv + 2), "w");
41
42     /* Two functions are used for the cases n=1 and n>1 because for n=1 the
43     algorithm is simpler */
44     if (n == 1)
45         fprintf(fout, "N. of words at least 1 character long: %d", readerOne(fin));
46     else
47         fprintf(fout, "N. of words at least %d character long: %d", n, reader(fin, n));
48
49     fclose (fin);
50     fclose (fout);
51
52     printf("\nRead results on %s\n", (*(argv + 2)));
53
54 }
55
56
57
58
```

```
59 /*****
60 *                               readerOne
61 *****/
62
63 Counts the number of words with at least 1 character. The function has direct
64 access to the input text file.
65
66 I use 'st' to detect whether I'm reading a word (in) or not (out).
67
68 Empty spaces and punctuation marks are ignored when st=out.
69 When st=out and alphanumerical characters are read, 'st' is set to 'in'.
70 'st' remains in the same state as long as alphanumerical characters come.
71 When st=in and an empty space or an apostrophe are read, a word is counted and
72 'st' is set to 'out'.
73
74 Some care is needed for the last word: if the last word is immediately followed
75 by EOF 'st' remains 'in' and the last word wouldn't be counted. So a check at
76 the end makes a correction in that case.
77
78 NB: Special characters are not considered.
79 NB: Wrong characters at the beginning or at the end of a word are not treated.
80 */
81
82 unsigned readerOne (FILE *f) {
83     unsigned words = 0, c;
84     enum st{out, in} st;
85
86     st = out;
87
88     while ((c = getc(f)) != EOF) {
89         if ( st == out && (isspace(c) || ispunct(c)) )
90             continue;
91         else if ( st == in && ( isspace(c) || ispunct(c)) ) {
92             st = out;
93             words++;
94         }
95         else if ( st == out && !( isspace(c) || ispunct(c)) )
96             st = in;
97         /* Nothing to do in other cases */
98         else
99             continue;
100     }
101
102     if (st == out)
103         return words;
104     else
105         return (words + 1);
106 }
107
108
109
110
111
112
113
114
115
116
```

```
117
118 /******
119 *                                     Reader
120 *****/
121 /*
122 Counts words with at least n characters, reading from the input text file
123
124 Punctuation characters are ignored.
125 If it reads alphanumerical characters, it increments 'letters'.
126 Empty spaces at the beginning or end of file (letters=0) are ignored.
127 When empty spaces or punctuation marks are read immediately after the end of a
128 word (letters > n-1), 'words' is incremented and 'i' is reset to 0. If the same
129 happens when 0 < letters < n-1, the word read is too short and is not counted
130 but 'letters' is reset to 0.
131 As for the other function, the last word requires some care.
132 */
133
134
135 int reader (FILE * f, int n) {
136     int c;
137     unsigned short letters = 0;
138     unsigned short words = 0;
139     while ((c = getc(f)) != EOF) {
140         if ((isspace(c) || ispunct(c)) && letters == 0)
141             continue;
142         else if ((isspace(c) || ispunct(c)) && letters > (n - 1)) {
143             letters = 0;
144             words++;
145         }
146         else if ((isspace(c) || ispunct(c)) && letters > 0 && letters <= (n - 1))
147             letters = 0;
148         else if ( (c != '\\') && ispunct(c) )
149             continue;
150         else
151             letters++;
152     }
153
154     if (letters > n - 1)
155         words++;
156
157     return words;
158 }
```