

USER AWK FUNCTIONS LIBRARY

Boris Bocharov, Maria Voevodina, Anton Yeryomin

We got the information from the Moodle database and proceed to its processing, using, as always, a universal program for working with text files **awk** [1-7].

During the implementation of the information processing task, it turned out that the amount of code (lines of the program) is growing catastrophically. In addition, the information processing is pretty unified and individual blocks are used repeatedly.

All this made us to realize the idea, which sometimes arose for several years - to create a library of functions that are added to the text of the program before it is executed (user-defined functions library). Before, the lack of a task of sufficient complexity and volume interfered to implement this idea, but not natural laziness (without which the programmer easily turns into a pedant-literal).

So, the task of creating a user library has appeared. The implementation took much less time than the description.

Let's explain the work with the user functions library on a concrete simple example.

Let two programs must be developed within the framework of one task (project). The first program deletes extra spaces from each line of the input file, and the second one also replaces the special **html** characters in the sequence that these characters encode (for example, the quote is encoded by the sequence **"**;

Ex1. **_aw** file implements the first task.

The line, that begins with **#libpath** characters, defines the path to the library of functions (#1).

The line, that begins with **#include** characters, defines the name of file, from which we insert the information into the source program (#2).

```
1 #libpath c:/awklib
2 #include trim.f
4 {
5     print (trim($0));
6 }
```

Ex2._aw file implements the second task.

```

1 #libpath c:/awklib
2 #include trim.f
3 #include HTML_spec.f
4
5 {
6     print HTML_spec(trim($0));
7 }
```

These functions, that are inserted into the source programs must be placed in the library.

Trim.f file contains the function **trim**, that deletes extra spaces from the string (all spaces at the beginning and end of the line, any sequence of spaces in the middle of the line is replaced by one space).

```

1 function trim(str){
2     trim_str = str;
3     gsub(/^ +/, "", trim_str);
4     gsub(/ +$/, "", trim_str);
5     gsub(/ {2,}/, " ", trim_str);
6     return trim_str;
7 }
```

HTML_spec.f file contains a function that encodes special **html** characters (for more details about this function see [5]).

```

1 function HTML_spec(str, s){
2     s = str;
3     gsub(/\&/, "\\\\\\\&amp;", s);
4     gsub(/\"/, "\\\\\\\&quot;", s);
5     gsub(</, "\\\\\\\&lt;", s);
6     gsub(>/, "\\\\\\\&gt;", s);
7     return s;
8 }
```

__prj.bat file performs "preprocessing" working of project files, i.e. inserts the contents of files from the **#include** lines into the program texts.

```

1 gawk --re-interval -f c:/awklib/pawk.mk ex1._aw >
ex1.awk
2 gawk --re-interval -f c:/awklib/pawk.mk ex2._aw >
ex2.awk
```

For adding lines from files, **pawk.mk** program (in **awk** language) answers. This program is very simple and does not need special comments.

```

1 function trim(str){
```

```

2      trim_str = str;
3      gsub(/^ +/, "", trim_str);
4      gsub(/ +$/, "", trim_str);
5      gsub(/ {2,}/, " ", trim_str);
6      return trim_str;
7  }
9  function print_file(fn){
10     while(r = getline s < fn){
11         if(r == -1){
12             print "ERROR: " fn " - " ERRNO;
13             return -1;
14         }
15         print s
16     }
17     return 0;
18 }
20 BEGIN{
21     lp = "";
22 }
24 /^#libpath/{
25     s = $0;
26     gsub(/^#libpath/, "", s);
27     lp = trim(s) "/";
28     next;
29 }
31 /^#include/{
32     s = $0;
33     gsub(/^#include/, "", s);
34     fn = lp trim(s);
35     print_file(fn);
36     next;
37 }
39 {
40     print $0;
41 }

```

As a result, we get two ready-to-run program files.

Ex1.awk file.

```

1      function trim(str){

```

```

2      trim_str = str;
3      gsub(/^ +/, "", trim_str);
4      gsub(/ +$/, "", trim_str);
5      gsub(/ {2,}/, " ", trim_str);
6      return trim_str;
7  }
10 {
11     print (trim($0));
12 }

```

Ex2.awk file.

```

1  function trim(str){
2      trim_str = str;
3      gsub(/^ +/, "", trim_str);
4      gsub(/ +$/, "", trim_str);
5      gsub(/ {2,}/, " ", trim_str);
6      return trim_str;
7  }
9  function HTML_spec(str, s){
10     s = str;
11     gsub(/\&/, "\\\\\\\\\\\&amp;", s);
12     gsub(/\"/, "\\\\\\\\\\\&quot;", s);
13     gsub(</, "\\\\\\\\\\\&lt;", s);
14     gsub(>/, "\\\\\\\\\\\&gt;", s);
15     return s;
16 }
19 {
20     print HTML_spec(trim($0));
21 }

```

References:

1. Bocharov Boris. AUTOMATIZED WEB PAGES PARSING AND CREATION / Boris Bocharov, Maria Voevodina // Information technologies in education: electronic supplement to the journal "Educational Institutions Libraries". – 2017. – N5, p. 1-5.

2. Бочаров Б. Особливості формування фонду електронної бібліотеки навчального закладу // Б. Бочаров, М. Воєводіна, І. Рябченко // Вісник Книжк. палати. – 2006. – №. 12. – С. 37-38

3. Бочаров Б.П. Формирование отчетов в электронных каталогах / Б.П. Бочаров, М.Ю. Воєводина // Библиотеки учебных заведений. – 2003. – №. 10. – С. 41-61.

4. Бочаров Б. П. АWK-универсальная программа работы с текстовыми файлами / Б.П. Бочаров, М.Ю. Воєводина // Библиотеки учебных заведений. – 2002. – №. 4. – С. 39-51.

5. Попов В.М. Модель адаптивной системы техногенной безопасности региона / В.М. Попов, И.А. Чуб, М.В. Новожилова // Системи управління, навігації та зв'язку.–2013.–Вип. 2 (26). – 2012. – С. 120-123.

6. Попов В.М. Показатели эффективности региональной системы техногенной безопасности / В.М. Попов, И.А. Чуб, М.В. Новожилова // Вестник университета гражданской защиты МЧС Беларуси. – 2014. – №. 2. – С. 32-41.

7. Chub I.A., Novozhylov M. V., Murin M. N. Optimization problem of allocating limited project resources with separable constraints / I.A. Chub, M.V. Novozhylov, M.N. Murin // Cybernetics and Systems Analysis. – 2013. – Т. 49. – №. 4. – С. 632-642.