

Using regular expressions for input data verification

A.M. Nefedova, Y.E. Mishina

Samara National Research University, Samara, Russia

Background. Currently, there is an acute issue of checking the correctness of the obtained value in programs of various kinds. Thus, writing a function responsible for validating input data becomes a necessity. Due to the abundance of variations in input data patterns, the programmer faces the problem of choosing the optimal tool for implementing validation. In this case, regular expressions and functions from the <regex> library become a useful and convenient tool with rich functionality. If you know the syntax of the formal regular expression language and how some of the functions in the <regex> library work, you have ample opportunity to check for valid values. At present regular expressions are not very common. We would like to test this hypothesis and show in practice the usefulness of regular expressions.

Aim. To develop a function to validate input against a registry of regular expressions. To find out whether the students are familiar with regular expressions.

Methods. The set aim was solved with the help of the whole complex of methods. Literature review on the topic and the selection of materials on which further research was to be carried out was followed by a number of attempts to develop a well-functioning code. A regular expression is discussed in terms of a pattern that the regular expression engine attempts to match in the input text [1]. A pattern may consist of one or several characters (literals, operators, or constructs), which are divided into several categories. There are also special regular expression language elements. In order to match them in a regular expression, they must be either escaped or included in a special group called positive character group [2].

The resulting code underwent the steps of testing and refinement until proved to be valid. In order to get information on students' awareness on the problem of regular expressions an electronic survey was designed. The survey data was processed and analyzed.

Results. The first step of the research resulted in developing the pattern of the floating point arithmetic. Further analysis of the work of functions from the regex library resulted in design of the function. The function helps to validate input data using regular expressions. The regular expression was then tested and refined.

A survey was also conducted among students mainly of technical specialties. The survey data revealed that the majority of respondents are aware of regular manifestations and know how to use them, but rarely put them into practice. In general, the results of the survey showed a positive attitude towards the topic of regular manifestations, as well as the fact that many students would like to expand their knowledge on this topic.

Conclusions. The input validation feature was implemented using regular expressions. Data obtained through a survey of students, most of whom major in sciences, demonstrated that the majority of respondents have a positive attitude towards regular expressions.

Keywords: regular expressions; input data verification; pattern of the floating point arithmetic.

References

1. learn.microsoft.com [Электронный ресурс]. Regular Expression Language — Quick Reference [дата обращения: 01.03.2023]. Доступ по: <https://learn.microsoft.com/en-us/docs/>
2. learn.microsoft.com [Электронный ресурс]. <regex> functions [дата обращения: 01.03.2023]. Доступ по: <https://learn.microsoft.com/en-us/docs/> (date of the application: 01.03.2023).

Information about authors:

Anastasia M. Nefedova — student, group 6105-010302D, Informatics and Cybernetics Institute; Samara National Research University, Samara, Russia. E-mail: newpinksnow@gmail.com

Yulia E. Mishina — PhD (Philology), Associate Professor, Department of Foreign Languages and Russian as a Foreign Language, Samara National Research University, Samara, Russia. E-mail: mishina.yue@ssau.ru