OPINION

by Prof. Dr. Juliana Peneva Pashkova

Lecturer at New Bulgarian University, dept. "Informatics"

of a dissertation for awarding the educational and scientific degree 'doctor'
In the field of higher education 4. Natural sciences, mathematics and informatics.
Professional direction 4.6 Informatics and computer sciences.
Doctoral program Informatics

Author: Nikolay Georgiev Handzhiyski
Topic: An iterative parsing algorithm with application in the profiling of parsers
Scientific supervisor: *Prof. Elena Petrova Somova, PhD*
                                Plovdiv University "P. Hilendarski"

1.   General presentation of the procedure and the doctoral student

By order No. PD-21-654 dated 21.03.2024 of the Rector of the Plovdiv University "Paisiy Hilendarski" (PU), I have been appointed as a member of the scientific jury to ensure a procedure for the defense of a dissertation on the topic " An iterative parsing algorithm with application in the profiling of parsers " for the acquisition of the educational and scientific degree 'doctor' in the field of higher education 4. Natural sciences, mathematics and informatics, professional direction 4.6 Informatics and computer sciences, doctoral program Informatics. The author of the dissertation is Nikolay Georgiev Handzhiyski. He is a full-time doctoral student at the Department of Computer Informatics with scientific supervisor Prof. Elena Petrova Somova, PhD from PU "Paisiy Hilendarski" - FMI.

The presented set of materials on paper is in accordance with Article 36 (1) of the Regulations for the Development of the Academic Staff of the PU, and contains:

- request to the Rector of the PU to open the procedure for the defence of a dissertation work;
- CV in European format;
- protocol from the departmental council, related to reporting the readiness to open the procedure and preliminary discussion of the dissertation work;
- abstract - in Bulgarian and English;
- declaration of originality and authenticity of the attached documents;
- verification for compliance with the minimal national requirements;
- list of scientific publications on the topic of the dissertation;
- dissertation work;
- copies of scientific publications;

Six publications of the PhD student are attached.

PhD student Nikolay Handzhiyski has a Master's degree in Informatics from FMI. During his studies he was an Erasmus student at the University of Koblenz-Landau, Koblenz, Germany and a Cepus student at the University of Johannes Kepler, Linz, Austria. Currently, he works as a software engineer in a Bulgarian company developing software for management of retail outlets and as a manager of a German company with the same subject of activity. He has pedagogical experience as a part-time lecturer in Computer science disciplines at the FMI of Plovdiv University as well as participation in research at Koblenz-Landau University, Koblenz, Germany

2. Relevance of the dissertation subject

The dissertation is dedicated to the automation of the parsing process related to the extraction of structured information from unstructured data. Parsing specific data means recognizing their belonging to a particular language and displaying specific structural information about these data. Parsing algorithms build a data structure that can be used by a compiler, interpreter, or translator to create an executable program or library. Parsing algorithms are fundamental to the analysis and processing of linguistic structures, especially in the context of the increasing amount of data generated by different applications and platforms.

The objective of this research is the development of tools (parsing machines and algorithms) that are suitable for linear translation of data based on a given grammar. Given the applicability of these tools for the development of technologies and the interest of leading companies in them, I find the subject of the dissertation a problem of the present day.

3. Knowledge of the problem

The review of the cited literature (a total of 191 titles, 4 in Cyrillic and 188 in Latin, of which 10 are Internet sources) makes it possible to claim that the PhD student has entered the issues well enough. The list is representative by number, by year and by distribution of authors. The Ph.D.'s overview of the formal tools related to the mathematical foundations of translators and programming languages shows an in-depth study of the state of research in the field under consideration. The basic concepts such as combinatorial schemes, regular expressions, context-free grammars, finite and stack automata are presented. Translators as a means of converting programs from one formal language to another and the associated lexical and syntactic analysis have been studied. The possibilities of various compilers, lexer and parser generators have been investigated. The author has thoroughly studied the problems of the dissertation study.

4. Research methodology

The methodology applied by PhD student Nikolay Handzhiyski derives from the set goals and the separate research tasks. The methodological components of a dissertation development with an applied focus such as: state analysis, theoretical model, software implementation, an experiment proving the usability of the developed tools are remarked.

After conducting research in the field and defining the problem, a well-founded parsing machine architecture was created and a new type of context-free grammars was defined. A corresponding parsing algorithm built into the parsing engine has been developed. A prototype of a tool for measuring and comparing the resources used by different parsers was proposed and a series of experiments were carried out with it, proving the practical applicability and usability of the development. It can be asserted that the doctoral student used methodological techniques inherent in correct scientific research.

5. Characterization and evaluation of the dissertation and contributions

The dissertation with a total volume of 181 pages contains the following components: list of tables, list of figures, table of used terms from the literature, table of summarized terms in the dissertation, list of used abbreviations, introduction, four chapters, conclusion, contributions, perspectives, list of author's publications on the subject, a list of noted citations, participation in projects, appendices, a list of the literature used and a declaration of originality.

The work is illustrated by appropriately selected figures (total 56) and tables (total 8), and as a volume it is completely in the norms. The formulated six tasks arise from the aim of the thesis research, namely to develop tools that are suitable for linear data translation based on some multivalued context-free grammars. The contributions are scientific, scientific-applied and applied, covering: development of a conceptual model of a parsing machine with new functionalities compared to the machines used so far; define specific grammars with phrase symbols; creation of a tunnel parsing algorithm with linear parsing time for certain grammars; design of a metaprogramming language for grammars; realization of a parser generator, allowing the performance of tests (measurement of the resources used during recognition and parsing) with parsing machines created by different parser generators and compilers.

Perspectives for the development of the topic include: expanding the list of commands that are sent by the parser module and accepted by subsequent modules in order to generate syntax structures other than syntax trees; upgrading the proposed Tunneling algorithm for parsing based on left recursion grammars; adding new features to the profiler (maintenance of additional parser generators and compilers); measuring various characteristics of the "parsing" process.

6. Evaluation of the publications and personal contributions of the doctoral student.

The doctoral student has presented a list of six publications on dissertation research: 4 - in specialized journals, 1 - in the proceedings of an international conference and 1 - in the papers of an international forum. The publications are indexed: 4 - in Web of Science and 5 - in Scopus, a total of 5 are in editions with SJR, reflecting the main results obtained in the dissertation. All publications are co-authored with the supervisor. A citation of 1 of the publications on the topic in a scientific study that is indexed in the world's databases was spotted.

The results have been sufficiently presented to a specialized scientific audience. As a number, publications are sufficient. The specific requirements of the FMI have been met. The subject matter of the publications emphasizes the personal participation of the doctoral student

7. Abstract

The abstract is compiled according to the requirements and presents sufficiently all aspects of the dissertation research. The main results achieved and the contributions of the author are summarized.

8. Recommendations for future use of dissertation contributions and results

Parsing algorithms are a basic tool for data processing. Dissertation contributions and results can find application in any processes for analysis and structuring of input data.

CONCLUSION

The dissertation *contains scientific, scientific-applied and applied results, which represent an original contribution to science* and **meet all the requirements** of the Law on the Development of the Academic Staff in the Republic of Bulgaria (ZRASRB), the Regulations for the Implementation of the ZRASRB and the relevant Regulations of PU "Paisiy Hilendarski". The presented materials and dissertation results fully correspond to the specific requirements of the Faculty of Mathematics and Informatics, adopted in connection with the Regulations of the PU for the application of the ZRASRB.

The results and contributions are original and not borrowed from foreign research and publications.

The dissertation shows that PhD student Nikolay Georgiev Handzhiyski **possesses** in-depth theoretical knowledge and professional skills in the scientific specialty of informatics, **demonstrating** qualities and skills for independent conduct of scientific research.

Due to the above, I confidently give my *positive assessment* of the conducted research, presented by the above-reviewed dissertation work, abstract, achieved results and contributions, and I *propose to the honorable scientific jury to award the educational and scientific degree "doctor"* to Nikolay Georgiev Handzhiyski in the field of higher education: 4. Natural sciences, mathematics and informatics, professional direction 4.6 Informatics and computer sciences, doctoral program "Informatics".

May 01, 2024

Opinion prepared by: ....

Assoc. Prof. Juliana Peneva