O P I N I O N

for the procedure for defence of PhD thesis
for obtaining the educational and scientific degree "Doctor"

Higher education field: *4. Natural sciences, Mathematics and Informatics*
Professional area: *4.6. Informatics and computer science*
Doctoral program: *Informatics*

Author: **Nikolay Georgiev Handzhiyski**, full time PhD student at the Department of Computer Informatics, Faculty of Mathematics and Informatics, Plovdiv University Paisii Hilendarski
Topic: "**An Iterative Parsing Algorithm with Application in the Profiling of Parsers**"
Scientific supervisor: **Prof. Elena Petrova Somova**, PhD

Opinion author: Galina Mircheva Ilieva, Professor in 3.8. Economics, PhD in Informatics, Plovdiv University Paisii Hilendarski

1. General description of the procedure and the PhD candidate

This opinion has been prepared according to order No. RD–21–654/21.03.2024 of the Rector of Plovdiv University Paisii Hilendarski (PU), Prof. Dr. Rumen Mladenov, based on a decision (Protocol No. 5/20.03.2024) of the Faculty Council of the Faculty of Mathematics and Informatics (FMI) and a report by Prof. Dr. Angel Atanasov Golev, dean of the FMI, in accordance with Art. 4. of ZRASRB, Art. 2. (2), Art. 30. (3) of the Regulations for the Application of the RSARB and Art. 37. (1) of the Regulations for the Development of the Academic Staff of Plovdiv University Paisii Hilendarski (PRASPU). In the above-mentioned order, I have been appointed as a member of the scientific jury to ensure a procedure for the defence of a dissertation titled "An Iterative Parsing Algorithm with Application in the Profiling of Parsers" for obtaining the educational and scientific degree "Doctor" in higher education field 4. Natural sciences, mathematics and informatics, Professional area: 4.6. Informatics and computer science, Doctoral program: Informatics. The scientific supervisor is Professor Elena Petrova Somova, PhD.

Nikolay Georgiev Handzhiyski completed his Master's degree in "Software Technologies" at Plovdiv University Paisii Hilendarski. He is fluent in English and German.

2. Relevance of the topic and expediency of the set goals and objectives

Improving parsing algorithms is a current issue in modern computer science for several reasons. First, they enhance the software development process by providing advanced capabilities for syntactic analysis. Embedded within parsers, they enable developers to detect errors more effectively and offer features for automatic code completion (autocomplete) in Integrated Development Environments (IDEs). These changes lead to improved code quality, faster development cycles, and increased developer productivity. Second, these algorithms play a significant role in computational linguistics by

enabling accurate syntactic analysis and correct semantic understanding of the language. They address various tasks, such as machine translation and information extraction, contributing to the development of complex language processing systems in various fields.

The dissertation is focused on the study of the theoretical foundation and practical application of parsers for a specific type of ambiguous context-free grammars. The following elements have been developed: 1) a special theoretical framework; 2) a parsing machine; 3) a parsing algorithm; and 4) a profiler. The latter facilitates the verification of the author's ideas by generating context-free grammars and input data for experiments with them, with automated complexity tracking of parsing. The doctoral study encompasses six tasks.

3. Knowledge of the problem

Nikolay Handzhiyski explores a variety of scholarly literature on the investigated topic. He is well-versed in the fundamental aspects of translator design, particularly algorithms for recognizing and parsing languages defined by context-free grammars.

4. Research methodology

In order to accomplish the objectives of the dissertation, the doctoral student employs the following scheme: 1) research and analysis of specialized literature related to the subject of investigation; 2) development of a parsing machine to explore the capabilities of various parsers; 3) presentation of the tunnel parsing algorithm and the necessary objects for its operation; and 4) verification of the proposed methodology through a series of experiments conducted using a profiler tool specially created for this purpose.

The methodology applied by Nikolay Handzhiyski is appropriate and effectively addresses the tasks and goals of the doctoral research.

5. Characteristics and evaluation of the PhD thesis and its contributions and significance

The thesis consists of a total of 181 pages, organized into an introduction, four chapters, a conclusion, and a bibliography. The main body of the thesis spans 163 pages, supplemented by two appendices about the profiler, each one pages long. The list of references contains 191 titles, with 187 in English, two in Russian, and two in Bulgarian.

In the first chapter, fundamental terms from the theory of formal languages and abstract machines related to the process of translation and processing of program code are defined. The characteristics and applications of finite automata and regular expressions are presented, types of grammars

are classified, the components of translators are described, and an overview of existing parsing algorithms for context-free grammars is provided. The second chapter contains a detailed description of the proposed parsing machine. It is of a hybrid type, allowing parsing without or with lexical analysis. Each component of the parsing machine is described in detail based on its functional capabilities. The third chapter presents the pseudocode of a new parsing algorithm for non-left recursive advanced grammars – Tunnel Parsing (TP). The algorithm is demonstrated based on a specific example with a given input string. The fourth chapter describes a software tool called a profiler for generating context-free grammars and input data for parsing. Additionally, this tool facilitates the reporting of experiments with parsing machines. The chapter also introduces the newly created language for creating pattern grammars. The results obtained from the four experiments conducted with the profiler are analyzed. In the Conclusion, the obtained results are summarized and systematized, highlighting the main scientific, scientific-applied, and applied contributions of the dissertation work. Perspectives for future development of the dissertation topic are formulated.

The obtained results confirm the hypothesis of the dissertation work. They are significant and represent a contribution to the development of efficient parsers based on certain types of advanced grammars with linear time complexity.

The main contributions of the doctoral candidate are scientific, scientific-applied, and applied. The scientific contributions of the dissertation research are as follows: A conceptual model of a parsing machine with enhanced capabilities compared to existing analogs has been developed. Advanced grammars with phrase symbols have been defined using notation similar to the augmented Backus-Naur form. A model of a phrase machine has been proposed, which speeds up the analysis by pre-categorizing phrases from the parser grammar. The scientific-applied contributions are: the functionality of the parsing machine has been described; an algorithm for “tunnel parsing” has been created, parsing based on non-left recursive advanced grammars with linear time complexity; a metalanguage for grammar metaprogramming has been designed and implemented; a profiler of parser generators has been designed and implemented, allowing tests with parsing machines created by different parser generators and accounting the computational resources used. The applied contributions of the dissertation research are: 1) a prototype of a software tool – a profiler of parser generators has been developed for experiments with both directly input grammars and a set of grammars programmatically generated by the tool itself, with a built-in module for visualization of the results; and 2) experiments to assess the efficiency and complexity of parsers have been conducted using the profiler.

6. Evaluation of the scientific publications

Regarding the doctoral dissertation, the doctoral candidate has presented 6 publications. The six publications are in English and are co-authored, with Nikolay Handzhiyski as the first author. Four of the publications are journal articles, while the remaining two have been published in proceedings of scientific conferences. All publications are indexed in the global bibliographic databases – Web of Science and/or Scopus. It is impressive that two of the articles have been published in a journal with impact factor.

I am convinced that the proposed doctoral dissertation is the personal work of the author.

No plagiarism in the submitted scientific works has been proven according to the legally established procedure.

7. Abstract

The presented abstract meets the requirements of the law and relevant regulations, presenting the achieved research results and contributions.

8. Recommendations for future use of thesis contributions and results

I would recommend that Nikolay Handzhiyski continue the future research he declared.

CONCLUSION

From the above it can be seen that Nikolay Georgiev Handzhiyski is a highly qualified expert in the field of computer science with accumulated experience and proven achievements in this field. He has knowledge and skills for scientific research. This gives me reason to conclude that, in addition to meeting the minimum national requirements, the dissertation, the abstract and the presented scientific production satisfy the requirements of the ZRASRB, the Regulations for its implementation, as well as the Regulations of the PU Paisii Hilendarski and the specific requirements of the Faculty of Mathematics and Informatics at the PU for the development of the academic staff, required for candidates to obtain the educational and scientific degree "Doctor".

Due to the above, I give my **positive assessment** of the conducted research, presented by the above-reviewed dissertation thesis, abstract, achieved results and contributions, and I invite highly respectable Scientific Jury to award the educational and scientific degree "Doctor" to Nikolay Georgiev Handzhiyski in the Higher education field 4. Natural sciences, mathematics and informatics, Professional area: 4.6. Informatics and computer science, Doctoral program Informatics.

09.05.2024 **Opinion prepared by**: ……

Plovdiv (Prof. Galina Ilieva)