

## OPINION

on  
the dissertation of **Alexander Nikolaev Popov** on “Modeling Lexical Knowledge for Natural Language Processing” for acquiring PhD degree in Informatics and Computer Sciences

by

**Assoc. Prof. Dr. Gennady Pavlovich Agre**

Institute of Information and Communication Technologies at Bulgarian Academy of Sciences

By order № 127/12.07.2018 of the Director of the Institute of Information and Communication Technologies - Bulgarian Academy of Sciences, I have been approved as a member of the Scientific Jury in connection with the procedure for acquiring educational and scientific degree "doctor" in the doctoral program "Informatics" in professional field 4.6 "Informatics and Computer Sciences" by Alexander Nikolaev Popov with a dissertation entitled "Modeling Lexical Knowledge for Natural Language Processing" with the Scientific Supervisor Assoc. Prof. Kiril Simov.

As a member of the jury I received:

1. A copy of the dissertation for acquiring educational and scientific degree "doctor" in professional field 4.6 "Informatics and Computer Sciences" (in English)
2. An abstract of the dissertation (in Bulgarian)
3. Professional Curriculum Vitae
4. Copies of the author's publications related to the dissertation.

The dissertation consists of 145 pages, structured in: Introduction - Chapter One, 7 chapters, Conclusion, Bibliography and three Annexes.

### **1. Actuality of the dissertation topic and relevance of its objectives and tasks**

The dissertation is related to the modeling of lexical knowledge in natural languages, a task underlying for solving such important problems in the field of natural language processing as machine translation, word sense disambiguation, word similarity calculation, etc. From this point of view, the topic of the dissertation is completely up to date, and the objectives and tasks to be solved are expedient.

### **2. Knowledge of the state-of-the-art in the field**

The PhD candidate showed a very good knowledge of the thematic field as well as of topics related to the realization of the dissertation goals. The third chapter analyzes the main approaches for representation of the lexical knowledge intended for automatic processing. The bibliography contains 131 sources, and it should be noted that Alexander Popov is the author of a comprehensive review of the use of recurrent neural networks for word sense disambiguation published in a journal having SJR rank of SCOPUS.

### **3. Methodology of the study**

The research methodology chosen by the author derives from the dissertation objectives and tasks. The adequacy of the use of relational knowledge for lexicon modeling, as well as the

use of architectures based on recurrent neural networks, has been convincingly demonstrated by proposed solutions of the task of removing lexical ambiguity.

#### **4. Short analytical resume of the dissertation**

The first chapter is an introduction to the thematic field of the study. It gives the motivation for choosing the dissertation topic and points out the objectives and tasks of the dissertation. In the second chapter more formal definitions of the dissertation tasks are given: solution of the parts-of-speech (POS) tagging task, removing the lexical ambiguity as well as determining the similarity and the relations between words. The third chapter analyzes the main results achieved in the thematic field chosen. At the beginning, the explicit models for representing lexical knowledge are examined in details. Then an analytical overview of the existing methods for solving word sense disambiguation task and of the use of neural networks to solve various tasks related to natural language processing are considered. The main dissertation contributions are presented in the next four chapters. Chapter 4 described an original solution for the POS tagging task. Chapter 5 presents the author's work on enriching the structure of the popular semantic resource WordNet, as well as on evaluating some new resources from the point of view of their use by knowledge-based methods for removing lexical ambiguity. Chapter 6 enhances the work presented in the previous chapter by approaches for learning distributed patterns (embeddings) of words and their senses. Chapter 7 presents two alternative approaches for solving word sense disambiguation task by using recurrent neural networks. Chapter 8 analyzes the possibility of using hybrid architectures based on multi-task learning for solving various tasks related to the modeling of lexical knowledge. The conclusion summarizes the content and the results of the dissertation and presents the dissertation contributions, the publication activity of the author, the information about his participation in various projects and conferences for presenting the results, as well as future directions of the work.

#### **5. Contributions of the dissertation**

The main contributions of the author are stated correctly. The results were obtained and used within two international projects under the 7th Framework Program and one national project funded by the National Science Fund, which is an additional guarantee for the result quality.

#### **6. Assessment of publications related to the dissertation**

The author presented 14 publications related to the dissertation, all of which were published in referenced journals and conference proceedings: 8 publications are indexed in SCOPUS (5 of them - in the Web of Science) and the other 6 - in proceedings of international conferences and workshops. The publications reflect the content of the dissertation and represent the original achievements of the dissertation' author.

According to the "Regulations for the specific conditions for acquiring academic degrees and positions at IICT-BAS", the PhD candidate must have at least 3 publications related to the dissertation and at least one of them should be published in a journal with impact factor/rank or in a proceedings of a specialized international conference.

According to the Rules for applying the Law on the Development of the Academic Staff in the Republic of Bulgaria, dated to 06.07.2018, the minimum national requirements for

acquiring a PhD degree is 30 points collected from publications referenced and indexed in the world well-known databases for scientific information, as well as from articles published in non-referenced journals with scientific review and from edited proceedings. According to the formulas presented in the Rules, the publications of Alexander Popov, indexed in SCOPUS, have 156.65 points, while the rest of the publications - 59 points.

These facts show that the publishing activity of Alexander Popov significantly exceeds both the minimum national requirements and the specific requirements of the Institute of Information and Communication Technologies - BAS for obtaining the PhD degree.

In addition, it should be noted that up to now 5 of the Alexander Popov's papers are cited 8 times (in total) by foreign scientists.

### **7. Personal contribution of the PhD candidate**

The personal contribution of the PhD candidate is significant, which is proved by the fact that he is the only author of 4 publications (2 of them are indexed in SCOPUS and Web of Science), in one publication he is pointed as the first author, in 5 publications – as the second author and in 4 – as the third one.

### **8. Abstract of the dissertation**

The abstract of the dissertation correctly presents the objectives, tasks and the results of the dissertation.

### **9. Conclusion**

From all mentioned above I can to conclude that all requirements of the Law on the Development of the Academic Staff in the Republic of Bulgaria, the Rules for applying the Low, as well as of the Regulations for the specific conditions for acquiring academic degrees and positions at IICT-BAS are fully satisfied. I am confident that the dissertation of Alexander Nikolaev Popov has all the characteristics required by a PhD Thesis. The results presented in the dissertation are significant and contribute to the development of the existing methods for modeling lexical knowledge.

All this facts give me the reasons for a positive evaluation of the dissertation and I strongly recommend the honored Scientific Jury to award the educational and scientific degree "doctor" in professional field 4.6 "Informatics and Computer Sciences" to Alexander Nikolaev Popov.

October 10, 2018

Sofia