

## Opinion

of Prof. Dr. Maria Nisheva – FMI, Sofia University St. Kliment Ohridski  
on Doctoral Thesis in Professional Area 4.6 Informatics and Computer Science  
Thesis Title: “Modeling Lexical Knowledge for Natural Language Processing”,  
Professional Area: 4.6 Informatics and Computer Science  
Author: Alexander Nikolaev Popov

Alexander Popov has Bachelor's degree in English Philology from Sofia University St. Kliment Ohridski, and Master's degree in Computer Science from Imperial College, London.

He has worked and continues to work as a developer of natural language processing software. He has also significant teaching experience and experience in working on international research projects.

Alexander Popov has strong research results in the fields of natural language processing, knowledge extraction, semantic technologies.

### Doctoral Thesis: Content

The doctoral thesis of Alexander Popov is devoted to research in the field of natural language processing. Its main objective is to analyze different aspects of lexical knowledge and to develop distributional models of lexical knowledge that encompass knowledge derived from the theory.

The dissertation consists of 150 pages of English text, including nine chapters, a list of references, and three appendixes.

The introductory first chapter presents in brief the field of research. The objectives of the dissertation are formulated and argued.

The second chapter defines, at the appropriate level of abstraction, the main tasks that derive from the aims of the dissertation.

The third chapter contains an overview of the area of research, with discussion of the main achievements and still open questions in it.

The fourth chapter is devoted to the study of recurrent neural networks (RNN) with respect to their applicability as a tool for part-of-speech tagging. Distributed presentations for Bulgarian language and a model for presentation of morphological information are proposed, which are of key importance for the correct solution of the task of determining the parts of the speech.

Chapter 5 presents the results of the author's research on the hypothesis that the enrichment of so-called knowledge graphs with additional paradigmatic and syntagmatic

relations can be accomplished by extracting information from textual, semi-structured and structured resources.

The sixth chapter discusses results obtained in relation with the construction of distributed representations of different types of lexical units: words, lemmas and senses.

Chapter 7 presents results from the study of the applicability of RNN for word sense disambiguation.

The eighth chapter is devoted to research results in parallel training of RNN on various natural language processing tasks based on lexical knowledge.

Chapter 9 summarizes the results of the author's research on the topic of his doctoral thesis.

### **Doctoral Thesis: Results**

The main scientific and applied scientific contributions of the doctoral thesis of Alexander Popov may be summarized as follows:

- An approach to modeling the lexicon through relational knowledge has been developed and the validity of this approach has been shown;
- It has been shown that encoding relational lexical knowledge in probabilistic distributed models is a powerful tool for lexical modeling;
- Several novel contributions in terms of representing the lexicon via vector space models are put forward and an original approach to encoding morphological information has been developed;
- A new approach to supervised word sense disambiguation has been proposed;
- Several RNN architectures for sequence-to-sequence tagging have been designed and implemented;
- Multiple sets of new relations between word senses have been generated and represented in the established format of WordNet;
- Various vector space models have been trained and evaluated.

The dissertation makes an excellent impression with the scope, depth and argumentation of the presentation. The field of research is actual and complex and the achievement of significant results in it requires very good interdisciplinary knowledge, permanent and intensive work. The results achieved by the author are original and meaningful and correspond to the declared goals.

### **Publications on the Doctoral Thesis. Reflection on the Works of Other Authors**

The author has worked on 14 publications related to parts of his doctoral thesis. He is the only author of four of these publications, the remaining 10 are co-authored. I have no doubt about the personal contribution of Alexander Popov in the collective publications.

Information about eight citations of publications on parts of the dissertation of Alexander Popov is provided by the author. In the meantime, another article citing two of the mentioned works of Al. Popov, was also published.

### **Critical Remarks and Recommendations**

I have no significant critical remarks on the content of the doctoral thesis. As a weak point side of the dissertation I would notice the lack of a short and general statement of its contributions.

### **Abstract**

The abstract corresponds to the content and the contributions of the dissertation, but its volume is too large. Probably this is explained by the author's decision to write his dissertation in English and the willingness to present a sufficiently detailed abstract in Bulgarian language.

### **Summary**

Summing up, I consider that the doctoral thesis of Alexander Popov fully satisfies and exceeds the requirements of the national regulations and the specific conditions and requirements of the Institute of Information and Communication Technologies at the Bulgarian Academy of Sciences. Its author has achieved significant research results that make an original contribution to the chosen field of study. Therefore, **I strongly advise the honorable scientific jury to award to Alexander Nikolaev Popov the educational and scientific degree "Doctor" in professional area 4.6 Informatics and Computer Science.**

Sofia, October 8, 2018

