Opinion

from Prof. Dr. Alexandra Andreeva Soskova
Faculty of Mathematics and Informatics, SU
(Member of scientific jury, order No. 109/29.04.2024
of the Director of IMI-BAN)

for the dissertation of ***Dimiter Dimitrov Dobrev***

"Artificial Intelligence – Definition, Implementation and Consequences"

presented for the acquisition of a scientific degree "doctor" in scientific field 4. "Natural sciences, mathematics and informatics", professional direction 4.5 "Mathematics", scientific specialty "Mathematical logic".

1. General characteristics of the dissertation work and the presented materials

The presented dissertation work is written in Bulgarian, 115 pages long, including a list of literature containing 92 titles. The dissertation, supervised by Prof. Dr. Lyubomir Ivanov, is in the field of Artificial Intelligence, an intensively developing scientific field. The dissertation begins with an introduction, followed by three chapters in which the author answers the questions: "What is AI?", "How to create it?" and "What will be the consequences of its creation?" and ends with a conclusion, publications related to the thesis, a statement of originality of results, and references. The abstract is 37 pages and correctly reflects the contributions of the PhD student.
The presented materials include all the documents required by the Law on the development of the academic staff in the Republic of Bulgaria and the Regulations for its application, as well as those required by the relevant Regulations of the BAS. The reference (document 8), attached in accordance with the requirements of NACID, is correctly made and shows that the applicant meets the minimum national requirements.

2. Data and personal impressions about the candidate

Dimiter Dobrev graduated in 1995, majoring in Mathematics, specialization "Mathematical logic and its applications" at the Faculty of Mathematics and Informatics of Sofia University. His diploma thesis "Periodic Loops in Prolog" is supervised by Prof. Dr. Dimiter Skordev. Since 1996, Dimiter Dobrev has been a mathematician in the Mathematical Logic section (now Algebra and Logic) at the Institute of Mathematics and Informatics, BAS. He teaches at FMI, SU on "Strawberry Prolog and Artificial Intelligence", "Mathematical Logic", "Logic Programming", "Discrete Mathematics", "Discrete Structures" and at the New Bulgarian University on "Strawberry Prolog and Artificial Intelligence". He participates in 5 research projects and has 6 recognized patents and an application for 3 new patents. He is the author of 21 scientific publications and 11 popular science publications.
I have known Dimiter Dobrev since his student years. He always impressed me with his new ideas, always unconventional. At first it was his passion for Prolog, he made a programming

system "Strawberry Prolog", which he still teaches at FMI to students. Our Logic Programming team, in which I participated, used his system successfully in the exercises. Next came his interest in creating Transliteration and Keyboard Layouts, which he implemented, then he was passionate about subway train movement schemes and Bonus-Malus-like systems. And in recent years, his interest has been exclusively in the definition and basic structure of General Artificial Intelligence. As a colleague, he was always benevolent and responsive. He actively participated with talks in our seminars on Mathematical Logic, and in the various conferences organized by the "Mathematical Logic and its Applications" department. He has participated as a part-time lecturer in several disciplines of the department for many years.

3. Content analysis of the candidate's scientific and scientific-applied achievements, contained in the presented dissertation work and the publications to it, included in the procedure

The theory of "general" artificial intelligence has been developing quite rapidly lately. Intelligence is too complex to be described by a single theory; instead, researchers construct a hierarchy of theories that characterize it at multiple levels of abstraction. At the lowest levels of this hierarchy, neural networks, genetic algorithms, and other forms of computation allow us to understand the processes of adaptation, perception, embodiment, and interaction with the physical world that must underlie any form of intelligent activity. At a higher level are logical formal inference, deduction, induction, algorithms, models, and many other ways of reasoning. Designers of expert systems, intelligent agents, and natural language understanding programs recognize the role of social processes in the creation, transmission, and maintenance of knowledge.
The study of mechanical or "formal" reasoning began with philosophers and mathematicians in antiquity. Logical research led to the Theory of Computability created by Alan Turing, Alonzo Church, Stephen Kleene, Emile Post, and others. This, along with simultaneous discoveries in cybernetics, information theory and neurobiology, led researchers to consider the possibility of building an "electronic brain". Modern research on Artificial Intelligence (AI) began in the mid-1950s. The term "Artificial Intelligence" itself was formally introduced by John McCarthy at the Dortmund workshop in 956.

The first generation of AI researchers were convinced that general artificial intelligence was possible and that it would exist in just a few decades. However, by the early 1970s it became apparent that researchers had greatly underestimated the difficulty of the project. In the early 1980s, the Japanese fifth-generation computer project revived interest in AI. But the goals of the fifth-generation computer project were never realized. In the first decades of the 21st century, access to large amounts of data, cheaper and faster computers have been successfully applied to many problems in all fields. The AI boom began with the initial development of key architectures and algorithms, leading to the development of large language models exhibiting human traits of reasoning, cognition, attention, and creativity.
The dissertation examines a complex model of the work and action of general artificial intelligence. It is a type of artificial intelligence that matches or surpasses human abilities in a wide range of tasks. In contrast, the weak AI can solve a specific problem but lacks general cognitive abilities.

In the first chapter of the dissertation, an informal definition of AI is considered. The author compares it with those known in the literature and shows what improvements he has made. A formal mathematical definition of AI is also given. The author improves on the definition of

AI, originally given by Hernández-Orallo and substantially improved by Marcus Hutter, making the definition of AI independent of lifespan and independent of language for describing the world.

In the second chapter, the author presents a new approach in the study of AI. This is the Event-Driven (ED) approach. The idea behind the ED approach is that the model should not take all the input-output information but should only use the important events. For the description of worlds, the most common tool is the Markov decision process (MDP). It is shown that the ED model is the natural generalization of the Markov decision process. The author gets a simpler model that describes more worlds. It also looks at an advanced model where the state knows everything, using which it introduces interpretation of events and Event-Driven models. A world description language has the advantage that the description can be looked up automatically. It is shown that the language for describing worlds can, through the simple modules of which it is composed, describe quite complex worlds with many agents and complex relationships between them. The dissertation describes a world (the game of chess). In the paper Dobrev (2020a) a computer program is made which, based on this description, emulates the world.

In the third chapter, the author discusses the issue of the consequences of creating general artificial intelligence. With several examples the author warns of avoidable dangers. It looks at the question: how our life will look like after the advent of AI.

4. Approbation of the results

Essentially, the content of the first chapter is the articles:

- Dobrev, D. (2000). AI - What is this. *PC Magazine - Bulgaria*, 11/2000, pp. 12-13
- Dobrev, D. (2022a). The AI Definition and a Program Which Satisfies this Definition. arXiv:2212.03184 [cs.AI].

The content of the second chapter is mainly presented in the articles:

- Dobrev D. (2022b). Language for Description of Worlds. Part 1: Theoretical Foundation. *Serdica Journal of Computing* 16(2), 2022, pp. 101-150.
- Dobrev D. (2023). Language for Description of Worlds. Part 2: The Sample World. *Serdica Journal of Computing* 17(1), 2023, pp. 17-54.

The content of the third chapter is presented in the article:

- Dobrev, D. (2019c). AI Should Not Be an Open-Source Project. *International Journal "Information Content and Processing"*, Volume 6, Number 1, 2019, pp. 34-48.

The author also has 19 publications and one patent that are related to the dissertation but are not part of the text of the dissertation. For example, the concept of Event-Driven model is introduced in Dobrev (2018), the extended model is introduced in Dobrev (2019a). Dobrev (2017a) reviews the test events. How an agent can also influence events that are not its actions is described in Dobrev (2021b). He has presented two publications, in the journals Mathematica Balkanica New Series, indexed in Zentralblatt, and Serdica Journal of Computing, indexed in MathSciNet.

There are 8 citations on the dissertation, of which 3 in Scopus and 119 on the topic, of which 34 in Scopus. In Google Scholar, the article Dobrev (2005a) alone has 127 citations (116 without self-citations). The author has reported his results at many conferences and workshops.

The scientific works meet the minimum national requirements (according to Article 2b, paragraphs 2 and 3 of the Law on the development of the academic staff in the Republic of Bulgaria) for the acquisition of an educational and scientific degree "doctor" in the scientific field 4."Natural sciences, mathematics and informatics" and professional direction 4.5 "Mathematics" ("Mathematical Logic"). The results presented by the candidate in the dissertation work and related scientific works do not repeat those from previous procedures for acquiring a scientific title and academic position. The author declared the originality of his work.

5. Qualities of the auto-reference

The abstract in the Bulgarian and English versions correctly reflects the results presented in the dissertation and is prepared according to the requirements.

6. Critical notes and recommendations

The dissertation is written in an unconventional style. In my opinion, it could be formulated with more mathematical rigor. What struck me is that various new concepts are commented on before they are formally defined. For me, the opposite is more logical. On the plus side, it is an easy read and accessible to a wider audience. Quite specific examples are given, again unconventionally, which, however, help to understand the ideas. But still, the impression of a popular science reading remains. In the last chapter, this is quite prevalent.

6. Conclusion

Having familiarized myself with the dissertation work presented in the procedure and the scientific works accompanying it and based on the analysis of their significance and the scientific and scientific-applied contributions contained in them, **I confirm** that the presented dissertation work and the scientific publications to it, as well as the quality and originality of the results and achievements presented in them meet the requirements of the Law on the development of the academic staff in the Republic of Bulgaria, the Regulations for its application and the corresponding Regulations of the BAS for the candidate's acquisition of the educational and scientific degree "doctor" in scientific field 4. "Natural sciences, mathematics and informatics", professional direction 4.5 "Mathematics", scientific specialty "Mathematical logic".
Based on the above, I recommend the scientific jury to award Dimiter Dobrev an educational and scientific degree "doctor" in scientific field 4. "Natural sciences, mathematics and informatics", professional direction 4.5 "Mathematics" ("Mathematical logic").

10.06.2024. Prepared the opinion: Prof., PhD, Alexandra Soskova
(Academic position, scientific degree, name, surname)