

PEER REVIEW

**on the thesis entitled “Artificial intelligence – definition, realization and consequences
What is it, how can we do it and what shall we do once we do it?”**

submitted for awarding

of the educational and scientific degree “doctor” (Ph.D.)

by Dimiter Dimitrov Dobrev

Scientific field: **4. Natural Sciences, Mathematics and Informatics**

Professional Field: **4.5. Mathematics**

Doctoral Program: **“Mathematical Logic”**

Scientific Advisor: **Assoc. Prof. Ph.D. Lyubomir Ivanov**

Section: **“Algebra and Logic”**

Institute of Mathematics and Informatics of Bulgarian Academy of Sciences

The peer review has been prepared by **Prof. D.Sci. Vesselin Stoyanov Drensky, Full Member of the BAS**, retired professor at IMI – BAS, member of the Scientific Jury by Order No. 108/29.04.2024 of the Director of IMI – BAS.

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

The Ph.D. thesis is presented in two versions – in Bulgarian and English. The English version contains 104 pages, of which 6 introductory pages, 91 pages of main text, 7 pages of final remarks, ending with 92 cited titles. The main part consists of an introduction, ending with a brief reference to the contributions, and three chapters. The rest of the presented materials contain the standard documents for such a procedure (curriculum vitae with a list of publications, data on the doctoral studies, information card of NACID, a list of 5 publications on which the dissertation is based and 20 other publications related to the work, copies of these 25 publications, a list of citations of the papers of the applicant, an author statement for the contributions of the dissertation and the five articles on which it is based, summary in Bulgarian and English).

All materials on the Ph.D. thesis are uploaded with free access to the home page of IMI. I have some critical remarks about the presentation of the publications relevant to the dissertation. Of the five papers included in the dissertation, only 2 are presented in original. One of the other 2 papers is in Bulgarian, and the candidate has submitted an English translation, and the other two are submitted in a preliminary form, although the originals are available online. Of the remaining 20 publications, only 5 are presented in original. Moreover, some of the others (for example, some of those presented in conference proceedings) are presented in a version that is much longer than the original. It would be reasonable, if some of the papers do not have an open access, that their original text is available only to the members of the Scientific Jury, and that the manuscript versions are with open access.

2. Short CV and personal impressions of the candidate

Dimiter Dimitrov Dobrev graduated from the Faculty of Mathematics and Informatics of the Sofia University “St. Kliment Ohridski”, specialization "Mathematical logic and its applications" and defended his Master thesis on the topic “Periodic loops in Prolog” under the supervision of Prof. D.Sci. Dimiter Skordev. During his studies, he was on a five months specialization at the University of Twente, the Netherlands under the Tempus program. From 1996 until now, he works as a mathematician at the Institute of Mathematics and Informatics at the BAS, first in the section “Mathematical Logic”, and after the cooperation of the sections “Mathematical Logic” and “Algebra” in the new founded section “Algebra and Logic”. He has classes (lectures and seminars) at the Chair of “Mathematical Logic and its Applications” at the FMI of the SU and at the New Bulgarian University. The scientific interests of Dimiter Dobrev are in the field of artificial

intelligence and logical programming, systems of the “Bonus-Malus” type. In addition, he is one of the authors of a project for the movement of subway trains without drivers and without stopping at all stations. Dimitar Dobrev was a member of the team for the transliteration and keyboard layouts adapted to the Bulgarian language. He is a developer of software and information products and has several registered patents as well as patent applications. Dimiter Dobrev has a well-maintained page on the Internet, from which one can obtain detailed information about his entire activity. In addition, he is the author of the book “Prophet Gocho”, for which the author promises to add new chapters, which is written in colorful language and for which there are many contraversal opinions. The laudatory go so far as to say that it is a modern version of “Bay Ganyo”. Dimiter Dobrev has also made a number of public appearances on television and other media. He is usually presented there as a writer and/or as an expert in artificial intelligence and very rarely as a mathematician. In his CV, Dimiter Dobrev presented a list of 21 scientific publications (for the period 2005 – 2022) and 11 popularising science publications (for 2000 – 2018). Of the scientific articles, 2 are in *Mathematica Balkanica*, 2 are in the *Serdica Journal of Computing*, 3 are in the *International Journal “Information Theories and Applications”*, 1 is in the *International Journal “Information Technologies and Knowledge”*, 5 are in proceedings of conferences in Bulgaria and Greece, 8 are in the arXiv and viXra preprint bases. (In the CV, article [19] is given as a preprint, but in the list of 20 articles related to the dissertation, the article appeared in 2021 in *Serdica Journal of Computing*.) Popularising science articles are: 8 in *PC Magazine – Bulgaria* and one in *eWeek Bulgaria, Culture and viXra*. I would recommend the applicant to better choose the journals where he publishes his works in order to get more recognition not only in our country but also abroad.

In conclusion, I will note that sometimes the behavior of Dimiter Dobrev is not typical for a mathematician, but nevertheless I have a positive opinion of him and consider him to be a useful member of the Bulgarian mathematical community.

3. Content analysis of the scientific and applied achievements of the candidate, contained in the presented Ph.D. thesis and the publications to it, included in the procedure

Topic of the dissertation. I would characterize the research presented in the dissertation as research in the field of artificial intelligence inspired by mathematical logic. The candidate has 4 articles referenced in *Zentralblatt für Mathematik*, which according to the *Mathematics Subject Classification – MSC2020* are in the section 68-XX Computer science, subsections 68Txx Artificial intelligence (68T01 General topics in artificial intelligence and 68T05 Learning and adaptive systems in artificial intelligence) and 68Wxx Algorithms in computer science (68W01 General topics in the theory of algorithms). I would also classify his publications under 68T27 Logic in artificial intelligence.

Actuality of the developed problems. I believe that important and current tasks are studied and considered in the dissertation. In confirmation of this, I will note that *Zentralblatt für Mathematik* has reviews for more than 115,000 publications in 68Txx Artificial intelligence, over 7,000 of them are books, and since 2020 there have been almost 13,000 publications, of which over 640 are books.

Scientific contributions. According to the applicant, the purpose of his dissertation is to dispel two misconceptions related to Artificial Intelligence (that AI is a memoryless function and that AI is pseudoscience) and to ask the three most important questions related to this topic: “What is AI?”, “How to create it?” and “What consequences will the creation of AI entail?”

In my opinion, formulated in this way, the statement about the two misconceptions is exaggerated. As for AI being a memoryless function for most researchers, we should state that simply AI theory has two subfields – memoryless AI (narrow but not at all fake AI as the thesis claims) and general AI or global AI. I am not convinced that the prevailing view is that AI is pseudoscience, as the thesis claims. I think AI, and global AI in particular, is not at all neglected by both society and

researchers. In confirmation, I will note two publications that appeared these days in a popular Bulgarian news page on the Internet. It commemorates the 70th anniversary of Alan Turing's death (June 7, 1954), where one of his main credits is said to be the father of AI. The other publication is an interview with Prof. Mariana Todorova and her participation in an international conference at the University of Information Science and Technology in Nanjing, China, where scientists from different countries will discuss precisely the future of humanity in the era of artificial intelligence in all spheres, including the possibilities of general artificial intelligence. I also disagree with the statement: "Even among mathematical logicians there is some skepticism about AI, but logic is the mathematical discipline closest to AI and is perfectly suitable for creating AI. One can say that mathematics is the basis of all exact sciences, while logic is the math of mathematics. Each mathematical discipline builds its fundament on logic." (For me, the claim about the importance of mathematical logic is inspired by the famous statement of Gauß: "Die Mathematik ist die Königin der Wissenschaften und die Zahlentheorie ist die Königin der Mathematik.") I would state the claim about the skepticism of logicians in a milder form: "There are mathematical logicians who are not convinced that AI is related to mathematical logic."

The first chapter of the dissertation is entitled "What is artificial intelligence?". The first rhetorical question the applicant asks is: "Do we have to know what is AI?" To this question he answered simply: *Yes, if we want to find it then our task will be a lot easier if we know what is the thing we are looking for.* Of course, this formulation is well known to researchers in all fields of knowledge. The first answer to the question of what AI is was given in the paper of Turing in 1950. The Turing test is based on a game called the "Imitation game". It is played with three people, a man, a woman, and an interrogator. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. The task of the man is to confuse the interrogator. The task of the woman is to help the interrogator make the right choice. Turing asked the question what would happen if the man were replaced by a machine. Will the interrogator be able to make the right choice and will the machine lie to him? The answer to this question indicates whether the machine can appear as thinking. In the thesis it is formulated as follows: *"If we can't make difference between the computer and a human being then it will be AI."* This definition is fine for AI without memory, but the applicant wants to define AI with memory. In the dissertation he accepts the thesis of Church that every calculating device can be modelled by a program, and gives the following informal definition: **AI will be such a program which in an arbitrary world will cope not worse than a human.** The condition that the program works successfully in an arbitrary world is essential. Usually, specific AI developments work in a specific world. For example, a few years ago a Turing test was created for financiers.

Jasmina Hasanhodzic, Andrew W. Lo, Emanuele Viola, Is It Real, or Is It Randomized?: A Financial Turing Test, arXiv:1002.4592v1 [q-fin.GN].

Experiments with this test show that a person is able to detect regularities in seemingly random events. In this case, this is verified with data from the financial markets, i.e. the development refers to a specific world.

The next step in the dissertation is to formalize the definition of AI. It is obvious to me that the formalization uses the idea of the Turing machine, but there are several fundamentally new moments in the construction, and technical difficulties have also been overcome in the considerations. In particular, the previously given definition of Hernández-Orallo and Minaya-Collado and its improvement by Marcus Hutter is clarified, because the definition of AI given in the thesis does not depend of the length of the life and of the language for describing the world. Different variants of AI are considered – when the world is deterministic and is structured simply enough, when it is deterministic but complex, and when there are elements of randomness. Algorithms of a strategy that

is close to the best strategy and that are most suitable for each of the cases are described. Several statements are formulated that are not proven, but arguments are given that they should be true.

The thesis notes that this definition has received a lot of acknowledgement because right now the first result returned by Google in response to a query for *Definition of Artificial Intelligence* is the paper of Dobrev I checked this. Typing the question in Google, the first two lines of the answer are:

“Scholarly articles for Definition of Artificial Intelligence

A definition of artificial intelligence – Dobrev – Cited by 131”.

The second chapter “How can we create AI?” reduces the task of creating AI to the task of finding the right language to describe the world. In contrast to programming languages that describe only computable functions, the thesis considers a language that describes a slightly wider class of functions. The idea is to create a new language to describe a particular world and then show that this language is suitable for describing an arbitrary world. The description of the world is seen as a structure consisting of many layers and which in the dissertation is represented as a pyramid, for which the first layer is the base. In particular, a generalization of the concept of algorithm is given. The discussions are illustrated with different variants of the game of chess – when the player plays against himself and when there are two players. As in the first chapter, the considerations here are formalized in a language that, in my opinion, is inspired by the considerations typical for mathematical logic and the theory of algorithms.

The third chapter, “What shall we do once we are done with creating AI?”, examines philosophical questions related to the consequences of creating AI. One of the questions asked at the end of the dissertation is “*Should AI technology be an Open Source Project?*” According to the applicant, it is a dangerous technology and serious AI papers should be classified. Other questions are also being discussed: *Can we not create AI? How do serious journals treat the subject? Who should have access to the computers? What will our life look like after the creation of AI?* and others of this kind. I disagree with many of the statements in this chapter of the dissertation, which I believe sound extremistic. For example, the applicant states that reviewing in paper-based magazines is dispensable because paper supplies are limited and they cannot afford publishing whatever comes by. Certainly, paper-based magazines are obsolete reminiscences of the past. Nevertheless, reviewers continue to have important roles even in e-zines, although not to say yes or no, but to evaluate the articles and advise readers whether an article is worth reading. That is, reviewers are needed, but in the capacity of critics rather than censors.

Another controversial claim in the thesis is that after the advent of AI, our life will appear deceitfully easy. We will not have to bother about our food or livelihood, will not have to work and even will not have to entertain each other because AI will deliver entertainment much better than any human entertainer could do. *I think a moderately intelligent reader will easily separate the useful stuff from the extremist stuff.* Personally, I am a supporter of Turing's view. By creating his test to see if you are a machine or a human, on which robotic experiments are still based today, Turing laid down the rules and was the first to say that computers cannot and should not make humans useless, only be in their help. It is a humane look at the future that inspires hope that despite progress, the machine will never be human.

I particularly want to point out the unconventional style used in the dissertation. The explanations are written in a style accessible to a general (including non-mathematical) audience, but which is far from the standard style for writing a paper in the field of mathematics. This way of presentation is useful for promoting achievements, but in my opinion, it is not the best way to write a thesis in the field of mathematics, computer science and their applications.

Another characteristic feature of the dissertation is that it does not contain a single essential mathematical statement that has been proven. For many other sciences, having a discussion on a

problem can be important to the development of science. To a lesser extent, this is also true for mathematics. For example, according to the general mathematical audience, one of the most important achievements of the Soviet mathematician Andrey Tikhonov was his 1926 definition of a product of topological spaces. Of course, I do not want to compare the applicant with Tikhonov, but I find that the issues addressed in the dissertation are important and the discussion in it is in a useful direction.

Application of the obtained results. The publications included in or related with the dissertation have echoed in the publications of other researchers. The documentation includes a list of 119 citations in monographs, articles, dissertations and theses in English, German, Italian, Spanish, Portuguese, Slovenian, Croatian, Slovak, Finnish and Turkish.

It is obvious to me that the applicant knows very well the results in the field of his research. In confirmation of the non-traditional style of exposition, I will note that the dissertation also cites materials, the presence of which is a bit strange in a mathematics dissertation. Isaac Asimov, The Terminator, The Matrix, President Macron, Pavlov's Dog and other of this kind are quoted.

4. Approbation of the results

As I already noted, the dissertation is based on 5 articles published in 2000 – 2023, and another 20 articles are directly or indirectly related to the dissertation. Two of the five articles were published in an respectable Bulgarian journal (with an impressive volume of 50 and 38 pages), one was published in the International Journal "Information Content and Processing", one was popular, one was in the preprint base. I have only indirect data on the presentation of results at seminars and conferences (from the list of publications in conference proceedings and from the meetings of the Algebra and Logic seminar).

My inspection shows that:

- a) The scientific works meet the minimum national requirements of the Act on Development of the Academic Staff in the Republic of Bulgaria and respectively the additional requirements of Institute of Mathematics and Informatics for obtaining the educational and scientific degree “Doctor” (Ph.D.) in the scientific field and professional direction of the procedure. Moreover, with minimum requirements of 30 points in one of the groups of indicators, the applicant has submitted evidence for 36 points.
- b) The results presented by the applicant in the dissertation and scientific works to it do not repeat such from previous procedures for acquiring a scientific title and academic position.
- c) There is no plagiarism proven in the legally established order in the submitted dissertation work and scientific papers under this procedure.

5. Qualities of the abstract

The abstract is quite detailed (37 pages) and is in the style of the dissertation. I think that it gives a clear and adequate idea of the content and main results of the dissertation. The statement of contributions accurately reflects the contributions of the applicant. I think that the abstract and the summary of the contributions meet all the requirements for their preparation.

6. Critical notes and recommendations

I have no particular criticisms other than those noted in the previous points of my review. I would like to especially point out that the things I noted do not spoil my overall positive impression of the results and their presentation in the dissertation.

7. Conclusion

Having become acquainted with the Ph.D. thesis presented in the procedure and the accompanying scientific papers and on the basis of the analysis of their importance and the scientific and applied contributions contained therein, **I confirm** that the presented Ph.D. thesis and the scientific publications to it, as well as the quality and originality of the results and achievements presented in

them, completely satisfy the requirements of the law and the additional requirements of the BAS and IMI – BAS for a Ph.D. thesis in the area of mathematics. In particular, the candidate satisfies the minimum national requirements in the professional field and the additional requirements of IMI - BAS, and no plagiarism has been found in the scientific works submitted for the competition.

Based on the above, I **strongly recommend** to the respectable Scientific Jury to award Dimiter Dimitrov Dobrev with the educational and scientific degree “Doctor” (Ph.D.) in the Scientific Field 4. Natural Sciences, Mathematics and Informatics, Professional Field 4.5. Mathematics, (Doctoral Program: “Mathematical Logic”).

Date: June 11, 2024

Signature:

(Vesselin Drensky, Prof. D.Sci., Full Member of the BAS)