

Programming methodology for trusted artificial intelligence

Goncharov S.^{1†*}, Nechesov A.^{1†**}, Sviridenko D.^{2***}

¹ S.L. Sobolev Institute of Mathematics, SB RAS, Novosibirsk, Russia

² Institute of Philosophy and Law, SB RAS, Novosibirsk, Russia

*s.s.goncharov@math.nsc.ru; ** nechesov@math.nsc.ru; ***dsviridenko47@gmail.com

Key words: trusted AI; programming methodology; learning theory

The task of building reliable, transparent and explainable artificial intelligence attracts the attention of all the world's researchers. Today it becomes clear that all advanced large language models do not satisfy even the basic requirements put forward for trusted intelligent systems.

The report will touch upon the development of trusted artificial intelligence algorithms [1]. A programming methodology in Turing complete languages will be presented, which is based on the task approach and the concept of semantic programming [2–5].

Our task is to make the thinking process of intelligent machines transparent and human-understandable. For these purposes, it is also necessary to build a coherent and high-quality concept of learning theory [6]. This concept should allow, based on logical rules and basic statements, to show the user the final result produced by the intelligent system. For these purposes, a new logical-probabilistic method will be presented, which allows us to close all these gaps.

Funding: The work was performed within the state task of the S.L. Sobolev Institute of Mathematics (project No. FWNF-2022-0011).

References

1. Vityaev E.E., Goncharov S.S., Gumirov V.S., Mantsivoda A.V., Nechesov A.V., Sviridenko D.I. Task approach: on the way to trusting artificial intelligence. In: World Congress Systems Theory, Algebraic Biology, Artificial Intelligence: Mathematical Foundations and Applications. Selected Works. 2023;171-243. doi 10.18699/sblai2023-41
2. Goncharov S., Nechesov A. Polynomial Analogue of Gandy's Fixed Point Theorem. *Mathematics*. 2021;9:2102. doi 10.3390/math9172102
3. Goncharov S., Sviridenko D. Logical Language of Description of Polynomial Computing. *Dokl Math*. 2019;99:121-124. doi 10.1134/S1064562419020030
4. Goncharov S., Nechesov A. Solution of the Problem $P = L$. *Mathematics*. 2022;10:113. doi 10.3390/math10010113
5. Goncharov S., Nechesov A. Semantic programming for AI and Robotics. In: IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), Yekaterinburg, Russian Federation. 2022;810-815. doi 10.1109/SIBIRCON56155.2022.10017077
6. Goncharov S., Nechesov A. AI-Driven Digital Twins for Smart Cities. *Eng Proc*. 2023;58:94. doi 10.3390/ecsa-10-16223