

HOW AND WHY TO OVERCOME THE BARRIERS BETWEEN US AND AGI

Smolin V.S., Shen H.¹

KIAM RAS, 125047, Moscow, Miuskaya sq., 4, +7 499 978-13-14, smolin@keldysh.ru

¹MSU, Faculty of Space Research, 119991, Moscow, GSP-1, 1-52, Leninskiye Gory,
+7 903 170-86-66, haoyuishen@yandex.ru

Since about 2017, the topic of AGI has gone from the distant realm of science fiction to a practical goal [1]. Today, the creation of AGI has become a primary objective of AI research at prominent companies such as OpenAI, DeepMind, and Anthropic PBC. However, the scientific community at large still believes that machines can never replace humans with souls, feelings, empathy, and most importantly, trustworthiness.

However, there is another extreme view of AGI: some argue that “intelligent” machines could supplant humans, especially as generative AI [2] now creates human-level text and realistic images, potentially surpassing human abilities over time.

The challenge is to choose a balanced, mathematically based approach to assessing the capabilities of an already created AI. This requires overcoming the established stereotypes that serve as barriers to AGI creation. It should also be understood that, like any human invention, AGI will not be useful or harmful in itself, but in accordance with its goals and capabilities.

Overcoming barriers, namely such ideas as a) not everything can be automated; b) only a person can gain new knowledge; c) a machine will never be able to feel, understand, or set tasks consistent with human needs [3]; d) be more energy-efficient than a person when solving cognitive problems and many others should be based on demonstrating the algorithmicity of all previously seemingly unsolvable tasks using formal decomposition methods.

The research findings suggest that AGI is feasible and necessary, as human limitations currently constrain civilization’s cognitive development. However, AGI could either accelerate progress or deepen inequality, depending on the social context and accepted notions of fairness.

References

1. *Burtsev M.S., Bukhvalov O.L., Vedyakhin A.A. et al. Strong Artificial Intelligence: On the Approaches to Superintelligence. - Moscow: Knowledge Literature Publishing House, 2021. - 232 c.*
2. *Foster D. Generative deep learning: how we don't paint pictures, write novels and music / 2nd ed. - Sprint Book, 2024. - 448 c.*
3. *Russell S. Human compatible: Artificial intelligence and the problem of control. - Penguin, 2019. – 336 c.*