

Социальные аспекты создания общего искусственного интеллекта

Philip Matveev Mikhailovich

Студент (бакалавр)

Московский государственный университет имени М.В.Ломоносова, Философский факультет, Кафедра социальной философии, Москва, Россия

E-mail: mphm@inbox.ru

Большинство исследований AI, начиная с 1970-х годов, нацелены не на создание AI человеческого уровня, а на применение теорий и методов AI к конкретным проблемам [Russell, Norvig 2021]. Так, стремительный прогресс в современном AI связан с результатами, полученными благодаря применению методов обучения нейронных сетей (прежде всего, в области компьютерного зрения, распознавания речи, навигации и др.). При этом создатели ИИ как области, исследователи 1950-х годов, действительно думали об интеллекте человеческого уровня [McCarthy 2008]. Алан Тьюринг, пионер ИИ, первым подчеркнул, что ИИ будет реализован с помощью компьютерных программ. Сейчас интерес к ИИ человеческого уровня и методам его достижения больше, чем за последние 40 лет.

С нашей точки зрения, философские вопросы актуальны для ИИ именно в связи с проблемой создания общего ИИ (*Artificial General Intelligence*), то есть искусственного интеллектуального агента, способного действовать в динамической среде, решая тот же спектр задач, который способен решать человек [Efimov 2021].

В этом докладе мы попытаемся (1) показать, что роль философии важна именно в исследованиях общего, а не «узкого» ИИ, (2) привести аргументы в пользу невозможности создания общего ИИ методами, ответственными за текущий прогресс в «узком» ИИ, а также (3) предложить гипотезу о том, как достичь результатов в функциональной области, которую, условно можно назвать кластером социально-эмоциональных способностей искусственного агента.

Источники и литература

- 1) McCarthy, John. (2008). The Philosophy of AI and the AI of Philosophy. 10.1016/B978-0-444-51726-5.50022-4.
- 2) Efimov, Albert and others. (2021). Artificial General Intelligence: on approaches to the supermind. Intellectualnaya Litteratura.
- 3) Russell, Stuart, Norvig, Peter (2021) Artificial Intelligence: A Modern Approach, 4th US ed., Pearson, Hoboken.

Иллюстрации

Narrow AI vs. AGI

NARROW AI	ARTIFICIAL GENERAL INTELLIGENCE (AGI)
<ul style="list-style-type: none">■ AI focused on a specific, singular or limited task■ Examples include image recognition, hyper-personalization, chatbots, predictive text■ Trained on specific tasks by data scientists■ Correlates questions or assignments to a specific data set to accomplish a task■ No self-awareness, consciousness, ability to think	<ul style="list-style-type: none">■ Not fully realized, with some developers questioning if it will be possible■ Seeks machines that can handle a range of cognitive tasks with little oversight■ The ability to learn, generalize, apply knowledge and plan for the future■ Must consistently pass the Turing Test■ Single, general intelligence that possesses common sense and creativity and expresses emotions


©2020 TECHTARGET. ALL RIGHTS RESERVED 

Рис. 1. Общая информация об "узком" и общем ИИ.

Current technologies have achieved different levels of human performance across 18 capabilities

■ Below median ■ Median ■ Top quartile

	Automation capability	Capability level ¹	Description (ability to ...)
Sensory perception	Sensory perception		Autonomously infer and integrate complex external perception using sensors
Cognitive capabilities	Recognizing known patterns/categories (supervised learning)		Recognize simple/complex known patterns and categories other than sensory perception
	Generating novel patterns/categories		Create and recognize new patterns/categories (e.g., hypothesized categories)
	Logical reasoning/ problem solving		Solve problems in an organized way using contextual information and increasingly complex input variables other than optimization and planning
	Optimization and planning		Optimize and plan for objective outcomes across various constraints
	Creativity		Create diverse and novel ideas, or novel combinations of ideas
	Information retrieval		Search and retrieve information from a large scale of sources (breadth, depth, and degree of integration)
	Coordination with multiple agents		Interact with others, including humans, to coordinate group activity
	Output articulation/presentation		Deliver outputs/visualizations across a variety of mediums other than natural language
Natural language processing	Natural language generation		Deliver messages in natural language, including nuanced human interaction and some quasi language (e.g., gestures)
	Natural language understanding		Comprehend language, including nuanced human interaction
Social and emotional capabilities	Social and emotional sensing		Identify social and emotional state
	Social and emotional reasoning		Accurately draw conclusions about social and emotional state, and determine appropriate response/action
	Social and emotional output		Produce emotionally appropriate output (e.g., speech, body language)
Physical capabilities	Fine motor skills/dexterity		Manipulate objects with dexterity and sensitivity
	Gross motor skills		Move objects with multidimensional motor skills
	Navigation		Autonomously navigate in various environments
	Mobility		Move within and across various environments and terrain

¹ Assumes technical capabilities demonstrated in commercial products, R&D, and academic settings; compared against human performance.

Рис. 2. Прогресс в разных функциональных областях, согласно отчёту McKinsey от 2017 года.