ARTIFICIAL INTELLIGENCE vs ARTIFICIAL HUMAN

Mykola Rabchevskiy

HIDDEN BELIEF IN ANTHROPOMORPHISM

The focus of the public interested in the developments and prospects of **AGI** and/or **strong AI** is security issues directly or indirectly related to the behavior of such systems. In this regard, there are discussions about ethics, empathy, the civil rights of artificial intelligence, the copyright for discoveries and inventions made by AI, and so on.

A common feature of most (if not all) such discussions is the ***implied anthropomorphism*** of future systems in the sense that future systems will have a set of feelings, ethical principles, and social relationships. This assumption looks like a matter of course against the background of the introduction of voice assistants, computer translators, emotion detectors in text and images, and the construction of anthropomorphic (in terms of appearance) robots.

AGI engineering is a reader-supported publication. To receive new posts and support my work, consider becoming a free or paid subscriber.

A detailed analysis of the intellectual abilities of such existing systems tagged "**AI**" reveals that they do not yet have intelligence in the original sense laid down by the creators of the term AI. Sometimes there is an outward resemblance to a person, sometimes an imitation of meaningful dialogue, sometimes an imitation of emotions, but the level of abilities comparable to a human has not yet been reached. The presence of systems capable of performing individual intellectual functions (arithmetic calculations, finding the optimal path from one point to another, etc.) has not yet led to the creation of systems capable of combining all human intellectual abilities in one system.

To analyze ***potential opportunities*** - as well as ***future problems and dangers*** - it is crucial to examine the factors that shape the specificity of a human on the one hand and potential AI systems on the other hand. The results of such an analysis do not always correspond to "obvious" assumptions.

FOR WHAT?

The creation of artificial systems is an engineering business. Any significant development must be helpful, and the degree of usefulness must be commensurate with the costs. And the first question that arises when creating this or that device is "for what?". If tomorrow someone proposes to make a copy of the Egyptian pyramid with a kilometer in size - which is feasible in principle - then the first question will be, "what will the corresponding funds be spent for?". When the first attempts to build an airplane took place, the **goal was not to copy birds** but to create an apparatus for transporting people and goods through the air; **copying some principles** that allow birds to fly was a **tool, not a goal**.

Creating an exact full-functional copy of a human by technical means seems pointless. The natural intelligence of a person requires that any developed intelligent system **surpasses a human in some aspects; otherwise, there is no point in creating it**. Computers, for example, were designed to speed up calculations and logical analyses that a human can do but does slowly and with frequent errors. The presence of advantages over a person in some aspects is necessary, and the presence of advantages automatically means that **all such systems will differ from a person**. That is, we should not talk about a functional copy of a human but about a system with intelligence, which in some way surpasses a human and **differs from a human**.

FUNDAMENTAL DIFFERENCES

Since any proper AI system, as we found out, must necessarily be different from a human, the question arises of what exactly. In some aspects of AI, the system should be superior to humans, but at the same time, there are **differences dictated by technology**. The main difference, dictated by modern technology, is that **human-made devices cannot self-evolve from the embryo** to "full size," nor are they able to **reproduce**. The second radical difference is that the **information in a technical system can be copied from one system to another**, while biological systems are deprived of such an opportunity. The implications of these two differences are significant.

The function of supporting reproduction requires means that form the appropriate behavior. This requires proper emotions and ethical rules as a kind of superstructure on emotions. AI-based systems are obviously devoid of corresponding emotions and do not need ethical regulation in this area.

The lack of an effective possibility of copying knowledge from one individual to another leads to the development of social methods of education that form a certain amount of knowledge, skills, and habits, including basic knowledge about the world, communication skills, ethical principles, ways of getting pleasure and professional knowledge. As one result, personal

differences are unavoidable - unlike technical systems, which can differ only by serial number.

CONSEQUENCES

As a natural intellectual system, humans have a set of features that are either impossible to implement with modern technology or unnecessary. On the other hand, technical systems must have capabilities that humans do not have to be helpful: creating an AI system aims to provide superiority in some aspects over human capabilities.

Any AGI-equipped system will obviously differ from a human in many ways and in each individual case differently depending on the purpose.

The purpose of an AI-controlled system dictates a **set** of real or virtual **sensors**. The set of sensors affects the collection of **primary emotions**. This, in turn, affects the **behavior** of the system. Accordingly, the set of emotions differs, requiring corresponding differences in the **ethical sphere**.

The implied anthropomorphism leads to ignorance of differences and contributes to an inadequate understanding of AI systems' opportunities, problems, and dangers.

As an example of the consequences of ignoring differences, we can cite a recent publication On the Measure of Intelligence by *Francois Chollet*.

The detailed and reasoned analysis concludes with a chapter that proposes a set of tests for detecting/assessing the level of intelligence applicable to both humans and AI systems. The collection includes tests that are feasible for most people: the subject (human or AI system) must find dependencies in the configuration of images of such kind:
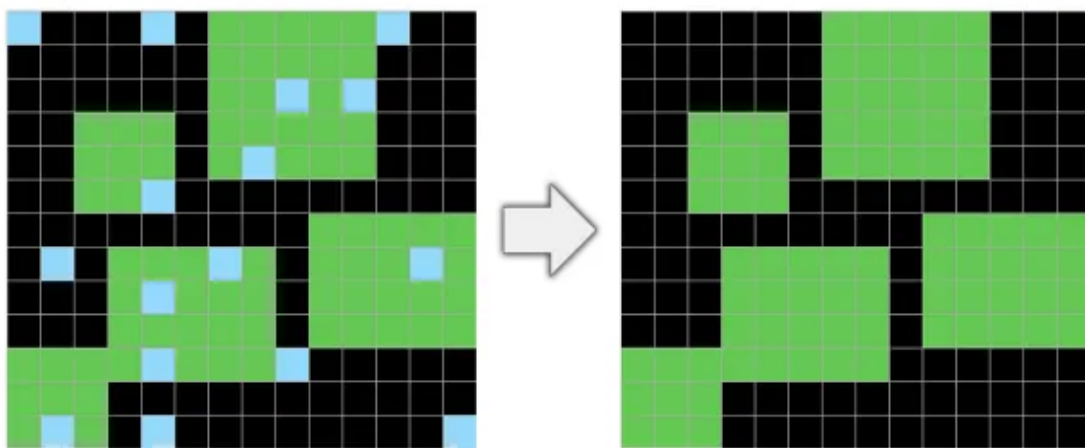
Implicit adherence to the anthropomorphism hypothesis gives the impression that tests allow you to compare the level of human intelligence with that of artificial systems: if people pass

the tests, then the ability or inability to cope with the tests depends only on the level of intelligence, which is required from the tests. However, the consequence of the fact that the AI system is obviously different from a person is the fallacy of the assumption that the test results depend only on the level of intelligence.

The authors declare the need for prior knowledge ("**Object cohesion**" and "**Object persistence**," and assume that there is no need to have other prior abilities.

But suppose we offer these tests that are quite feasible for people, giving descriptions of tests **not in the form of pictures,** but in the form as they are provided by the authors in Github, **in the form of a text description** (enumeration of picture cells with coordinates and colors, which does not require knowledge of JASON grammar). In that case, it turns out that **most people will fail the tests despite the apparent presence of the required knowledge**. This means that to solve problems successfully, a person needs a **specific set of skills that allows one to associate visual information with the corresponding topological concepts**.

If the AI system under test satisfies the hypothesis of anthropomorphism, this does not interfere with the application of tests. But since the AI system is necessarily different from a human, the test result becomes dependent on whether the AI has a sensory system identical to a human. But there is no such identity, so the test result depends not only on the availability of the declared prior knowledge. Ultimately, the test does not correspond to the stated goal: it tests the **ability of the combination of the intellect and the sensory system** and not the intelligence in isolation from everything else.

This sort of error can be avoided by eliminating the implicit use of the assumption that AI systems are anthropomorphic.

AGI engineering is a reader-supported publication. To receive new posts and support my work, consider becoming a free or paid subscriber.