Qeios

Peer Review

Review of: "No Consciousness? No Meaning (and no AGI)!"

Alexandru Tăbușcă[1]

1. Romanian-American University, Romania

This paper tackles one of the most fundamental questions in artificial intelligence: Can machines truly understand meaning, or are they just sophisticated pattern-matching systems? The author argues that despite the impressive capabilities of modern AI systems like ChatGPT, there's a crucial missing ingredient that prevents them from achieving genuine understanding—consciousness. The central thesis is bold and clear: without conscious experience, no system can truly grasp meaning, and therefore, true Artificial General Intelligence (AGI) that matches human-level understanding is impossible without first solving the problem of machine consciousness.

The Gap Between Symbols and Meaning

The author builds on classic philosophical arguments, particularly John Searle's famous "Chinese Room" thought experiment from 1980. Imagine someone locked in a room with a rulebook for manipulating Chinese characters. They can follow the rules perfectly and produce appropriate Chinese responses to Chinese inputs, but they don't actually understand Chinese—they're just manipulating symbols according to rules. This, the author argues, is exactly what modern AI does, no matter how sophisticated. Large Language Models (LLMs) process symbols (words, tokens) according to mathematical rules, but they don't actually understand what those symbols mean. The paper's most provocative claim is that meaning-making requires conscious experience—what philosophers call "qualia." These are the subjective, felt qualities of experience: the redness of red, the pain of a toothache, the taste of chocolate. The author argues that without these conscious experiences, symbols remain just symbols, never connecting to genuine understanding. For example, an AI might process millions of descriptions about apples, but without ever experiencing the taste, texture, or visual appearance of an apple, it can't truly understand what "apple" means in the way humans do.

Evidence and Analysis

doi.org/10.32388/95S7GL

The paper presents compelling evidence of AI limitations:

Current AI Struggles With:

- Simple counting tasks (like counting words in a sentence)
- Negation (understanding "no elephant in the room")
- Mathematical reasoning beyond pattern matching
- Common sense that differs from training data
- Distinguishing between people and billboards (self-driving cars)

These Failures Reveal:

- AI relies on probabilistic pattern matching, not genuine understanding
- Performance drops dramatically when problems differ from training examples
- AI cannot correct its own reasoning errors like humans can

Some researchers argue that giving AI bodies and sensory experiences (embodiment) could solve the meaning problem. The author acknowledges this might help but argues it's insufficient. Even with sensors and motors, an AI without consciousness would still just be processing more complex patterns—now including sensory data rather than just text.

Strengths of the Paper

- Philosophical Rigor

The paper demonstrates impressive breadth, connecting ideas from linguistics (Wittgenstein, Chomsky), information theory (Shannon), and consciousness studies. The author skillfully weaves together technical and philosophical perspectives.

- Relevance

With the recent explosion of interest in AI capabilities and AGI predictions, this paper addresses crucial questions about what we're actually building and what limitations we might be overlooking.

The paper uses accessible examples (Chinese Room, apple tasting, visual recognition) to illustrate complex philosophical concepts, and the author cites extensive recent research showing specific limitations in current AI systems, from mathematical reasoning to basic counting tasks.

While the paper's central argument depends on consciousness being necessary for meaning, it doesn't adequately address how consciousness arises in biological systems. If we don't understand how brains generate consciousness, how can we be certain it's impossible in machines?

The paper could better explore whether there might be forms of understanding that don't require human-like consciousness. Could machines develop their own type of semantic awareness that we don't recognize?

Practical Implications

The paper focuses heavily on philosophical arguments but offers limited guidance on how these insights might inform AI development practices or research directions. While the author suggests these ideas could lead to better tests of intelligence than the Turing Test, the paper doesn't provide concrete, implementable testing frameworks.

Broader Implications

- **For AI Development**

If the author is correct, the current approach of scaling up AI systems (more data, more parameters, more computing power) will hit a fundamental ceiling. True AGI would require breakthroughs in understanding and creating consciousness—a much harder problem than currently anticipated.

- **For Consciousness Research**

The paper suggests that studying AI limitations might provide new insights into human consciousness and the nature of understanding itself.

- **For Society**

The implications are profound: if true AGI requires consciousness, then current AI systems, no matter how impressive, are sophisticated tools rather than minds. This affects how we should think about AI rights, capabilities, and societal impact.

Critical Assessment

The Strength of the Argument

The paper makes a compelling case that current AI systems lack genuine understanding. The evidence of AI failures in seemingly simple tasks is particularly convincing. The connection between meaning and

conscious experience, while debatable, is well-argued and draws on substantial philosophical tradition.

Potential Weaknesses

1. **Definitional Issues**: The paper doesn't provide a precise definition of consciousness, making it difficult to evaluate when the threshold would be crossed.
2. **Biological Chauvinism**: There's a risk of assuming that only biological consciousness counts, potentially missing other forms of awareness.
3. **Practical vs. Philosophical Understanding**: For many practical purposes, current AI systems demonstrate understanding-like behavior. The paper could better address whether philosophical understanding is necessary for practical applications.

Scientific Rigor

While philosophically sophisticated, the paper sometimes conflates empirical claims with philosophical arguments. Some conclusions about AI limitations might reflect current technical constraints rather than fundamental impossibilities.

Why This Matters

This paper arrives at a crucial moment when AI capabilities are advancing rapidly, and predictions about the AGI timeline are becoming increasingly optimistic. If the author is even partially correct, we may be overestimating how close we are to human-level AI. The paper also highlights important questions about what we mean by "intelligence" and "understanding." As AI systems become more capable, distinguishing between sophisticated pattern matching and genuine comprehension becomes increasingly important.

Future Directions

The paper opens several important research questions:

1. **Developing better tests** for genuine understanding versus sophisticated mimicry
2. **Investigating the relationship** between consciousness and intelligence in biological systems
3. **Exploring whether** machine consciousness is theoretically possible
4. **Considering the implications** for AI safety and alignment if current systems lack genuine understanding.

This is a thought-provoking and important paper that challenges prevalent assumptions about AI progress. While not all arguments are equally convincing, the author succeeds in highlighting fundamental questions that the AI community must address.

The paper's central insight—that there may be a qualitative difference between pattern matching and understanding—deserves serious consideration. Whether or not consciousness is the key, the evidence suggests current AI systems have significant limitations that scaling alone may not overcome.

Strengths: Philosophically sophisticated, timely, well-evidenced critique of current AI limitations

Weaknesses: Some definitional ambiguity, limited practical guidance, potential biological bias

Overall Assessment: Important contribution to AI philosophy that should influence how we think about intelligence, understanding, and the path to AGI

This paper won't settle the debate about AI consciousness, but it provides a valuable framework for thinking about these crucial questions as AI continues to advance. For anyone interested in the future of artificial intelligence, the nature of mind, or the relationship between consciousness and understanding, this paper offers essential reading.

Even if the paper is actually not a “technical” one, related to the broader AI field, it is a very interesting and provocative product that should be read carefully by programmers, developers, and software engineers as well.

Declarations

Potential competing interests: No potential competing interests to declare.