Agentic AI for Scientific Research

Olatomiwa O. Bifarin

School of Chemistry and Biochemistry
Georgia Institute of Technology

August 17, 2024

Abstract

Note: This manuscript is very much a work in progress, with currently mostly outlines. Many of the papers I will be reviewing here, I have reviewed on my AI blog https://theepsilon.substack.com/ This paper will allow me to congeal the reviews and articulate my thoughts for a more academic audience. The advent of Large Language Model (LLM) based AI agents is revolutionizing scientific research, offering (semi) autonomous systems capable of accelerating research across diverse fields. This paper will double as a tutorial and review that examines the current landscape of AI agents in scientific research, exploring their rapidly evoloving architecture, capabilities, applications, and future prospects. The review begins with the foundational elements of these agents, LLMs and key techniques like prompt engineering, Retrieval-Augmented Generation (RAG) and tool use. The review then dissects the anatomy of AI agents, elucidating the interplay between perception, memory, reasoning, and action modules that enable autonomous functionality. Both single-agent and multi-agent systems are presented, with illustrative examples from chemistry, biology, and data science. The discussion encompasses varying levels of agent autonomy, from research assistants automating routine tasks to envisioned future agents capable of independent scientific breakthroughs. Critical challenges and opportunities are addressed, including safety concerns, and the imperative for seamless integration within the scientific community. This review highlights the transformative potential of agentic AI in accelerating scientific progress and heralds a new era of human-AI synergy in research endeavors.

1 Introduction

1.1 The Evolution of AI in Science

The evolution of artificial intelligence (AI) in scientific research marks a significant transformation in the landscape of scientific discovery [1]. Initially serving as auxiliary tools, AI systems have progressed to become autonomous agents capable of executing complex scientific tasks with minimal human oversight [2]. This shift represents a paradigm change in research methodologies across various scientific disciplines [3, 4, 5]. These systems, while powerful, often necessitated substantial human expertise for their development, training, and result interpretation, which constrained their broader applicability and impact.

1.2 Foundations of LLM-Based Agents

- **Foundational Models:** Introduce the concept of foundational models, particularly LLMs, as the core of many modern AI agents.
- **Key LLM Techniques:** Briefly explain the essential techniques that enable LLMs to function effectively in scientific domains:
  - **Prompt Engineering:** Crafting instructions (prompts) that guide LLMs to perform specific tasks. Mention various prompting strategies (few-shot learning, Chain-of-Thought) and their impact on performance.
  - **Retrieval-Augmented Generation (RAG):** Combining LLM generation with information retrieval from external knowledge sources (e.g., databases, scientific literature) to improve factual grounding and reduce hallucinations.

1.3 The Promise of AI Agents

- State the purpose of the review: To provide a comprehensive overview of current research on AI agents in scientific research, focusing on their capabilities, challenges, and future directions.

2 The Anatomy of an AI Agent

2.1 What Defines an AI Agent?

- Define AI agents as autonomous systems capable of perceiving their environment, reasoning, making decisions, and taking actions to achieve specific goals.
- Contrast AI agents with traditional AI tools, highlighting their greater autonomy, adaptability, and potential for complex problem-solving.

2.2 Key Agent Modules

- **Perception:**
  - Explain how agents acquire information from their environment, including multi-modal data streams (e.g., text, images, experimental data).
  - Discuss different perception strategies, such as natural language understanding, visual perception modules, and cross-modal alignment.
  - Provide examples of perception modules used in scientific agents (e.g., ChemCrow's multi-modal data processing, BioPlanner's protocol analysis).
- **Memory:**
  - Discuss how agents store and recall information, differentiating between short-term memory for immediate context and long-term memory for knowledge retention.

  - Explain different memory implementations, including internal memory encoded in LLM weights and external memory in the form of knowledge bases (e.g., databases, literature).
  - Provide examples of memory modules in scientific agents (e.g., PaperQA's literature retrieval for factual grounding, Coscientist's use of experimental data for decision-making).

- **Reasoning:**
  - Describe how agents make decisions and plan actions based on the information they perceive and retrieve from memory.
  - Highlight different reasoning techniques, such as Chain-of-Thought reasoning, multi-path reasoning, and planning strategies (e.g., Least-to-Most prompting, Tree-of-Thoughts).
  - Provide examples of reasoning modules in scientific agents (e.g., ChemCrow's reasoning for choosing synthesis pathways, STORM's question-asking strategy, TAIS's analysis of gene expression data).

- **Action:**
  - Explain how agents interact with their environment, taking actions to achieve their goals.
  - Discuss different types of actions, such as interacting with humans (e.g., through dialogue, feedback), using tools (e.g., code execution, database queries), and manipulating physical environments (e.g., robotic actions in self-driving labs).
  - Provide examples of action modules in scientific agents (e.g., Coscientist's interaction with robotic liquid handlers, ChemCrow's execution of chemical syntheses).

3 Architectures and Applications of Scientific AI Agents

3.1 Single-Agent Systems

- Introduce single-agent systems where one LLM is programmed with multiple roles to address different aspects of a scientific task.
- Provide detailed examples from the literature, categorizing them by scientific domain:
  - **Chemistry:** Discuss agents like ChemCrow, Chemist-X, Coscientist, and ChatMOF, analyzing their specific functionalities, strengths, and limitations.
  - **Biology:** Explore agents such as BioPlanner, CRISPR-GPT, and ProtAgent, examining their capabilities in protocol planning, gene editing, and protein design.
  - **Data Science:** Discuss agents like AutoBa and LLaMP, highlighting their roles in automating bioinformatics analyses and enhancing scientific feedback generation.

3.2 Multi-Agent Systems

- Introduce the concept of multi-agent systems, where multiple AI agents with specialized roles collaborate to solve complex scientific problems.
- Discuss the advantages of multi-agent systems over single-agent systems, emphasizing their greater flexibility, adaptability, and potential for interdisciplinary research.
- Provide detailed examples from the literature, focusing on how different agents collaborate and interact:
  - **TAIS (Team of AI-made Scientists):** Analyze how TAIS agents (Project Manager, Data Engineer, Domain Expert, Statistician, Code Reviewer) work together to identify disease-predictive genes from gene expression data (Tu et al., 2023).
  - **ProtAgent:** Discuss the collaboration between Planner, Assistant, and Critic agents to automate protein design tasks (Alipanahi et al., 2015; Dauparas et al., 2022; Watson et al., 2023).
  - **Other Multi-Agent Systems:** Explore examples from other domains (e.g., self-driving lab agents, as in Sanders et al., 2023), highlighting their potential applicability to chemistry and biology.

3.3 Levels of Autonomy in Scientific AI Agents

- **Research Assistant:** Describe agents that execute predefined tasks specified by scientists, providing support and automating routine procedures. (e.g. ChemCrow, AutoBa).
- **Collaborator:** Discuss agents that can actively participate in the scientific process, refining hypotheses, suggesting experiments, and interpreting results alongside scientists.
- **Scientist:** Explore the potential for future agents that can independently formulate hypotheses, design experiments, and make scientific discoveries, though this level of autonomy remains largely aspirational.
- Analyze the current state of autonomy in existing scientific agents, discussing the challenges and opportunities for achieving higher levels of autonomy.

4 Challenges and Opportunities for Scientific AI Agents

4.1 Ensuring Safety and Responsible Use

- Discuss the potential risks associated with autonomous AI agents in science, particularly concerning the synthesis of hazardous substances, bias in scientific discovery, and overreliance on AI.
- Highlight mitigation strategies, such as incorporating safety checks, ethical guidelines, human oversight, and robust evaluation protocols (e.g., ChemCrow's safety checks, TAIS's emphasis on human validation).

4.2 Improving Explainability and Transparency

- Underscore the need for more transparent and interpretable AI agents, making their reasoning process understandable to scientists and fostering trust in their outputs.
- Discuss strategies for improving explainability, such as integrating XAI techniques (e.g., XpertAI), providing detailed justifications for actions (e.g., WikiCrow's citations), and enabling interactive explanations that address user queries.

4.3 Enhancing Generalization and Adaptability

- Discuss the challenges of creating AI agents that can generalize to new scientific tasks, adapt to evolving data and knowledge, and overcome domain-specific limitations.
- Highlight potential strategies, such as transfer learning, continual learning, and the development of more flexible and adaptable agent architectures.

4.4 Fostering Collaboration and Integration

- Emphasize the importance of collaboration between AI researchers and scientists, ensuring that AI agents are aligned with scientific goals and values.
- Discuss the need for greater integration between AI agents and existing scientific tools and platforms, creating seamless workflows that enhance research productivity and innovation.

5 Conclusion

- Summarize the key findings of the review, highlighting the transformative potential of AI agents for scientific research across diverse disciplines.
- Reiterate the critical challenges that need to be addressed to ensure the safe, reliable, and beneficial use of AI agents in science.
- Offer a forward-looking perspective on the future of this rapidly evolving field, emphasizing the potential for greater automation, increased discovery rates, and a new era of human-AI collaboration in scientific research.

Acknowledgements

Acknowledgements will go here.

References

[1] Hanchen Wang, Tianfan Fu, Yuanqi Du, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, et al. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, 2023.

[2] Daniil A Boiko, Robert MacKnight, Ben Kline, and Gabe Gomes. Autonomous chemical research with large language models. *Nature*, 624(7992):570–578, 2023.

[3] Jessica Vamathevan, Dominic Clark, Paul Czodrowski, Ian Dunham, Edgardo Ferran, George Lee, Bin Li, Anant Madabhushi, Parantu Shah, Michaela Spitzer, et al. Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery*, 18(6):463–477, 2019.

[4] Michael I Jordan and Tom M Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

[5] Olatomiwa O Bifarin, David A Gaul, Samyukta Sah, Rebecca S Arnold, Kenneth Ogan, Viraj A Master, David L Roberts, Sharon H Bergquist, John A Petros, Facundo M Fernandez, et al. Machine learning-enabled renal cell carcinoma status prediction using multiplatform urine-based metabolomics. *Journal of proteome research*, 20(7):3629–3641, 2021.