

LLM-based agents can autonomously generate, execute and refine in silico models for biology

Author:

Prerana Chandratre

Co-authors:

Lutz Brusch, Simon Syga

Many open biological questions, from human embryogenesis to complex diseases like cancer, cannot be solved by experiments alone but require integration with biophysical and computational modeling. However, generating, executing and refining such in silico models remains challenging, requiring cross-domain expertise from mathematics, computer science, and cell biology. Here, we investigate the capabilities of large language model (LLM)-based agents to assist in this difficult task.

As a starting point we use the modeling framework Morpheus, an open-source platform supporting various mathematical paradigms [1]. It is based on the declarative modeling language MorpheusML for specifying individual in silico experiments in a modular and reproducible way, a unique strength compared to other simulation frameworks. Building on the Model Context Protocol we equip an LLM-based agent with various tools that allow it to autonomously consult a library of example models, and create, run, debug, and refine Morpheus simulations [2]. We evaluate our setup on a benchmark of fifteen scientific papers containing in silico models from diverse subdomains of cell biology, including embryogenesis and cancer biology.

We find that - with our tools - LLM-agents can produce valid in silico models in all the cases. Our results suggest that future AI agents might be able to not only assist in data analysis and hypothesis generation, but also to autonomously test their own ideas with in silico experiments.

References

[1] <https://morpheus.gitlab.io>

[2] <https://github.com/Prerana17/Morpheus.AI>