

Sistemas Multi-Agentes Cognitivos Baseados em LLMs para Resolução Colaborativa de Tarefas

Jean Reinholt¹

¹Centro de Desenvolvimento Tecnológico (CDTec)
Universidade Federal de Pelotas (UFPel)
Pelotas – RS – Brasil

jreinholt@inf.ufpel.edu.br

Resumo. Sistemas multi-agentes baseados em Modelos de Linguagem de Grande Escala (LLMs) têm se destacado como uma abordagem promissora para a resolução colaborativa de problemas complexos. Este trabalho analisa a arquitetura proposta por Silva et al. (2025), que introduz um sistema multi-agente conversacional cognitivo onde um agente orquestrador distribui dinamicamente tarefas entre agentes especializados. Cada agente é equipado com cinco módulos cognitivos: Memória Híbrida, Planejamento, Execução, Percepção e Raciocínio. Realizou-se uma análise crítica da arquitetura, identificando seus pontos fortes — como a modularidade e o uso de memória compartilhada — e suas limitações, incluindo desafios de escalabilidade e avaliação. Propõem-se melhorias relacionadas à gestão de memória e aos mecanismos de coordenação entre agentes.

Abstract. Multi-agent systems powered by Large Language Models (LLMs) have emerged as a promising approach for collaborative complex problem-solving. This work analyzes the architecture proposed by Silva et al. (2025), which introduces a conversational cognitive multi-agent system where an orchestrator agent dynamically distributes tasks among specialized agents. Each agent is equipped with five cognitive modules: Hybrid Memory, Planning, Execution, Perception, and Reasoning. A critical analysis of the architecture was conducted, identifying its strengths — such as modularity and shared memory usage — and its limitations, including scalability and evaluation challenges. Improvements related to memory management and inter-agent coordination mechanisms are proposed.

1. Introdução

Arquiteturas de sistemas multi-agentes baseadas em Modelos de Linguagem de Grande Escala (LLMs) têm emergido como uma abordagem poderosa para a resolução de problemas complexos, permitindo que múltiplos agentes autônomos operem em um ambiente compartilhado [Wang et al. 2024]. Esses agentes se comunicam, colaboram e coordenam suas ações para alcançar objetivos específicos, tornando tais sistemas particularmente vantajosos em domínios dinâmicos e complexos.

Apesar dos avanços, sistemas multi-agentes baseados em LLMs ainda enfrentam desafios significativos que impactam sua efetividade e escalabilidade. A garantia de comunicação e coordenação fluidas entre agentes, a manutenção de consistência comporta-

mental e a gestão eficaz de memória em ambientes dinâmicos permanecem como questões em aberto [Sumers et al. 2024].

Silva et al. [Silva et al. 2025] propõem uma arquitetura multi-agente inovadora na qual um agente orquestrador distribui dinamicamente tarefas entre agentes cognitivos especializados, denominados *Task-driven Agents*. Cada agente é composto por cinco módulos: Memória Híbrida, Planejamento, Execução, Percepção e Raciocínio. O sistema utiliza uma memória compartilhada e um paradigma conversacional para coordenação.

Este trabalho tem como objetivo analisar criticamente a arquitetura proposta, identificando seus pontos fortes e limitações, e propor melhorias para os mecanismos de coordenação e gestão de memória. O restante do artigo está organizado da seguinte forma: a Seção 2 apresenta o referencial teórico; a Seção 3 descreve a metodologia do artigo de referência; a Seção 4 discute os resultados e propõe melhorias; e a Seção 5 apresenta as conclusões.

2. Referencial Teórico

2.1. Agentes Inteligentes e Sistemas Multi-Agentes

Um agente inteligente é uma entidade computacional autônoma capaz de perceber seu ambiente e agir sobre ele de forma a alcançar seus objetivos [Wooldridge and Jennings 1995]. Wooldridge e Jennings definem quatro propriedades fundamentais: autonomia, reatividade, proatividade e habilidade social. Sistemas Multi-Agentes (SMA) estendem essa definição ao considerar múltiplos agentes interagindo em um ambiente compartilhado, onde a coordenação, comunicação e cooperação são essenciais [Weiss 1999, Russell and Norvig 2021].

2.2. Arquiteturas Cognitivas e Agentes Baseados em LLMs

Sumers et al. [Sumers et al. 2024] propuseram o framework CoALA (*Cognitive Architectures for Language Agents*), que estrutura agentes de linguagem com base em princípios de arquiteturas cognitivas. O CoALA define agentes ao longo de três dimensões: memória (de trabalho, episódica, semântica e procedural), espaço de ação e tomada de decisão.

Wang et al. [Wang et al. 2024] apresentam uma revisão sistemática de agentes autônomos baseados em LLMs, introduzindo um framework unificado com quatro módulos: *Profile*, *Memory*, *Planning* e *Action*. Tais frameworks fornecem bases teóricas para o desenvolvimento de agentes mais adaptativos e conscientes de memória.

2.3. Frameworks Multi-Agentes Colaborativos

Qian et al. [Qian et al. 2024] propuseram o *ChatDev*, um framework de desenvolvimento de software baseado em chat onde múltiplos agentes LLM com papéis predefinidos colaboram iterativamente. Zhang et al. [Zhang et al. 2023] introduziram o CoELA (*Cooperative Embodied Language Agent*), projetado para agentes corporificados com integração de LLMs para raciocínio e geração de linguagem. Ambos os exemplos demonstram o potencial da integração de LLMs com componentes cognitivos para aprimorar a colaboração entre agentes [Becker 2024].

3. Metodologia

3.1. Arquitetura Proposta por Silva et al.

A arquitetura proposta por Silva et al. [Silva et al. 2025] consiste em um sistema multi-agente onde agentes cognitivos autônomos, denominados *Task-driven Agents*, colaboram para resolver problemas complexos. O sistema é composto por três elementos principais: (i) um agente Orquestrador, responsável por distribuir tarefas e mediar a comunicação; (ii) um conjunto de *Task-driven Agents* especializados; e (iii) uma Memória Compartilhada acessível a todos os agentes.

Cada *Task-driven Agent* possui cinco módulos cognitivos internos. A **Memória Híbrida** integra memória de curto prazo (*Working Memory*) e longo prazo, incluindo Memória Semântica (baseada em RAG — *Retrieval Augmented Generation* [Lewis et al. 2021]) e Memória Procedural. O módulo de **Planejamento** encapsula a tomada de decisão do agente. A **Execução** transforma planos em ações concretas utilizando ferramentas externas. A **Percepção** enriquece o conhecimento do agente a partir do ambiente. O **Raciocínio** gerencia as interações com o LLM e o código do agente.

3.2. Aplicação de Caso e Avaliação

A viabilidade da arquitetura foi demonstrada por meio de uma aplicação envolvendo três agentes: um Agente Moderador (orquestrador), um Agente Financeiro (coleta de dados via *yfinance*) e um Agente de Código (geração de visualizações com *Plotly*). O sistema foi implementado em Python utilizando LangGraph, com gestão de memória via OpenSearch e modelo GPT-4o. O paradigma conversacional *Report* foi utilizado para coordenação entre agentes.

4. Resultados e Discussões

4.1. Pontos Fortes

A arquitetura apresenta contribuições relevantes. A abordagem de agentes orientados por tarefa (*task-driven*) promove modularidade e escalabilidade, permitindo que novos agentes sejam adicionados sem modificar a estrutura existente. A Memória Compartilhada funciona como um *blackboard*, facilitando a troca de informações entre agentes sem comunicação direta. Além disso, a integração de múltiplos tipos de memória (Working, Semantic e Procedural) alinha-se com frameworks cognitivos estabelecidos como o CoALA [Sumers et al. 2024].

4.2. Limitações e Melhorias Propostas

Apesar dos avanços, identificam-se limitações importantes. A avaliação foi conduzida com apenas um cenário de aplicação, o que limita a generalização dos resultados. A Memória Compartilhada não possui mecanismos de resolução de conflitos ou controle de consistência. Ademais, a comunicação entre agentes ocorre exclusivamente via Orquestrador, sem possibilidade de interação direta.

Como melhorias, propõe-se: (i) a incorporação de *Knowledge Graphs* para enriquecer a Memória Semântica, conforme sugerido pelos próprios autores; (ii) a implementação de protocolos de comunicação direta entre agentes para cenários que exijam colaboração intensiva; (iii) a realização de estudos ablativos para avaliar a contribuição individual de cada módulo cognitivo; e (iv) a avaliação da arquitetura com múltiplos cenários e *benchmarks* padronizados para validar sua generalidade.

5. Conclusões

Este trabalho analisou criticamente a arquitetura multi-agente cognitiva proposta por Silva et al. [Silva et al. 2025], que introduz um sistema conversacional baseado em LLMs para resolução colaborativa de tarefas. A análise identificou pontos fortes significativos, como a modularidade proporcionada por agentes orientados por tarefa e a integração de múltiplos tipos de memória inspirada em arquiteturas cognitivas [Sumers et al. 2024].

Foram também identificadas oportunidades de melhoria, incluindo a necessidade de avaliação mais abrangente, mecanismos de resolução de conflitos na memória compartilhada e protocolos de comunicação direta entre agentes. Tais melhorias são essenciais para que a arquitetura alcance seu potencial em cenários de maior escala e complexidade.

Como trabalhos futuros, destacam-se a integração de *Knowledge Graphs* para enriquecer a memória semântica dos agentes, o desenvolvimento de *benchmarks* padronizados para avaliação de arquiteturas multi-agentes cognitivas e a investigação de paradigmas alternativos de comunicação entre agentes.

Referências

- Becker, J. (2024). Multi-agent large language models for conversational task-solving. *arXiv preprint arXiv:2410.22932*.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., tau Yih, W., Rocktäschel, T., Riedel, S., and Kiela, D. (2021). Retrieval-augmented generation for knowledge-intensive NLP tasks. *arXiv preprint arXiv:2005.11401*.
- Qian, C., Liu, W., Liu, H., Chen, N., Dang, Y., Li, J., Yang, C., Chen, W., Su, Y., Cong, X., et al. (2024). ChatDev: Communicative agents for software development. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15174–15186.
- Russell, S. and Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson, 4th edition.
- Silva, E., Santos, F. A., Thompson, P., and dos Reis, J. C. (2025). LLM-powered conversational multi-agent cognitive system for collaborative task solving. In *19th Workshop-School on Agents, Environments, and Applications (WESAAC)*, pages 59–70.
- Sumers, T. R., Yao, S., Narasimhan, K., and Griffiths, T. L. (2024). Cognitive architectures for language agents. *Transactions on Machine Learning Research*. Survey Certification.
- Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., et al. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):186345.
- Weiss, G. (1999). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press.
- Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, 10(2):115–152.
- Zhang, H., Du, W., Shan, W., Zhou, J., Du, Q., Du, Y., Tenenbaum, J. B., Tianmin, S., and Chuang, G. (2023). Building cooperative embodied agents modularly with large language models. *arXiv preprint arXiv:2307.02485*.