

Sistemas Multi-Agentes Cognitivos Baseados em LLMs

Para Resolução Colaborativa de Tarefas

Jean Reinhold

Centro de Desenvolvimento Tecnológico (CDTec)
Universidade Federal de Pelotas (UFPel)

jreinhold@inf.ufpel.edu.br

22 de fevereiro de 2026

- 1 Introdução
- 2 Referencial Teórico
- 3 Metodologia
- 4 Resultados
- 5 Propostas e Conclusão

- Arquiteturas de SMA baseadas em LLMs: abordagem poderosa para resolução de problemas complexos
- Múltiplos agentes autônomos em ambiente compartilhado
- Comunicação, colaboração e coordenação de ações
- Vantajosos em domínios dinâmicos e complexos

Referência Principal

Silva et al. (2025): arquitetura multi-agente cognitiva com orquestrador e agentes especializados (*Task-driven Agents*)

Problemas em Aberto

- 1 Comunicação e coordenação fluida entre agentes
- 2 Manutenção de consistência comportamental
- 3 Gestão eficaz de memória em ambientes dinâmicos
- 4 Escalabilidade para cenários de maior complexidade

Objetivo deste Trabalho

Analisar criticamente a arquitetura proposta por Silva et al. (2025), identificar pontos fortes e limitações, e propor melhorias para coordenação e gestão de memória.

Definição (Wooldridge & Jennings, 1995)

Entidade computacional autônoma capaz de perceber seu ambiente e agir sobre ele de forma a alcançar seus objetivos.

Propriedades Fundamentais

- **Autonomia** — opera sem intervenção direta
- **Reatividade** — percebe e responde ao ambiente
- **Proatividade** — inicia comportamentos dirigidos a objetivos
- **Sociabilidade** — interage com outros agentes

Conceitos Fundamentais

Múltiplos agentes interagindo em ambiente compartilhado (Weiss, 1999; Russell & Norvig, 2021)

Coordenação Organização de ações entre agentes

Comunicação Troca de informações e intenções

Cooperação Trabalho conjunto para objetivos comuns

Frameworks Colaborativos

- **ChatDev** (Qian et al., 2024)
Agentes LLM com papéis predefinidos em desenvolvimento de software
- **CoELA** (Zhang et al., 2023)
Agentes corporificados com LLMs para raciocínio e linguagem

CoALA (Sumers et al., 2024)

Framework de arquiteturas cognitivas para agentes de linguagem:

- **Memória:** de trabalho, episódica, semântica, procedural
- **Espaço de ação**
- **Tomada de decisão**

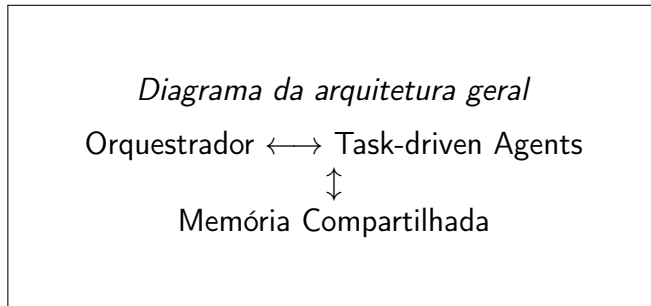
Wang et al. (2024)

Framework unificado para agentes LLM com quatro módulos:

- *Profile*
- *Memory*
- *Planning*
- *Action*

Base Teórica

Ambos os frameworks fornecem fundamentos para agentes mais adaptativos e conscientes de memória.



Orquestrador

Distribui tarefas e media
comunicação

Task-driven Agents

Agentes cognitivos
especializados

Memória

Compartilhada

Acessível a todos os
agentes

Agente Orquestrador

- Distribui dinamicamente tarefas
- Media a comunicação entre agentes
- Utiliza paradigma conversacional *Report*
- Garante coordenação global do sistema

Task-driven Agents

- Agentes cognitivos especializados
- Equipados com 5 módulos cognitivos
- Operam autonomamente em subtarefas
- Reportam resultados ao orquestrador

Diagrama dos 5 módulos cognitivos

Memória Híbrida

Working Memory +
Semântica (RAG) +
Procedural

Planejamento

Tomada de decisão do
agente

Execução

Transforma planos em ações
com ferramentas externas

Percepção

Enriquece conhecimento a
partir do ambiente

Raciocínio

Gerencia interações com o
LLM e código

Diagrama do fluxo entre os 3 agentes

Ag. Moderador

Orquestrador: distribui tarefas e coordena comunicação

Ag. Financeiro

Coleta de dados via *yfinance* para análise

Ag. de Código

Geração de visualizações com *Plotly*

Stack: Python + LangGraph + OpenSearch + GPT-4o

Contribuições Relevantes

- **Modularidade e escalabilidade:** abordagem *task-driven* permite adicionar novos agentes sem modificar a estrutura existente
- **Memória compartilhada (*blackboard*):** facilita troca de informações sem comunicação direta
- **Integração de múltiplos tipos de memória:** Working, Semantic (RAG) e Procedural, alinhada ao CoALA
- **Paradigma conversacional *Report*:** coordenação flexível entre orquestrador e agentes
- **Uso de personas:** especialização comportamental dos agentes

Desafios e Lacunas

- ❶ **Avaliação limitada:** apenas um cenário de aplicação, dificultando a generalização dos resultados
- ❷ **Memória compartilhada sem controle:** ausência de mecanismos de resolução de conflitos e consistência
- ❸ **Comunicação centralizada:** interação entre agentes ocorre exclusivamente via orquestrador
- ❹ **Dependência de modelo único:** avaliação realizada apenas com GPT-4o
- ❺ **Falta de *benchmarks*:** sem comparação com métricas padronizadas da literatura

Melhorias Propostas

- **Knowledge Graphs** para enriquecer a memória semântica
- Protocolos de **comunicação direta** entre agentes
- Estudos **ablativos** para avaliar contribuição individual dos módulos

Trabalhos Futuros

- Avaliação com **múltiplos cenários** e *benchmarks* padronizados
- Paradigmas **alternativos de comunicação** entre agentes
- Mecanismos de **resolução de conflitos** na memória compartilhada

Síntese

- Análise crítica da arquitetura multi-agente cognitiva de Silva et al. (2025)
- Pontos fortes: modularidade *task-driven*, integração de memória inspirada em arquiteturas cognitivas
- Oportunidades: avaliação mais abrangente, resolução de conflitos, comunicação direta entre agentes
- Contribuição: identificação de caminhos para amadurecimento de SMA cognitivos baseados em LLMs

Obrigado!

jreinhold@inf.ufpel.edu.br