CSE 598 Research Proposal: State Representation Learning for Long-Term Multi-Agent Interactions

{diangao, anishcha, linglong}@umich.edu
University of Michigan, Ann Arbor

February 23, 2025

Abstract

This research provides a systematic comparison of state representation architectures in multi-agent systems, analyzing GraphDB's relational modeling, VectorDB's continuous embeddings, and Semantic Memory's hybrid approach. Through controlled experiments in Tic-Tac-Toe variants and their latent-space extensions, we establish a framework for evaluating: (1) planning depth through -bench metrics, (2) cross-domain adaptability via representation similarity analysis, and (3) memory optimization through parameter-efficient fine-tuning. Our methodology combines theoretical analysis of representation spaces with empirical validation, offering practical guidelines for architecture selection in real-time coordination systems and adaptive AI applications.

1 Introduction and Motivation

The effectiveness of state representation learning constitutes a fundamental challenge in developing robust multi-agent systems, particularly for applications requiring long-term interaction and cross-environment adaptability. While current LLM-based agents demonstrate proficiency in isolated decision-making tasks, three critical gaps persist:

- **Representation-Environment Mismatch**: Fixed memory architectures struggle to adapt between discrete symbolic reasoning (e.g., game rules) and continuous latent-space decision-making
- **Multi-Horizon Coordination**: Existing systems lack mechanisms for maintaining coherent state representations across varying interaction timescales in multi-agent scenarios
- **Post-Training Instability**: Learned representations often degrade when deployed in environments differing from their training regimes

Our work addresses these challenges through a structured investigation of three state representation paradigms in Cognitive Language Agents:

- **GraphDB**: Explicit relational modeling for strategic game trees
- **VectorDB**: Continuous embedding spaces for probabilistic reasoning
- **Semantic Memory**: Hybrid neuro-symbolic representations via LLM abstraction

Using Tic-Tac-Toe variants as our experimental testbed, we establish a controlled environment to analyze:

- Discrete-to-continuous representation transitions through latent space projections
- Multi-agent coordination dynamics under constrained memory budgets
- Post-training optimization via parameter-efficient fine-tuning

This research delivers crucial insights for:

- *Architecture Designers*: Empirical guidelines for memory system selection based on environment characteristics

- *Agent Practitioners*: Strategies for maintaining representation consistency in real-world deployments
- *Theoreticians*: Quantitative framework for analyzing representation learning dynamics

2 Research Objectives

Our investigation establishes three principal research objectives that systematically address the core challenges in state representation learning for multi-agent systems:

1. **Architecture-Specific Reasoning Capacity Analysis (Experiment 1)**
   This objective focuses on quantifying how different memory architectures influence strategic decision-making in structured environments. Through controlled multi-agent Tic-Tac-Toe experiments, we will:
   - Compare the planning horizon supported by GraphDB's explicit game tree representations versus VectorDB's continuous embedding strategies
   - Measure coordination efficiency differentials through win rate analysis across grid sizes (3×3 to 4×4)
   - Evaluate the effectiveness of Semantic Memory's RAG mechanism in reducing redundant moves through move sequence entropy calculations
2. **Cross-Environment Representation Transfer Assessment (Experiment 2)**
   This objective examines the adaptability of learned state representations across decision-making regimes. Using our continuous Tic-Tac-Toe variants, we will:
   - Develop quantitative transferability metrics comparing discrete-to-smoothed and discrete-to-latent transitions
   - Analyze strategy consistency through KL divergence measurements between original and projected decision distributions
   - Validate failure recovery mechanisms by introducing controlled perturbations in continuous state spaces
3. **Post-Training Optimization Framework Validation (Experiment 3) (Tentative)**
   This tentative objective evaluates enhancement strategies for learned representations. Building on Experiments 1-2, we will:
   - Assess LoRA fine-tuning's capacity to preserve memory stability across extended interaction horizons (100+ game iterations)
   - Quantify COCONUT-style contrastive learning's impact on cross-architecture knowledge transfer
   - Establish adaptation speed benchmarks for novel task variations (Connect-4 rule adaptations)

These objectives are systematically explored through our experimental framework, incorporating both discrete and continuous decision-making scenarios in Tic-Tac-Toe variants.

3 Methodology

3.1 Data Collection

[Describe data sources] will be collected using [tools/methods]. Preprocessing will include [steps].

3.2 Analytical Approach

We propose the following model framework:

$$f(x) = \sum_{i=1}^{n} \alpha_i \phi(x_i) + \epsilon \tag{1}$$

where ϕ represents [feature mapping] and ϵ is [error term].

4 Expected Outcomes

1. Outcome 1: [Direct result of Objective 1]
2. Outcome 2: [Validation metric for Objective 2]
3. Outcome 3: [Theoretical contribution]

5 Timeline

Timeline	**Task**
Weeks 1-2	Literature review and problem formulation
Weeks 3-4	Data collection framework setup
Weeks 5-8	Model development and validation
Weeks 9-10	Experimental analysis
Weeks 11-12	Paper drafting and final submission

References