Project 16
Weekly Update: Jul 20, 2026

Cards on the Table:

The Winning Move is Not in the Rules

♠ ♥ ♦ ♣

REU Student: Jack Lamberg

Sabbir Mollah, Sarinda Samarasinghe, Dr. Mubarak Shah



Problem Formulation

Real card-game agents should improve after every game.

1. Card-game skill depends on learning what cards, effects, and combos exist.
2. Rules define legal moves, but card effects have higher precedence over rules.
3. A move that looks good now may be bad with knowledge on what cards exist.

Our contribution:

1. LLM/VLM Understanding: Evaluate LLM/VLM performance on in-game questions.
2. LLM/VLM Gameplay: Evaluate LLM/VLM performance in playing full games.
3. Memory-based agent: Evaluate impact of accumulated memory on these metrics.

Week 9-10 Deliverables

Task	Week 8-9
Implement 3 baselines memory modules	✅
Implement our memory module	90%
Evaluate module performance	Running
Improve our module	To-Do
Simulate more games to have data converge	To-Do

Our Memory Perspective

- We want to have a training-free memory module
- We want to learn from multiple games in parallel

- Three memory types:
  - **Strategy:** Overall ways to play the game well
  - **Card:** Specific memories about how to use each card
  - **Belief:** Guessing the opponent's plan based on their previous turns/visible cards

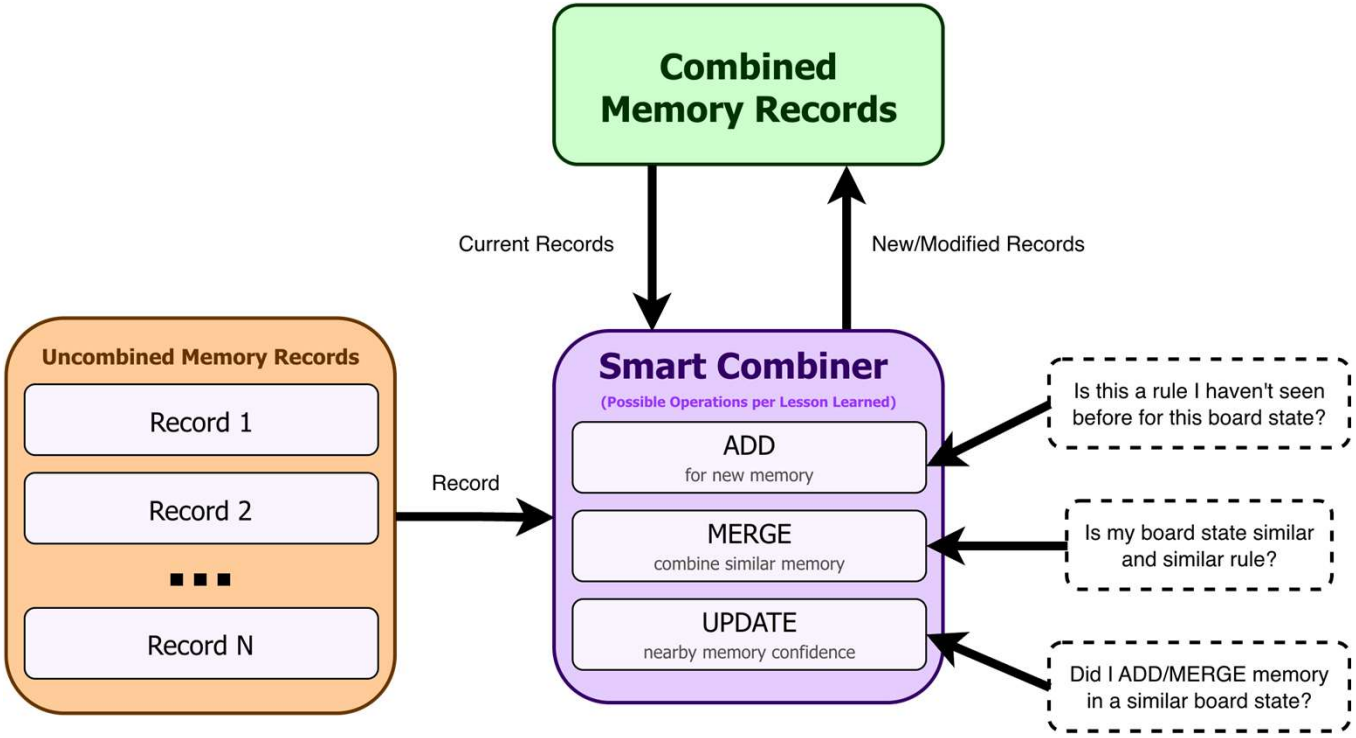
Our Memory Perspective- Three Stages

- **Build stage**:
  - Individually build memory documents from games
- **Combine stage:**
  - ADD/MERGE/UPDATE documents sequentially
- **Retrieval stage:**
  - Use board state tag to retrieve relevant memory to the current game state

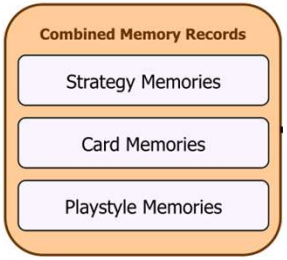
Build Stage

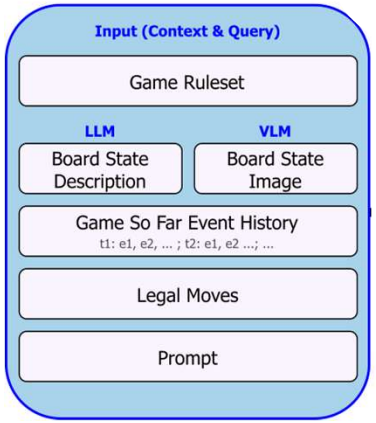


Combine Stage



Retrieval Stage





Our Memory Example

- **Strategy Memory**
  - **Rule:** Balance early board development with cautionary tactics to avoid overcommitment.
  - **Reasoning:** Neglecting early board development in favor of reactive tactics left the player vulnerable to the opponent's early aggression and limited their defensive capabilities
  - **Tag (In English):** Coilbind Reversal, Tidewall Fog, Tidewall Fog, Pearlfin Serpent, Skycoil Serpent, Undertow Drag
  - **Confidence:** 0.531441

Next Steps

- Need expert level games to use for memory build stage
- Evaluate expert level memory against all four bots
- Test memory system with multiple LLM/VLM models

Thanks for listening!

Questions?

