

Self-Reflection Structured Experience 프레임워크를 이용한 대형언어모델(LLM) 에이전트의 학습 최적화 연구

오현영, 이준행, 강영명*

성결대학교

ohy414@sungkyul.ac.kr, hkhk9495@sungkyul.ac.kr, *ykang@sungkyul.ac.kr

A Structured Self-Reflection Experience Framework for Adaptive Learning in LLM Agents

HyunYoung Oh, Junhaeng Lee, Young-myung Kang*

Sungkyul Univ.

요약

최근 대규모 언어 모델(Large Language Model, LLM)은 단순한 자연어 이해를 넘어 복잡한 태스크 수행이 가능한 지능형 에이전트로 활용되고 있다. 그러나 이러한 LLM 기반 에이전트는 막대한 비용과 시간이 소요되는 파인튜닝 과정이 요구되며 외부 환경 변화에 대한 적응력이 제한적이다. 따라서 본 논문에서는 Memento 기반 학습 구조에 자기 성찰(Self-reflection) 메커니즘을 통합한 SRSE(Self-Reflection Structured Experience) 프레임워크를 제안한다. SRSE는 실패 경험에서 원인과 교정 행동을 자동으로 분석하여 구조화된 학습이 가능하며 강화된 외부 메모리(Enhanced Case Bank)에 축적된 경험을 통해 유사 상황에서 경고하고 반복 가능한 의사결정을 지원한다. GAIA와 Deepresearcher 데이터셋을 활용하여 성능 평가를 진행한 결과 SRSE는 Memento Baseline 대비 성공률(Success Rate)을 약 7.9% 향상되었다. 이러한 결과는 SRSE의 구조화된 실패 경험이 에이전트의 계획 및 행동 교정에 효과적으로 기여하며 실패 경험을 단순 회피 정보가 아닌 전략적 자원으로 변환하는 본 프레임워크의 유효성을 입증한다.

I. 서론

최근 대규모 언어 모델(Large Language Model, LLM)은 단순한 자연어 이해 수준을 넘어 복잡한 태스크를 능동적으로 수행하는 지능형 에이전트의 핵심 기술로 자리매김하고 있다[1]. 그러나 이러한 LLM 기반 에이전트는 여전히 두 가지 주요 한계를 가진다. 첫째, 에이전트의 성능은 기반 모델 자체의 성능에 크게 의존하며 유연하고 고성능의 에이전트를 개발하기 위해서는 막대한 비용과 시간이 소요되는 파인튜닝(Fine-Tuning) 과정이 필수적으로 요구된다. 둘째, LLM 에이전트는 외부 환경 변화에 대한 적응력이 제한적이며 실행 결과의 사실적 정확도를 보장하기 어렵다. 이러한 문제를 해결하기 위해 최근 Memento[2]와 같은 외부 메모리 기반 접근 방식이 제안되었다. Memento는 에이전트의 실행 경험을 외부 메모리(Case Memory)에 저장하고 재활용함으로써 모델 재학습 없이 GAIA와 DeepResearcher 등 다양한 벤치마크에서 성능 향상을 가능하게 하였다. 그러나 Memento는 실행 궤적(Raw Trajectory)을 비구조적으로 저장하기 때문에 실패 사례에 대한 심층적이고 구조화된 표현이 제한되며 실패 경험이 단순 회피 대상으로만 축적된다는 한계를 지닌다[2].

이에 본 논문에서는 Memento의 경험 구조에 자기 성찰 메커니즘[3]을 통합한 SRSE(Self-Reflection Structured Experience) 프레임워크를 제안한다. SRSE는 에이전트의 실패 경험으로부터 실패 원인과 교정 행동을 자동으로 생성 및 구조화함으로써 실패를 단순 회피가 아닌 구조적 학습 자원으로 활용한다. 이를 통해 인간의 오답노트와 유사한 방식으로 과거의 실패를 개선된 행동 전략으로 재구성하여 에이전트의 지속적인 자기 개선을 효과적으로 지원한다. 성능 평가 결과 SRSE 프레임워크는 Memento Baseline 모델 대비 실패한 질문에 대한 재질의 시 약 7.9% 이상의 성공률 향상을 달성하며 실패 경험이 효율적인 전략적 자원으로 기능할 수 있음을 명확히 입증하였다.

II. 본론

1) Memento 구조

Memento는 에이전트 경험을 체계적으로 관리하는 비파라미터적 지식 저장 방식이다. 에이전트는 환경과 상호작용하며 수행한 실행 궤적을 외부 메모리에 저장하고, Planner는 이후 유사한 상황에서 이를 검색하여 의사결정에 활용한다. Memento의 주요 특징은 Table I 과 같다.

Feature	Description
경험 저장	Planner-Executor의 수행 결과를 실행 궤적에 저장
경험 검색	유사 상황에서 과거 경험을 검색하여 계획에 재활용
지식 재활용	별도의 모델 튜닝 없이 경험 기반 의사결정 가능

Table I Memento 주요 특징

Fig 1은 Memento의 핵심 기능을 설명한다. Planner는 외부 메모리를 참조하고, Executor는 MCP Server를 통해 외부 도구를 호출하고 결과를 Tool Memory에 기록한다. 이 구조는 경험 기반 의사결정이 가능하지만 실패 경험을 해석하거나 교정하는 기능은 부족하다. 따라서 외부 메모리에 저장된 실행 궤적에는 실패 원인이나 교정 행동 정보가 포함되어 있지 않아 Planner가 이를 교정적 경험으로 활용하는 데 근본적인 한계가 있다.

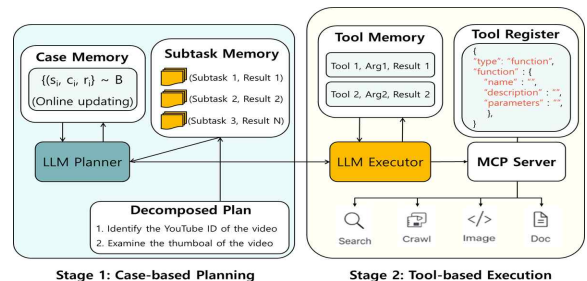


Fig. 1. Memento 구조

2) SRSE(Self-Reflection Structured Experience) 프레임워크

본 논문에서 제안하는 SRSE 프레임워크는 Memento의 경험 구조에 Reflection Layer를 결합한 자기 성찰 기반 학습 구조이다. 이는 실패 경험을 분석하고 교정 행동을 자동으로 생성한다. SRSE의 실행 궤적 구조(E_{SRSE})는 기존의 Case, Plan, Reward, Memory의 구조와 달리 Case, Plan, Reward, Reflection_Reason, Reflection_Correction 다섯 가지 요소로 구성된다. Fig. 2를 보면 Reflection Layer를 통해 Planner가 실패 경험을 자동으로 분석하고 교정된 경험을 강화된 외부 메모리(Enhanced Case Bank)에 저장하여 유사 상황에서 향상된 의사결정을 수행할 수 있다.

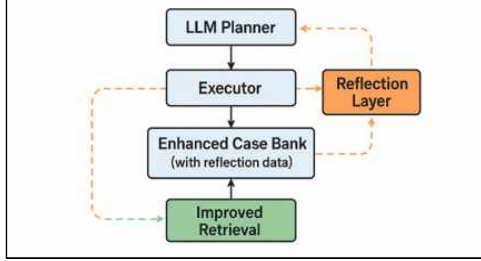


Fig. 2. SRSE 프레임워크 구조

III. 성능 평가

본 논문에서는 Memento Baseline 모델과 Reflection Layer를 추가한 SRSE 프레임워크의 성능을 비교하였다. Planner 모델은 gpt-4o-mini, Executor 모델은 o3 모델을 사용하였으며 MCP Tool 구성에는 검색 도구(SearxHG, playwright), 코딩 에이전트, 웹 크롤링 툴(crawl4ai), 수학 에이전트를 포함하였다. 성능 평가는 GAIA[4]와 Deepresearcher 데이터셋[5]을 활용하여 수행하였다. 성공률(Success Rate)은 실패한 사례에 대해 다음 시도에서의 동일 질문에 대한 응답 성공 정도를 의미한다. 또한 각 모델의 실행 궤적에 따라 검색기를 재학습하여 검색기의 성능도 함께 비교하였다. 모든 실험은 동일한 초기 경험 데이터(Case Bank)를 공유한 상태에서 진행하였으며 총 453개의 질문을 수행하였다. 최종적으로 Memento baseline 모델의 memory는 1,731개, SRSE 모델의 memory는 2,637행의 실행 궤적을 활용하였다.

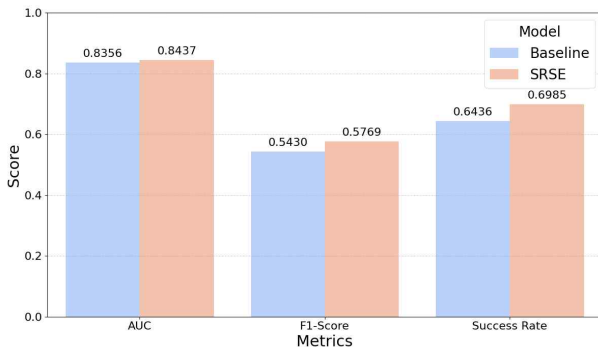


Fig. 3. Memento Baseline과 SRSE 검색기 성능 비교

Fig. 3에서 SRSE 검색기는 Memento Baseline 대비 Reflection 기반 구조화 경험 데이터를 활용함으로써 검색기의 품질이 향상되었음을 확인할 수 있다. SRSE 에이전트는 Memento Baseline 대비 7.9%의 성공률 향상을 달성하였다. Fig. 4는 동일한 질문에 대한 Memento Baseline과 SRSE 에이전트의 응답 예시를 보여준다. Memento Baseline은 단순히 계획만 제시하는 반면 SRSE는 실패 원인과 교정 제안을 함께 제공하여 체계적이고 반복 가능한 전략을 수립한다.

Question 1 How many pages if the 2023 IPCC report (85 pages version) mentions nuclear energy? Baseline: plan -> [{"id":1, description: "Attempt to access the 2023 IPCC report"}] SRSE: failure -> reason: 'inability to locate the 2023 IPCC report'; suggested correction: [{"id":1, description: "Reaccess the IPCC report by utilizing official sources like the website or academic databases."}]
Question 2 How many studio albums were published by Mercedes Sosa between 2000 and 2009 (included)? You can use the latest 2022 version of english wikipedia. Baseline: plan -> [{"id":1, description: "Access an music database"}] SRSE: failure -> reason: 'Access to external databases for information'; suggested correction: [{"id":1, description: "Utilize a music database such as Disoges or Allmusic for exact details"}]

Fig. 4. Memento Baseline과 SRSE 모델의 GAIA 응답

IV. 결론

본 논문에서는 Memento 기반 경험 학습 구조에 자기 성찰 매커니즘을 통합한 SRSE를 제안하였다. SRSE는 실패 경험을 구조화하여 외부 메모리에 축적함으로써 에이전트가 유사 상황에서 더 정교하고 반복 가능한 의사결정을 수행할 수 있음을 확인하였다. 또한 GAIA 데이터셋에 대한 응답 사례에서도 Reflection Layer가 계획과 행동 교정에 기여함을 확인하였다. 이는 SRSE가 에이전트의 자기 개선과 경험 재활용성을 동시에 강화함을 나타낸다. 향후 연구에서는 더욱 정교한 평가 지표를 개발하고 메모리 효율 저하 문제를 개선함으로써 다양한 도메인과 복잡한 환경에서도 높은 일반화 성능을 발휘할 수 있는 에이전트로 발전시킬 예정이다.

참고 문헌

- [1] X. Huang, W. Liu, X. Chen, X. Wang, H. Wang, D. Lian, ... and E. Chen, "Understanding the planning of LLM agents: A survey," arXiv preprint arXiv:2402.02716, 2024.
- [2] H. Zhou, Y. Chen, S. Guo, X. Yan, K. H. Lee, Z. Wang, K. Y. Lee, G. Zhang, K. Shao, L. Yang, and J. Wang, "Memento: Fine-tuning LLM Agents without Fine-tuning LLMs," arXiv preprint arXiv:2508.16153, 2025.
- [3] N. Shinn, F. Cassano, E. Berman, A. Gopinath, K. Narasimhan, and S. Yao, "Reflexion: Language Agents with Verbal Reinforcement Learning," arXiv preprint arXiv:2303.11366, 2023.
- [5] J. Yang, A. Jayakumar, R. Li, et al., "GAIA: A benchmark for general AI assistants," Hugging Face, Nov. 22, 2023. [Online]. Available: <https://huggingface.co/datasets/gaia-benchmark/GAIA>
- [5] Agent on the-Fly, "Memento: Fine tuning LLM Agents without Fine tuning LLMs," GitHub repository, 2025. Available: <https://github.com/Agent-on-the-Fly/Memento>