

LLM を用いた人狼エージェントにおける動的記憶蓄積による発言の一貫性維持と効率化

大溝 達也¹, 眞部 雄介²
Tatsuya Ohmizo¹, Yusuke Manabe²
¹ 千葉工業大学大学院 情報科学研究科
² 千葉工業大学 認知情報科学科

概要

本稿では、LLM を用いた人狼エージェントにおいて、文脈維持とコスト削減を両立させる「動的記憶蓄積」手法を提案する。従来の Full History 法はトークン増加に伴うコスト肥大化や忘却 (Lost in the Middle) が課題であった。提案手法では、エージェントが自身のターンごとに直近の会話を要約し、内部記憶を動的に更新することでコンテキストサイズを抑制する。5人村でのシミュレーション (40ゲーム) の結果、提案手法はトークン消費量を劇的に削減しつつ、投票成功率においてベースラインの2倍 (0.44) のスコアを記録した。主観評価で自然さに課題を残したものの、役職行動は維持されており、コスト対効果の高い実用的な手法であると確認した。

1 はじめに

人狼ゲーム (Werewolf Game) は、会話を通じて他者の正体を見破る不完全情報ゲームであり、人工知能 (AI) 研究における重要なベンチマークの一つとして扱われてきた [1]。従来の人狼 AI は、ルールベースや統計的手法を用いたものが主流であったが、近年では大規模言語モデル (以下 LLM) の急速な発展により、人間のように自然言語を用いて議論を行う対話型エージェントの開発が可能となっている。LLM を用いたエージェントは、従来の AI では困難であった「説得」や「嘘」、「文脈に応じた柔軟な応答」を実現できる可能性があり、人狼知能研究の新たなフロンティアとして注目されている [2]。

しかし、LLM を人狼ゲームのような長期的な対話ゲームに適用する場合、「コンテキストウィンドウ (入力トークン数) の制限」と「情報の忘却・幻覚 (Hallucination)」が大きな課題となる。既存の単純なアプローチ (Full History 法) では、ゲームの進

行に伴って増大する会話ログをすべてプロンプトに入力するため、以下の問題が発生する。

- コストの増大: ターン経過に伴いトークン消費量が指数関数的に増加し運用コストが肥大化。
- 文脈維持の困難さ: 入力長大になると、LLM の注意機構 (Attention) が分散し、ゲーム序盤の重要な情報 (例: カミングアウトの結果や過去の投票行動) を忘却したり、矛盾した発言を行ったりするリスクが高まる [3]。

特に、数日間にわたる議論において「初日の発言と矛盾しない一貫性」を保つことは、現在の LLM にとっても容易ではない。

そこで本研究では、全会話履歴を入力する代わりに、「動的な記憶更新」メカニズムを導入した新たな人狼エージェントを提案する。本手法では、エージェントがターンごとに直近の会話を要約し、自身の「内部記憶」を動的に更新・圧縮して保持する。この処理を「動的記憶蓄積」と呼ぶ。次の発言生成時には、生の会話ログではなく、この圧縮された記憶のみを参照することで、以下の実現を目指す。

- 長期的な文脈維持: 重要な事実 (役職 CO, 処刑結果など) を要約して保持し続けることで、発言の論理的一貫性を向上させる。
- コスト効率化: 入力トークン数を抑え、長期戦においても安定した運用を可能にする。

本論文では、提案手法を実装したエージェントと、全会話履歴を読み込む従来型エージェント (ベースライン) を用いて比較実験を行う。5人村環境におけるシミュレーションを通じて、トークン消費量の削減効果 (コスト評価) および、投票行動の正確性 (定量評価) を検証する。さらに、生成された対話ログに対する主観評価 (定性評価) を行い、提案手法がエージェントの発言の一貫性や記憶の正

確性に与える影響について考察する。

2 関連研究

本章では、人狼ゲームにおける AI 研究の変遷および LLM を用いた対話エージェントの現状について概観する。その上で、既存手法における課題を整理し、本研究の立ち位置を明確にする。

2.1 LLM を用いた人狼エージェントの研究

近年、LLM を用いて、人間らしい振る舞いや高度な推論を行うエージェントの研究が盛んである。

Park らは、人間社会のシミュレーションを行う「Generative Agents」を提案した [4]。同手法では、エージェントの全経験を記憶ストリームに記録し、状況に応じて関連する記憶を検索する仕組みを採用している。さらに、検索した記憶から高次の洞察を導き出す内省のプロセスを導入することで、長期的な一貫性と人間らしい行動計画を実現している。

また、Yao らは、LLM の推論能力と行動決定を連携させる枠組みとして「ReAct」を提案している [5]。これは、エージェントが行動を起こす前に、その根拠となる思考過程を言語化して生成する手法である。思考と行動を交互に行うことで、外部情報の探索や、複雑なタスクにおける推論の修正が可能となり、LLM 単体よりも高い問題解決能力を示す。

2.2 ルールベースと LLM のハイブリッド手法

人狼ゲームのような不完全情報ゲームにおいては、嘘の生成や議論の制御といった高度な社会的スキルが求められる。LLM は柔軟な対話生成に優れる一方で、特定のルールを守ったり、戦略的な整合性を保ったりすることには課題が残る。Sato らは、この課題に対し、LLM とルールベースアルゴリズムを組み合わせたハイブリッドアーキテクチャを提案している [6]。彼らの手法では、会話履歴を分析し、状況に応じて LLM による生成と、事前に用意されたテンプレート発言（ルールベース）を動的に切り替えるアプローチをとっている。これにより、「特定のタイミングで反論する」「会話を適切に終了させる」といった制御が可能となり、エージェントの論理性やキャラクター性を担保している。

2.3 本研究の立ち位置

Park らのような検索・内省システムは強力であるが、記憶のデータベース構築や検索処理の実装コ

ストが高い。また、Sato らのハイブリッド手法は、ルールベースによる外部介入に依存しており、LLM 本来の柔軟性が制限される可能性がある。これに対し本研究は、外部データベースや複雑なルールセットを用いず、LLM 自身に「記憶の要約と更新」を行わせるという、よりシンプルかつ統合的なアプローチを提案するものである。これは、Yao らの ReAct のように LLM の推論能力を活用しつつ、Park らが目指した長期的な文脈維持を、より低コストな「記憶の圧縮」という形で実現する試みである。

3 提案手法

本章では、提案手法である「動的記憶蓄積」メカニズムの詳細について述べる。システム全体の構成、記憶更新のプロセス、および各エージェントの役割に応じたプロンプト設計について解説する。

3.1 概要

本研究では、全履歴を保持する代わりに、情報を圧縮・更新する「動的記憶蓄積」手法を提案する。

従来手法 (図 1) では、ターンが経過するごとに会話ログが増大する一方、提案手法 (図 2) では、エージェントに発言権が与えられたタイミングで、記憶の更新と行動決定の一連の処理を実行する。これにより、長期的な対話においても入力トークン数を抑制しつつ、文脈を維持した発言生成を実現する。

3.2 システム構成

提案システムは2つのフェーズで構成される。

記憶更新フェーズ：発言権が与えられると、まず「前回の自身の発言以降に追加された会話ログ（新規ログ）」と、現在保持している「内部記憶」を LLM に入力する。LLM はこれらの情報を統合・要約し、「更新された内部記憶」を生成する。この際、重要度の低い情報は削除され、重要な事実のみが圧縮された状態で継承される。

行動決定フェーズ：記憶の更新が完了した後、その「更新された内部記憶」のみを参照してプロンプトを作成する。これを LLM に入力し、現在のターンでの自身の行動（発言・投票）を生成し出力する。

このように、自身のターンが来るたびに「記憶を最新化してから発言する」サイクルを繰り返す。

3.3 プロンプト設計

エージェントの役割に応じた振る舞いを保証するため、各エージェントの初期記憶には「役職定義」と「戦略ガイドライン」が含まれる。特に狂人に対しては、記憶更新の段階から「自分は嘘をついて場を混乱させる役割である」ことを強く認識させるプロンプトを与え、対抗カミングアウトなどの高度な戦略を誘発する設計としている。

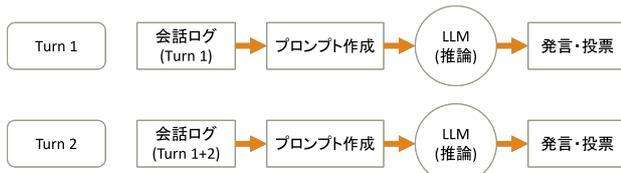


図1 ベースラインのコンテキスト処理

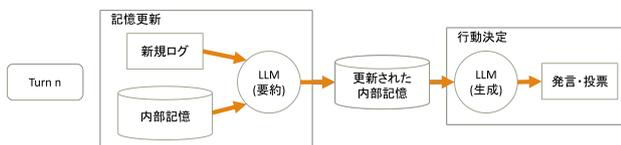


図2 提案手法のコンテキスト処理

4 実験設定

本節では、提案手法の有効性を検証するために行ったシミュレーション実験について述べる。

4.1 シミュレーション環境

提案手法の有効性を検証するため、以下の設定で人狼ゲームのシミュレーションを行った。

提案手法の有効性を検証するため、人狼知能プラットフォームをベースとしたテキストチャット環境でシミュレーションを行った。村構成は5人(村人2, 占い師1, 狂人1, 人狼1)とし、それぞれをLLM (GPT-4o-mini) によって制御される自律型エージェントとして実装した。試行回数は、乱数シードをランダムに設定し、各手法につき20ゲーム(計40ゲーム)を実施した。

4.2 評価指標

本実験では、以下の3つの観点から評価を行う。なお、5人村における勝敗はランダム要素の影響を強く受ける為、本研究では勝率評価は行わず、エージェントの振る舞いの質と効率性に焦点を当てる。

1. トークン消費量 (コスト評価) : 1ゲームあたりの平均総トークン数, およびターン経過に伴う

トークン増加量を比較する。

2. 投票行動の正確性 (定量評価) : 村人陣営のエージェントが、人狼に対して正しく投票できた割合を計測する。これにより、情報を圧縮しても推論能力が維持されているかを検証する。
3. 対話品質の主観評価 (定性評価) : 生成された対話ログを被験者(大学生14名)に提示し、「論理的一貫性」「記憶の正確性」「自然さ」等の項目について5段階で評価を行う。具体的な質問項目を表1に示す。

表1 主観評価におけるアンケート質問項目(全項目において1:低評価~5:高評価の5段階リッカート尺度を使用)

評価項目	質問内容
自然さ	エージェントの会話は人間らしく自然でしたか?
論理的一貫性	発言内容に矛盾や支離滅裂な点は見られませんでしたか?
記憶の正確性	過去のCO結果や処刑結果などの事実を正しく覚えていましたか?
役職行動	役職に応じた適切な行動(COや投票)が取れていましたか?
意図の明確さ	その発言をした意図や理由が理解できましたか?

以下の2手法を比較対象とする。

- 手法A (Baseline) : ゲーム開始時からの全会話履歴を毎回プロンプトに入力する「全読込型」。
- 手法B (Proposed) : 本研究で提案する「記憶蓄積型」。

5 実験結果

本節では、実験結果について述べる。

5.1 トークン消費量の削減効果

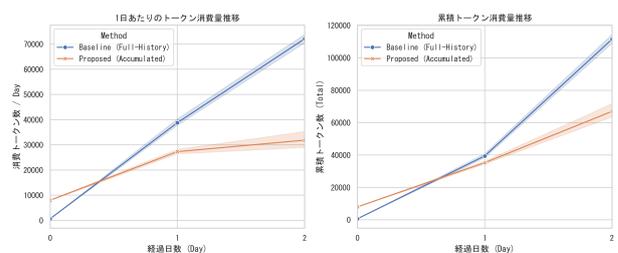


図3 トークン消費量

図3(左)に、ゲーム経過日数ごとの1日あたり

のトークン消費量の推移を示す。ベースライン手法 (Baseline) は、日数が経過するごとに会話ログが蓄積されるため、トークン消費量が線形に増加していることが確認できる。一方、提案手法は、記憶の要約・更新を行うことで、Day 1 以降は消費量が [約 30,000 トークン] 前後でほぼ横ばいに推移している。また、図 3 (右) の累積トークン量においても、提案手法はベースラインと比較して大幅なコスト削減を実現していることがわかる。これにより、提案手法は長期的なゲームにおいてもコストを一定水準に抑え、持続的な運用が可能であることが定量的に示された。

5.2 投票行動の正確性

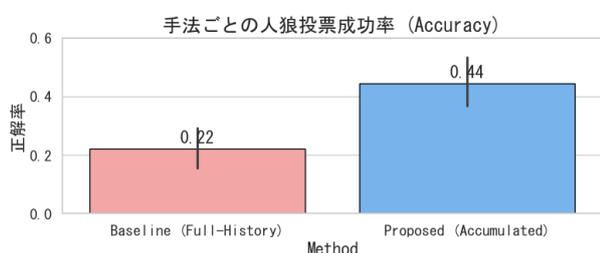


図 4 投票成功率

図 4 に、村人陣営による人狼への投票成功率を示す。実験の結果、ベースライン手法の正解率が 0.22 であったのに対し、提案手法は 0.44 を記録した。情報を要約・圧縮したにもかかわらず、提案手法はベースラインの 2 倍の正解率を達成している。これは、不要な情報が削除され、推論に必要な重要情報が整理された状態で LLM に入力されたことで、推論精度が向上した可能性を示唆している。

5.3 主観評価による対話品質

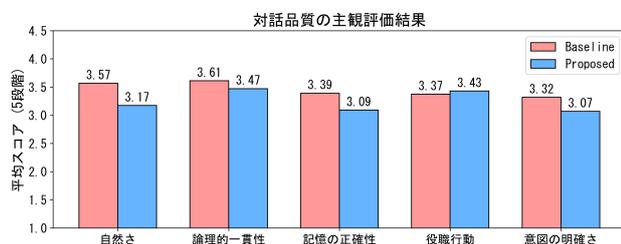


図 5 主観評価

図 5 に、被験者によるアンケート評価の平均スコアを示す。「役職行動」においては両手法ともに 3.4 と同等の評価を得ており、提案手法は情報の圧縮を行いつつも、ゲームプレイヤーとしての役割を十分に果たしていることがわかる。一方で、「自然

さ」や「記憶の正確性」等の項目では、ベースライン手法がわずかに高いスコア (自然さ: Baseline 3.6, Proposed 3.2) を示した。これは、要約の過程で会話の微細なニュアンスや文脈の一部が欠落した影響と考えられる。しかし、いずれの項目も 3.0 以上の評価を維持しており、対話エージェントとして破綻のない品質は保たれているといえる。

6 考察

本節では、実験結果をふまえた考察を述べる。

6.1 記憶圧縮と推論精度の関係

提案手法がコストを削減しつつ投票精度を向上させた点は特筆すべきである。Baseline では情報過多により重要手がかりが埋没する "Lost in the Middle" 現象が生じたのに対し、提案手法では記憶更新時に重要情報 (疑い先や矛盾点) が要約・保持されたため、LLM の推論リソースが核心に集中し、正解率向上に繋がったと考察できる。

6.2 実用性と対話品質のトレードオフ

「自然さ」の低下は、要約により文脈の機微が欠落したためと推察される。しかし「役職行動」は維持されており、勝敗に直結する戦略的振る舞いは機能していた。人狼 AI のような目的指向型システムでは、流暢さよりもコストと勝率の両立が重要視されるため、本手法は実用的なアプローチである。

6.3 課題と今後の展望

ログ分析の結果、CO や処刑結果が即座に記憶へ反映されない事例が確認された。これは記憶更新が「自身の発言ターン」にのみ実行される仕様に起因する。本実験のような短期決戦では致命傷に至らなかったが、長期戦では戦略ミスを誘発するリスクがある。今後は、重要イベント発生時に記憶を強制更新する「イベントトリガー型」メカニズムの導入が必要である。

7 おわりに

本研究では、文脈維持とコスト削減を両立させる「動的記憶蓄積」手法を提案した。実験の結果、トークン消費量を劇的に削減しつつ、投票成功率を倍増させることに成功した。今後は、認識遅延の課題を解決するイベントトリガー型更新の導入や、より複雑な環境での検証を進めたい。

参考文献

- [1] Fujio Toriumi, et al. : “AI Wolf Contest — Development of Game AI Using Collective Intelligence —”, Communications in Computer and Information Science ((CCIS, volume 705)) pp. 101–115
- [2] Jason Wei, et al. : “Chain-of-thought prompting elicits reasoning in large language models”, NIPS’22: Proceedings of the 36th International Conference on Neural Information Processing Systems Article No.: 1800, pp. 24824–24837
- [3] Nelson F. Liu, et al. : “Lost in the Middle: How Language Models Use Long Contexts”, Transactions of the Association for Computational Linguistics (2024) 12: 157–173.
- [4] Joon Sung Park, et al. : “Generative Agents: Interactive Simulacra of Human Behavior”, UIST ’23 : Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology Article No.: 2, pp. 1–22.
- [5] Shunyu Yao, et al. : “REACT: SYNERGIZING REASONING AND ACTING IN LANGUAGE MODELS”, 11th International Conference on Learning Representations, ICLR 2023.
- [6] Takehiro Sato, et al. : “An Implementation of Werewolf Agent That does not Truly Trust LLMs”, In Proceedings of the 2nd International AIWolfDial Workshop, pp. 58–67.