

長期タスクにおけるLLMエージェントの自律的文脈管理能力の学習

[S01-22] ~LLMエージェントは“賢く捨てる”ことを学習できるか？~

KUMDEE THANASAN¹ 馬 尤咪¹ 岡崎 直観^{1,2,3} 1) 東京科学大学, 2) 産総研, 3) NII LLMC
thanasan.kumdee@nlp.comp.isct.ac.jp {ma.y@, okazaki@}comp.isct.ac.jp



① 背景

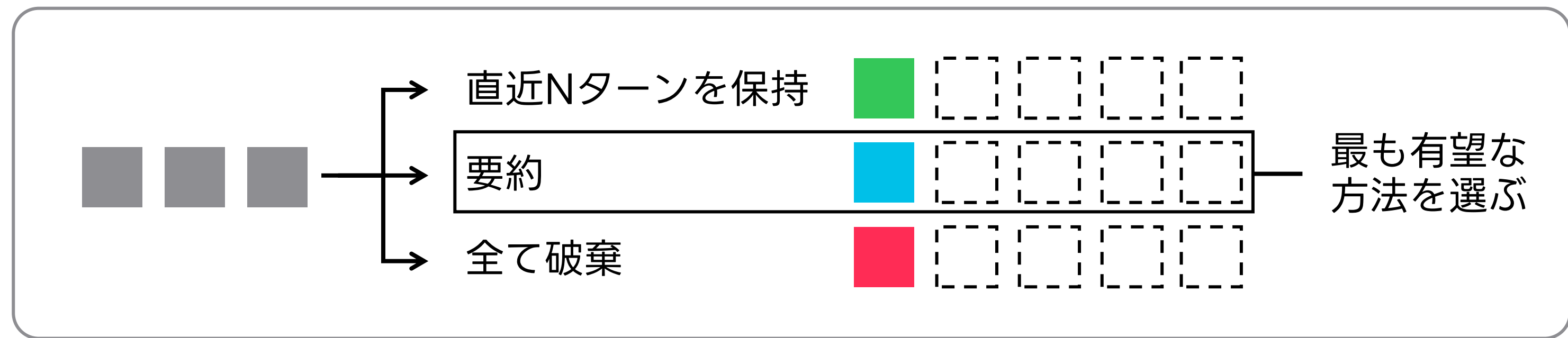
- LLMエージェントの応用場面として、ソフトウェア開発などの長い文脈を扱う「**長期タスク**」が主流になりつつある
- LLMが一気に処理できる文脈は限られているため、**常に文脈を管理する必要がある**
- 最適な文脈管理方法は過去の文脈の状態に依存する
 - 重要な情報が入っているなら「要約」
 - 不要な情報しかないなら「全て破棄」

③ 本研究のアイデア

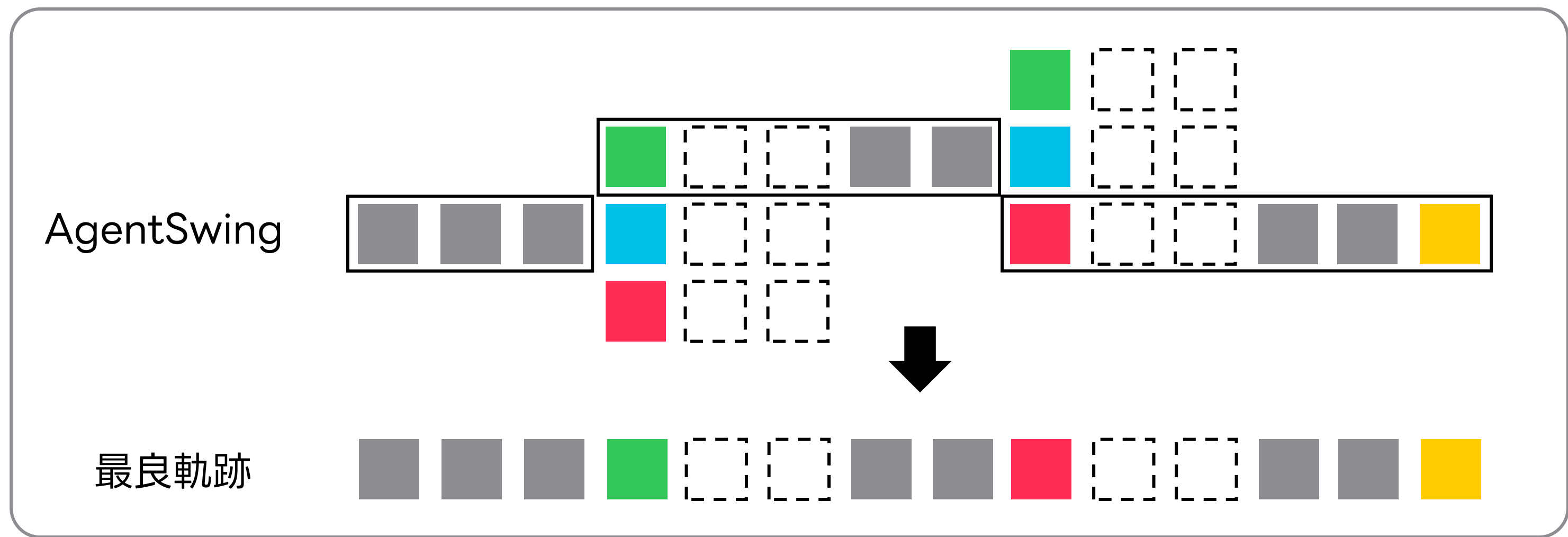
- AgentSwingは動的に文脈を管理し、**タスクの成功率を大幅に向上させる**一方、推論時の**コストが高い**
- 本研究では、並列試行により判明された「最も有望な文脈管理方法」を、試行せずに選択できるモデルを作る

⑤ アプローチ

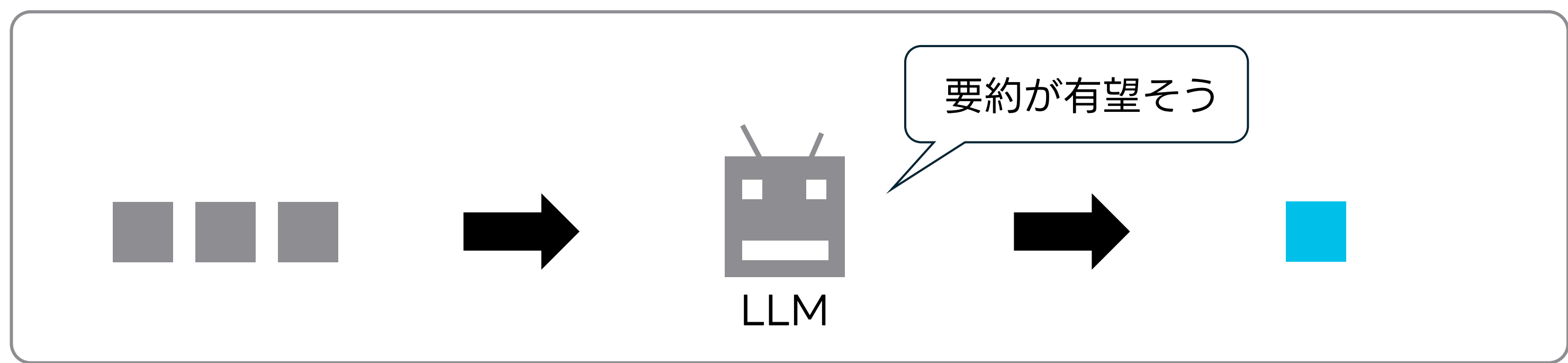
1. AgentSwingの再現



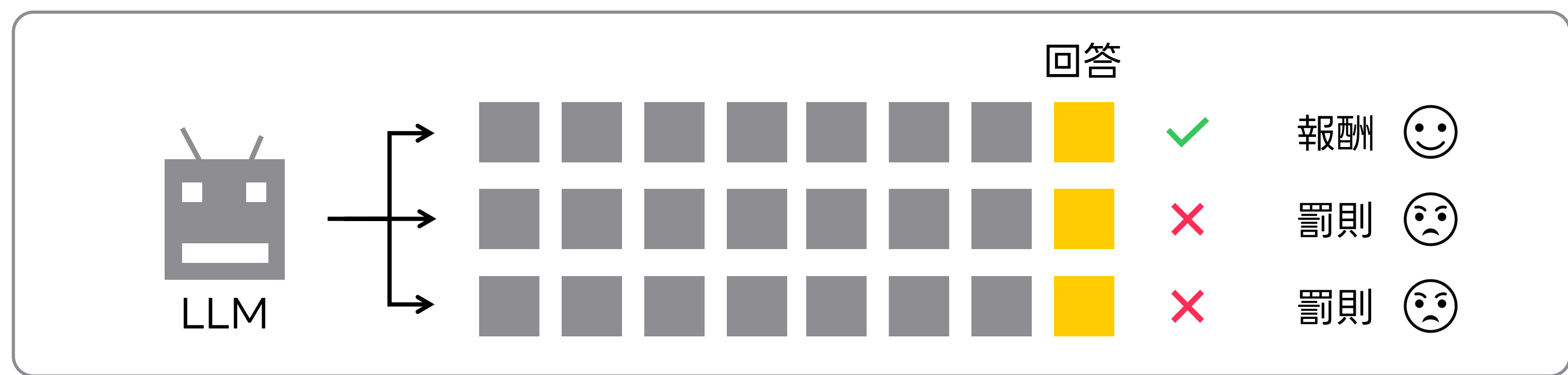
2. 最良軌跡の収集



3. SFTによる文脈管理のツール使用と管理方法選択の学習



4. RLによるタスク成功率の向上



② 先行研究

- エージェントは「思考 → 行動 → 観察」のステップを繰り返し、文脈を蓄積する [1]
- AgentSwingはエージェントの文脈を管理する最新の手法で、文脈がある程度蓄積されたら以下を行う [2]
 - 3つの文脈管理方法を並列に分岐して実行
 - それぞれ数ステップだけ試行して管理方法の有望度合いを確認
 - 最も有望な分岐だけを採用して他は破棄

④ 本研究の期待される貢献

賢い「捨てる」を学習する方法論

いつ・何を捨てるかという判断をLLM・エージェントに学習させる方法論の確立

文脈を自律的に管理できるLLM

あらゆるエージェントシステムに組み込める

⑥ 現在の進捗

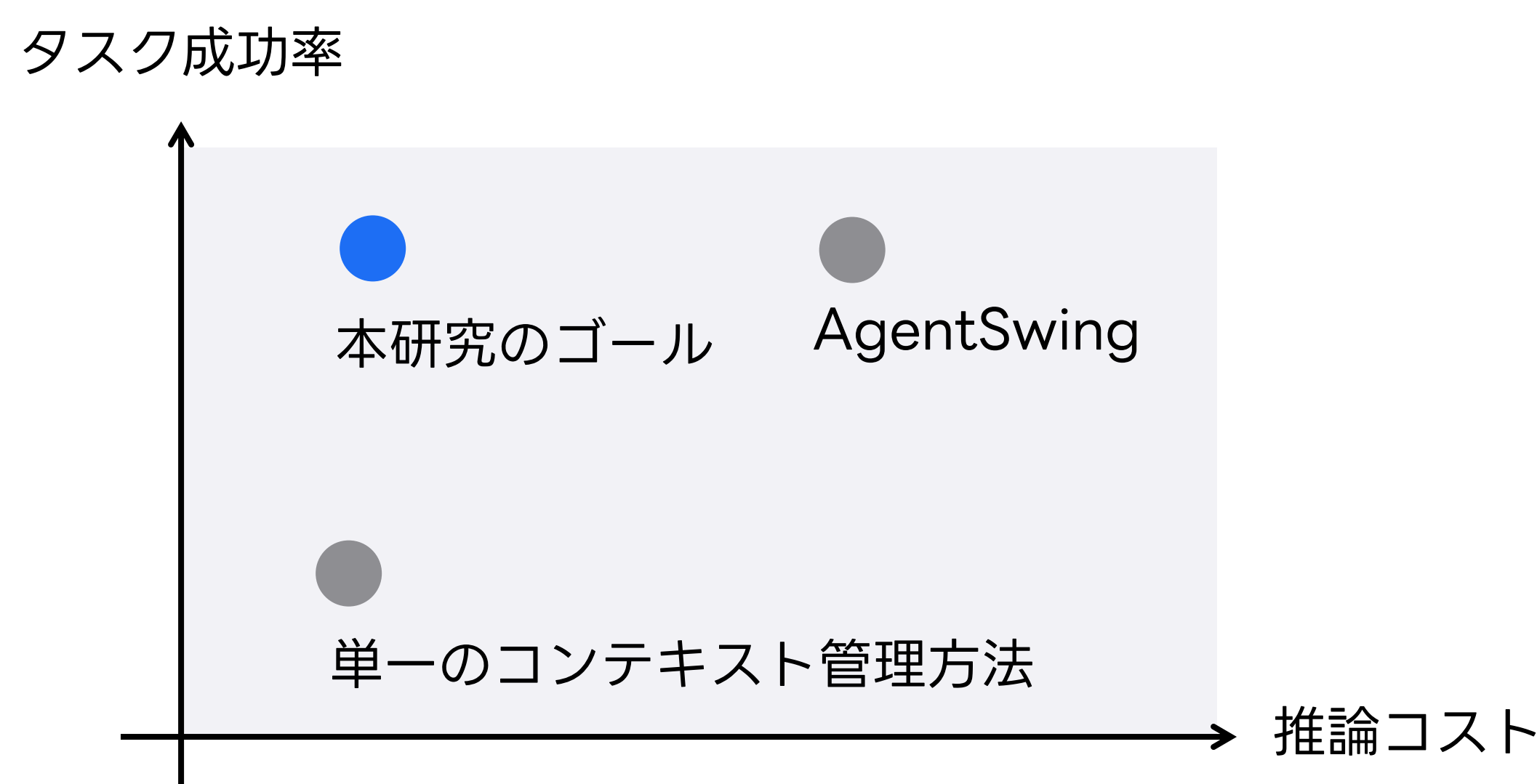
- AgentSwingの再現実験
 - そもそもウェブ検索タスクは再現しにくいのか？
 - まずは「全て破棄」の成功率を上げる方法を考える

モデル	文脈管理方法	BrowseComp	
		再現	論文
GPT-OSS-120B	ベースライン（なし）	32	39.5
	直近Nターンを保持	38.5	52.5
	要約	36	48
	全て破棄	27.5	50.5
	AgentSwing	34.5	60

- SFTの練習を行なっている

⑦ 期待される成果

- 試行しなくても文脈を適切に管理し、タスクを遂行できるモデルを作る
- 単一のコンテキスト管理手法に近い推論コストで、AgentSwingに近い成功率を達成できる



謝辞：この成果は、NEDO（国立研究開発法人新エネルギー・産業技術総合開発機構）の委託業務（JPNP26013）の結果得られたものです。
[1] : Shunyu Yao et al. ReAct: Synergizing Reasoning and Acting in Language Models. ICLR 2023. 2023.
[2] : Zhaopeng Feng et al. AgentSwing: Adaptive Parallel Context Management Routing for Long-Horizon Web Agents. arXiv:2603.27490. 2026.