

Can LLMs Learn from Previous Mistakes? Investigating LLMs' Errors to Boost for Reasoning

Yongqi Tong, Dawei Li, Sizhe Wang, Yujia Wang, Fei Teng, Jingbo Shang

ACL 2024

<https://aclanthology.org/2024.acl-long.169>

ACL読み会2024

名古屋大学 笹野研究室

M2 大鹿 雅史

概要

- LLMが過去の推論の誤りから推論の修正を行うことができるか調査
- Promptingとfine-tuningの両方の側面について調査
- 実験結果から過去の誤りを修正するように推論することで推論能力が向上

■ 選定理由

- LLMがどの程度人間らしく振る舞うかは注目すべき
- 人間の学習を模倣することで色々な発展ができそう

背景

- Chain-of-Thought(CoT)^[1]は人間のような論理で推論を行うことができる
 - CoT : "Let's think step-by-step" というプロンプトをもとに段階的な推論を行う
- 過去の誤りから学び同じ間違いをしないようにするのは人間の学習の特徴の一つ
 - LLMが同様な学習をできるかについて調査はされていない

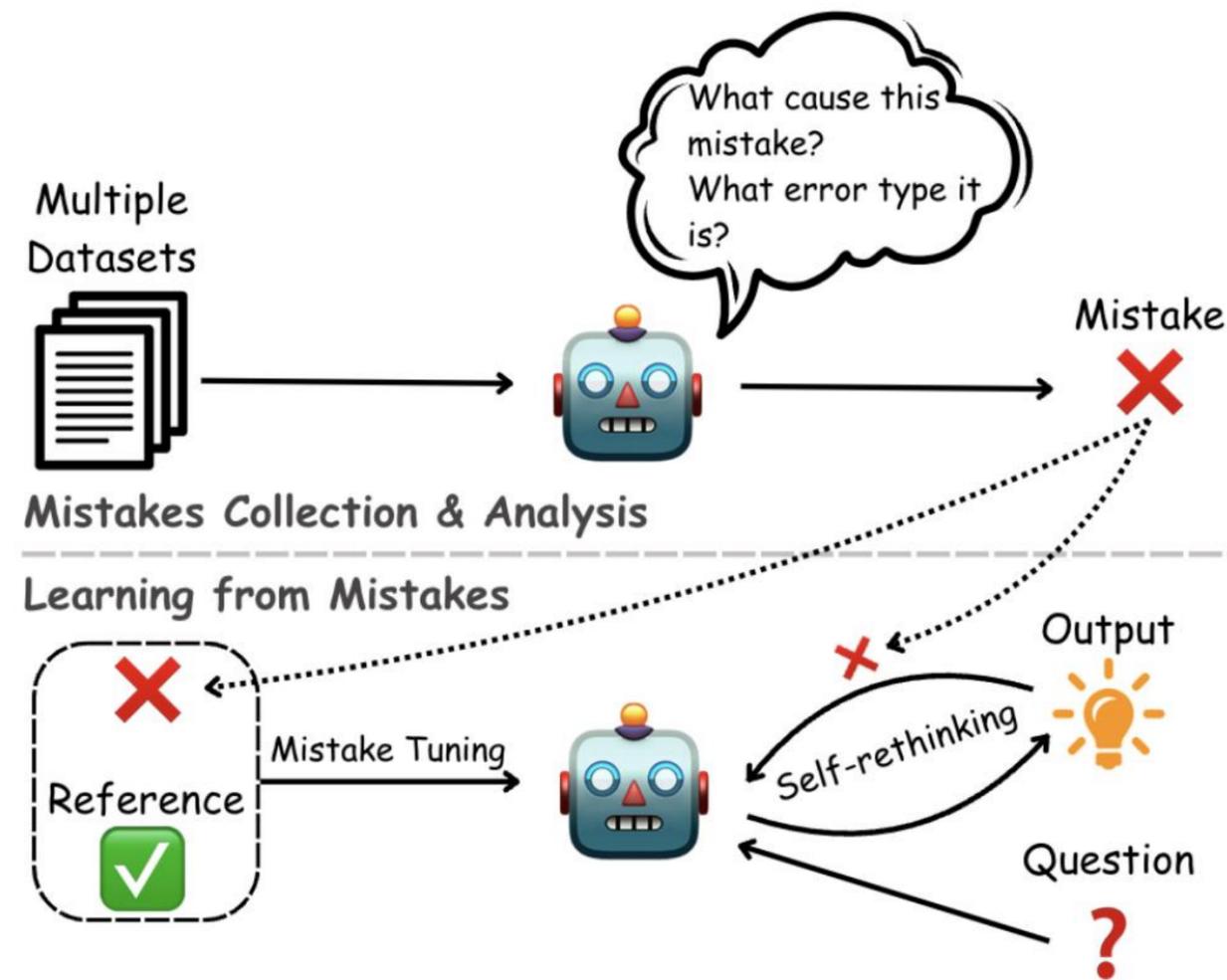
LLMは過去の誤りから学習し推論能力を向上することができるか？

本論文の貢献

1. データセット (COTERRORSET) の構築

2. 過去の誤りを用いた推論

- ① Self-rethinking
- ② Mistake Tuning

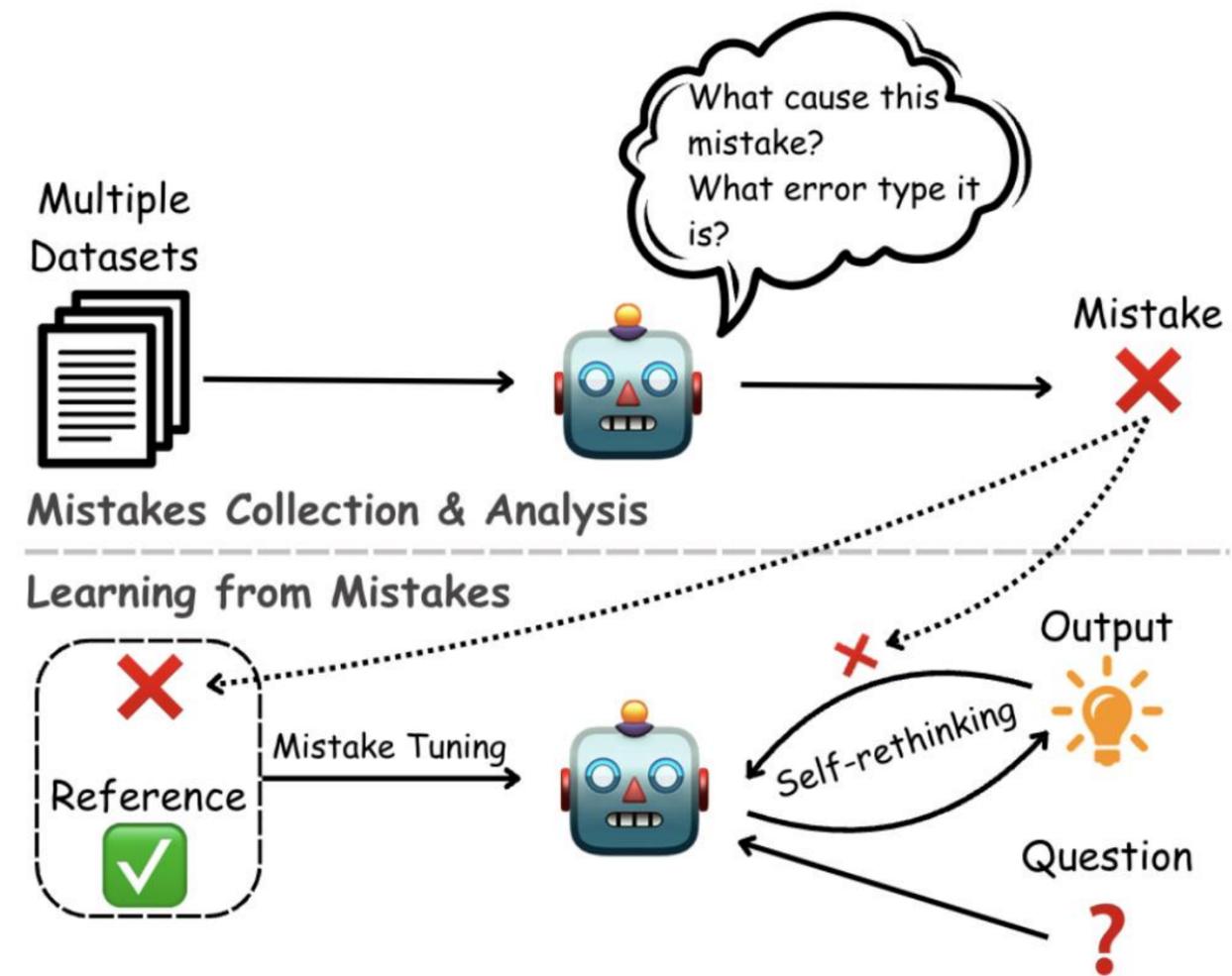


本論文の貢献

1. データセット (COTERRORSET) の構築

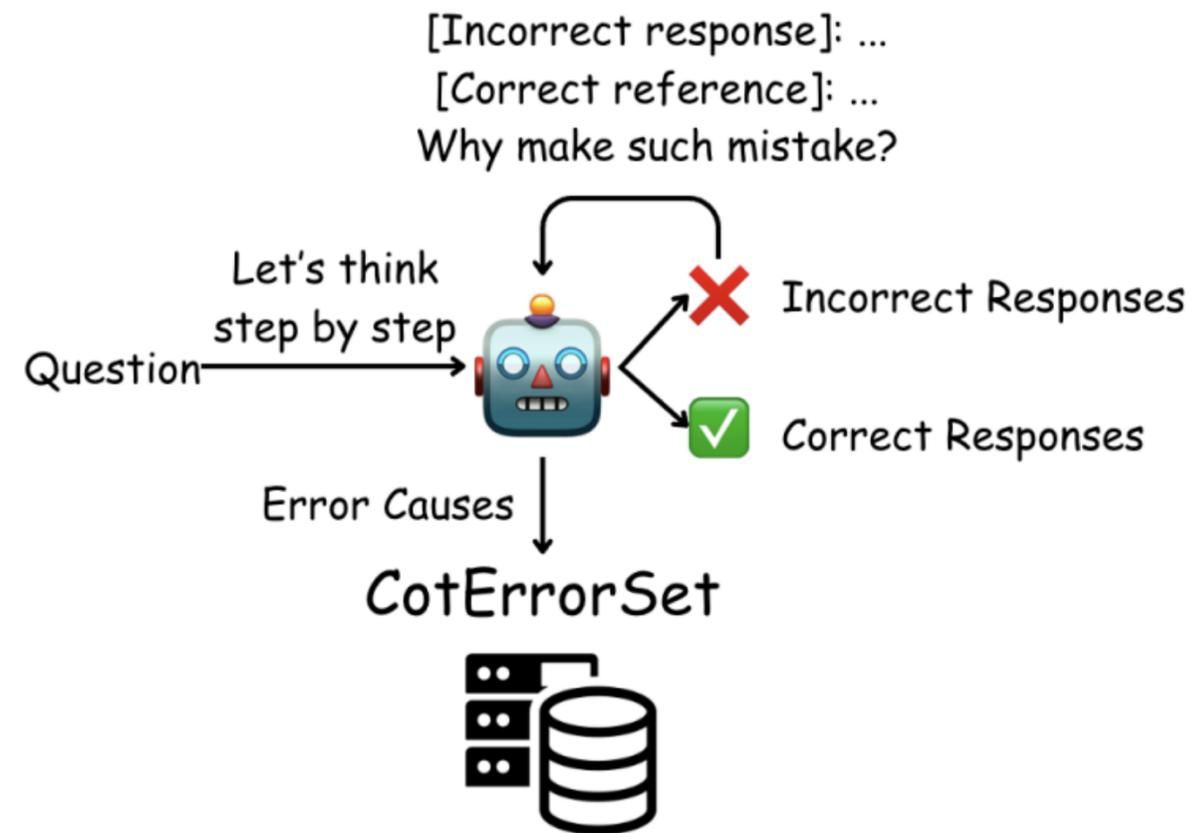
2. 過去の誤りを用いた推論

- ① Self-rethinking
- ② Mistake Tuning



データセット (CoTERRORSET) の構築手順

1. 既存データセットの推論タスクをPaLM2に回答
2. 間違えた回答に対してなぜ間違えたかをPaLM2で生成
3. 誤りの種類によってラベル付け

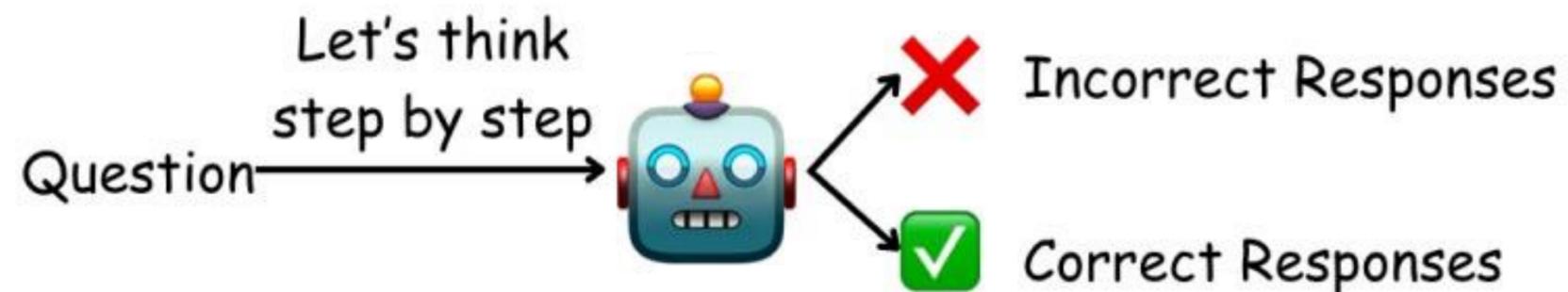


CoTERRORSET①

■ CoT-COLLECTIONのデータをPaLM2に回答

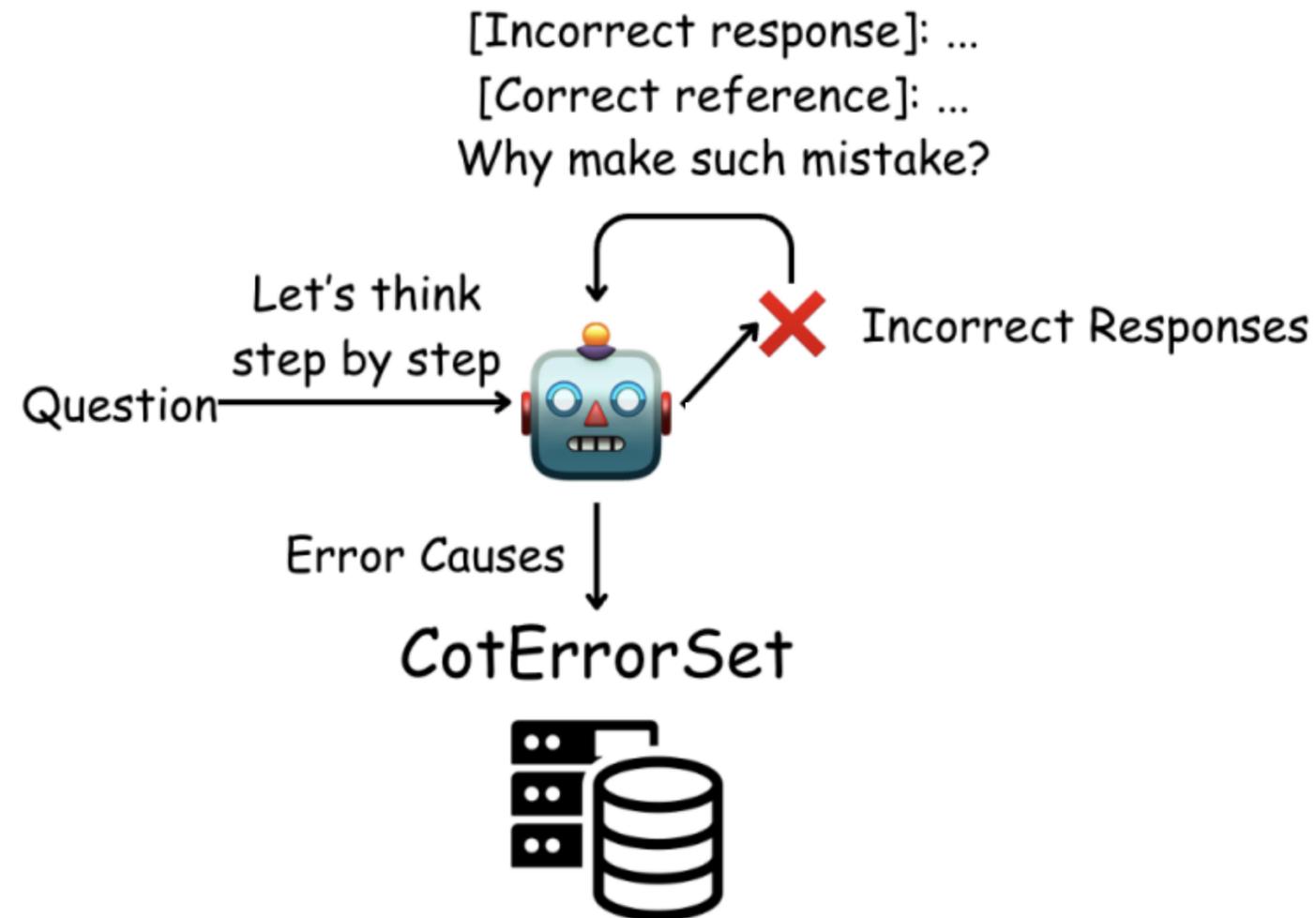
- CoT-COLLECTION：QAや論理推論、算術推論などのタスクを含んだデータセット
質問、回答、正しい回答への論理が組になっている

■ CoTを利用することで推論過程を生成



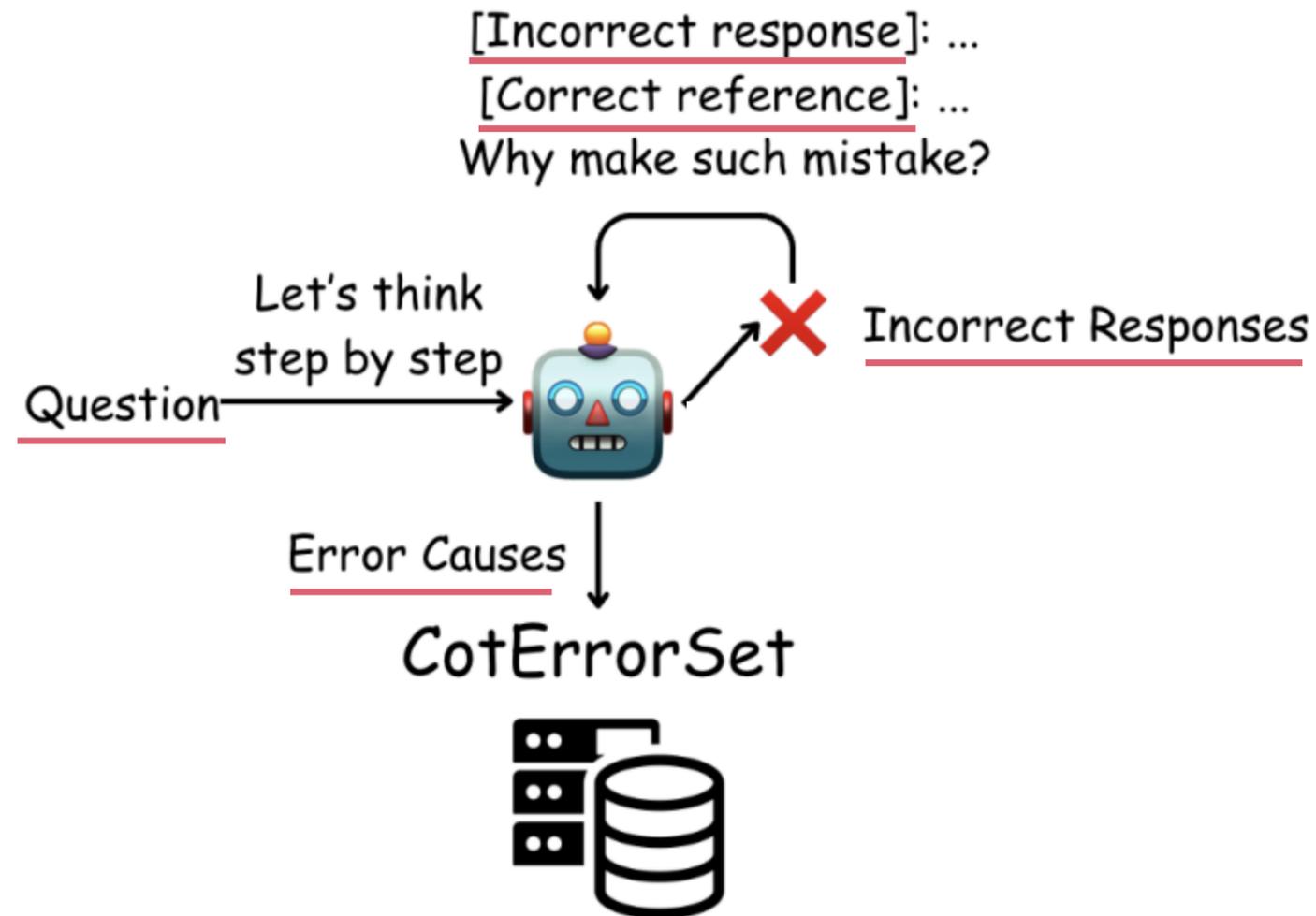
CoTERRORSET②

- 間違えた回答と正解を入力としてPaLM2でなぜ間違えたかを生成
- 質問、回答、正しい論理、誤った論理、誤った原因の組み合わせを構築



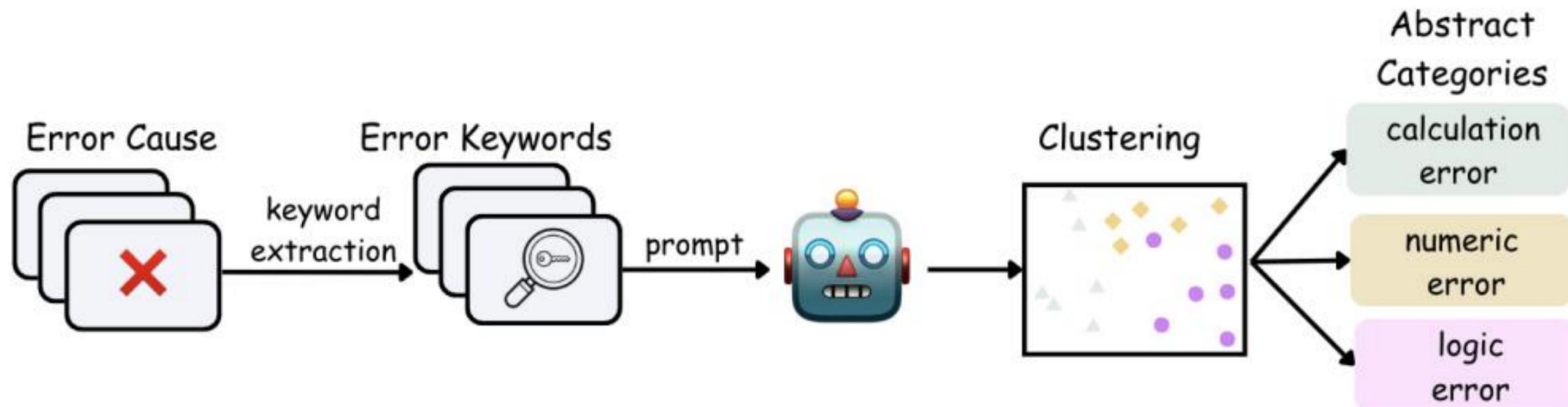
CoTERRORSET②

- 間違えた回答と正解を入力としてPaLM2でなぜ間違えたかを生成
- 質問、回答、正しい論理、誤った論理、誤った原因の組み合わせを構築



CoTERRORSET③

- データセットを誤りの種類についてLLMを利用して分類
- 異なる種類の誤りをPromptに用いることで効率よく誤りを検出

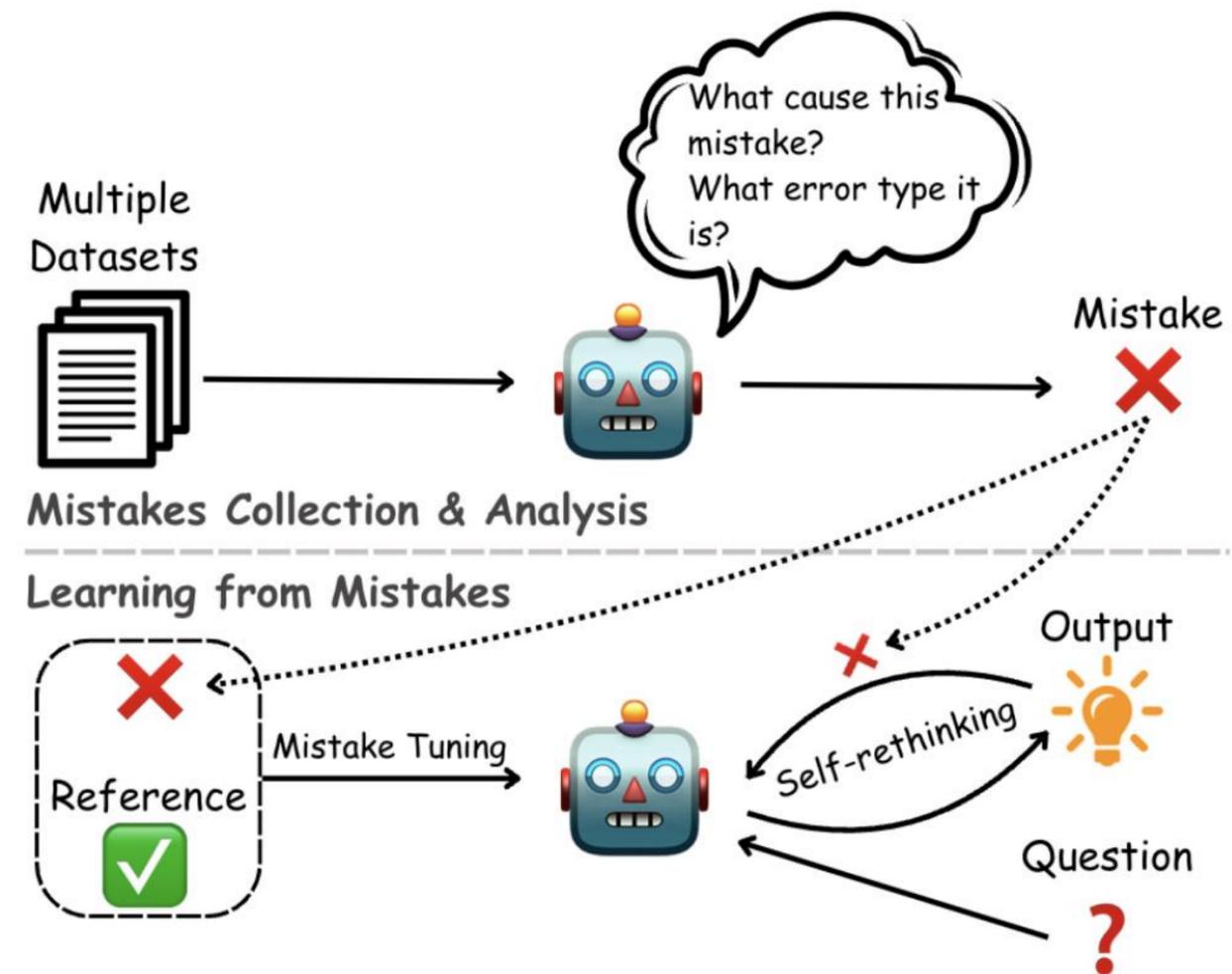


本論文の貢献

1. データセット (COTERRORSET) の構築

2. 過去の誤りを用いた推論

- ① Self-rethinking
- ② Mistake Tuning



提案手法 1 : Self-rethinking

1. 質問に対して回答を生成させる
2. 過去の誤りの中に同じ種類の誤りがあるかを判定
3. ない場合→終了、あった場合→再回答を行う

Step 1: CoT reasoning

- *Question:* Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
- *Answer:* Let's think step by step.
- **Response:** Natalia sold $48 * 2 = 96$ clips in May. Natalia sold $48+96 = 144$ clips altogether in April and May.

Step 2: self-rethinking

- *Question:* Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
- *Your output:* Natalia sold $48 * 2 = 96$ clips in May. Natalia sold $48+96 = 144$ clips altogether in April and May.
- *Do you make similar mistakes with the following examples:* ###Error Type 1: Misapplication of Algebraic Identities: ...
- **Response:** Yes, I make a mistake.

Step 3: correction (if they think they make mistakes)

- *So the correct answer is:*
- **Response:** Natalia sold $48/2 = 24$ clips in May. Natalia sold $48+24 = 72$ clips altogether in April and May.

提案手法 1 : Self-rethinking

1. 質問に対して回答を生成させる
2. 過去の誤りの中に同じ種類の誤りがあるかを判定
3. ない場合→終了、あった場合→再回答を行う

Step 1: CoT reasoning

- *Question:* Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
- *Answer:* Let's think step by step.
- **Response:** Natalia sold $48 * 2 = 96$ clips in May. Natalia sold $48+96 = 144$ clips altogether in April and May.

Step 2: self-rethinking

- *Question:* Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
- *Your output:* Natalia sold $48 * 2 = 96$ clips in May. Natalia sold $48+96 = 144$ clips altogether in April and May.
- *Do you make similar mistakes with the following examples: ###Error Type 1: Misapplication of Algebraic Identities: ...*
- **Response:** Yes, I make a mistake.

Step 3: correction (if they think they make mistakes)

- *So the correct answer is:*
- **Response:** Natalia sold $48/2 = 24$ clips in May. Natalia sold $48+24 = 72$ clips altogether in April and May.

提案手法 1 : Self-rethinking

1. 質問に対して回答を生成させる
2. 過去の誤りの中に同じ種類の誤りがあるかを判定
3. ない場合→終了、あった場合→再回答を行う

Step 1: CoT reasoning

- *Question:* Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
- *Answer:* Let's think step by step.
- **Response:** Natalia sold $48 * 2 = 96$ clips in May. Natalia sold $48+96 = 144$ clips altogether in April and May.

Step 2: self-rethinking

- *Question:* Natalia sold clips to 48 of her friends in April, and then she sold half as many clips in May. How many clips did Natalia sell altogether in April and May?
- *Your output:* Natalia sold $48 * 2 = 96$ clips in May. Natalia sold $48+96 = 144$ clips altogether in April and May.
- *Do you make similar mistakes with the following examples:* ###Error Type 1: Misapplication of Algebraic Identities: ...
- **Response:** Yes, I make a mistake.

Step 3: correction (if they think they make mistakes)

- *So the correct answer is:*
 - **Response:** Natalia sold $48/2 = 24$ clips in May. Natalia sold $48+24 = 72$ clips altogether in April and May.
-

提案手法 2 : Mistake Tuning

■ 正解論理と誤った論理の2種類でLLMをtuningする

- Prefixを元にどちらのデータを識別

■ 入出力

- 入力

正しい論理のデータ : 質問 + “CORRECT RATIONALE”

誤った論理のデータ : 質問 + “INCORRECT RATIONALE”

- 出力

論理 + 回答

実験

■ ベースライン

- Standard Prompting
- CoT
- Self-refine : 自身のfeedbackを元に回答を修正 (過去の誤りを利用していない)
- Self-consistency : 複数の推論経路から最も多かった回答を選ぶ手法

■ ベンチマーク

- 6つの数学タスク (GSM8K, AQuA, MathQA, OpenbookQA, LogiQA, Critical Reasoning)

■ 利用モデル

- PaLM2, GPT4, Flan-T5

実験結果①：Self-rethinking

- 他のベースライン手法に比べ多くのデータセットで高いスコア
- GPT4でも同様に高性能を示した

Methods	GSM8K	AQuA	MathQA	OpenbookQA	LogiQA	CR
Standard Prompting (Brown et al., 2020)	17.06	22.40	27.57	80.92	41.21	24.45
CoT (Madaan et al., 2023)	56.29	32.11	30.89	82.66	41.05	51.98
Self-refine (Madaan et al., 2023)	34.74	39.92	54.01	28.75	35.99	12.28
Self-consistency (Wang et al., 2022)	58.38	42.80	41.37	87.61	42.88	22.58
Self-rethinking (Ours)	65.13	44.72	43.95	87.71	49.12	54.53

実験結果①：Self-rethinking

- 他のベースライン手法に比べ多くのデータセットで高いスコア
- GPT4でも同様に高性能を示した

→ 過去の誤りから推論を修正することでより正確な回答が可能になる

Methods	GSM8K	AQuA	MathQA	OpenbookQA	LogiQA	CR
Standard Prompting (Brown et al., 2020)	17.06	22.40	27.57	80.92	41.21	24.45
CoT (Madaan et al., 2023)	56.29	32.11	30.89	82.66	41.05	51.98
Self-refine (Madaan et al., 2023)	34.74	39.92	54.01	28.75	35.99	12.28
Self-consistency (Wang et al., 2022)	58.38	42.80	41.37	87.61	42.88	22.58
Self-rethinking (Ours)	65.13	44.72	43.95	87.71	49.12	54.53

実験結果②：Mistake Tuning

- Mistake tuningは通常のfine-tuningよりも高い性能を示している
- 誤った論理のデータも明示的に利用することで推論に活かされる

Models	Methods	GSM8K	MathQA	AQuA
Flan-T5-large (780M)	Standard finetuning	14.28	42.79	13.10
	Mistake tuning	18.36	48.95	18.07
Flan-T5-xl (3B)	Standard finetuning	23.81	47.24	17.81
	Mistake tuning	24.29	52.22	20.99

実験結果②：Mistake Tuning

- Mistake tuningは通常のfine-tuningよりも高い性能を示している
- 誤った論理のデータも明示的に利用することで推論に活かされる

→ 誤りも利用できればデータセットの構築が容易になる

Models	Methods	GSM8K	MathQA	AQuA
Flan-T5-large (780M)	Standard finetuning	14.28	42.79	13.10
	Mistake tuning	18.36	48.95	18.07
Flan-T5-xl (3B)	Standard finetuning	23.81	47.24	17.81
	Mistake tuning	24.29	52.22	20.99

考察

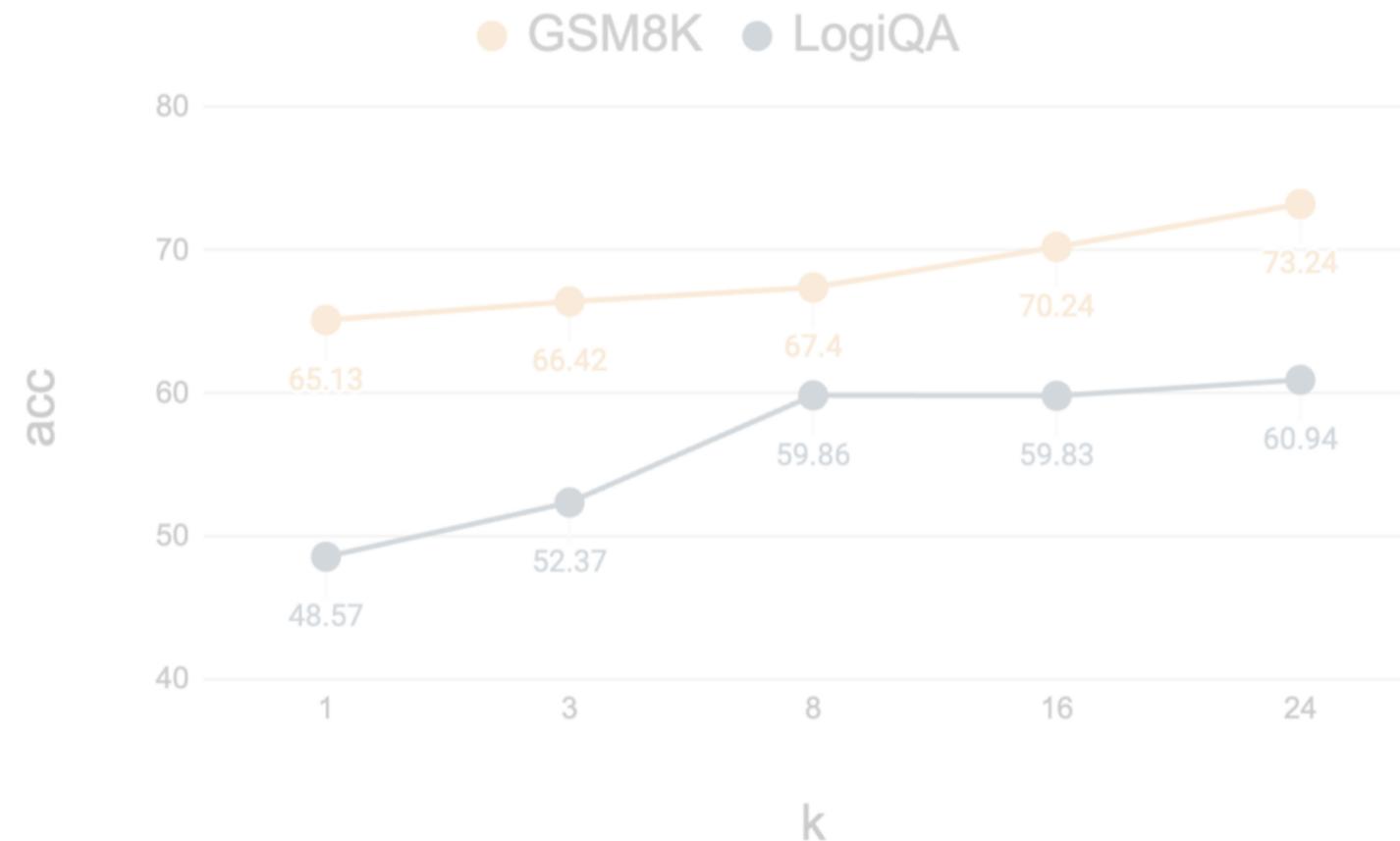
- 反復回数による性能差
- LLMによる推論の誤りの種類
 - データセット作成時に分類した誤りの種類について調査
 - 常識推論タスクと算術推論タスクについて調査

反復回数による性能差

■ Self-rethinkingの過程を反復することでAccuracyがどう変化するかを調査

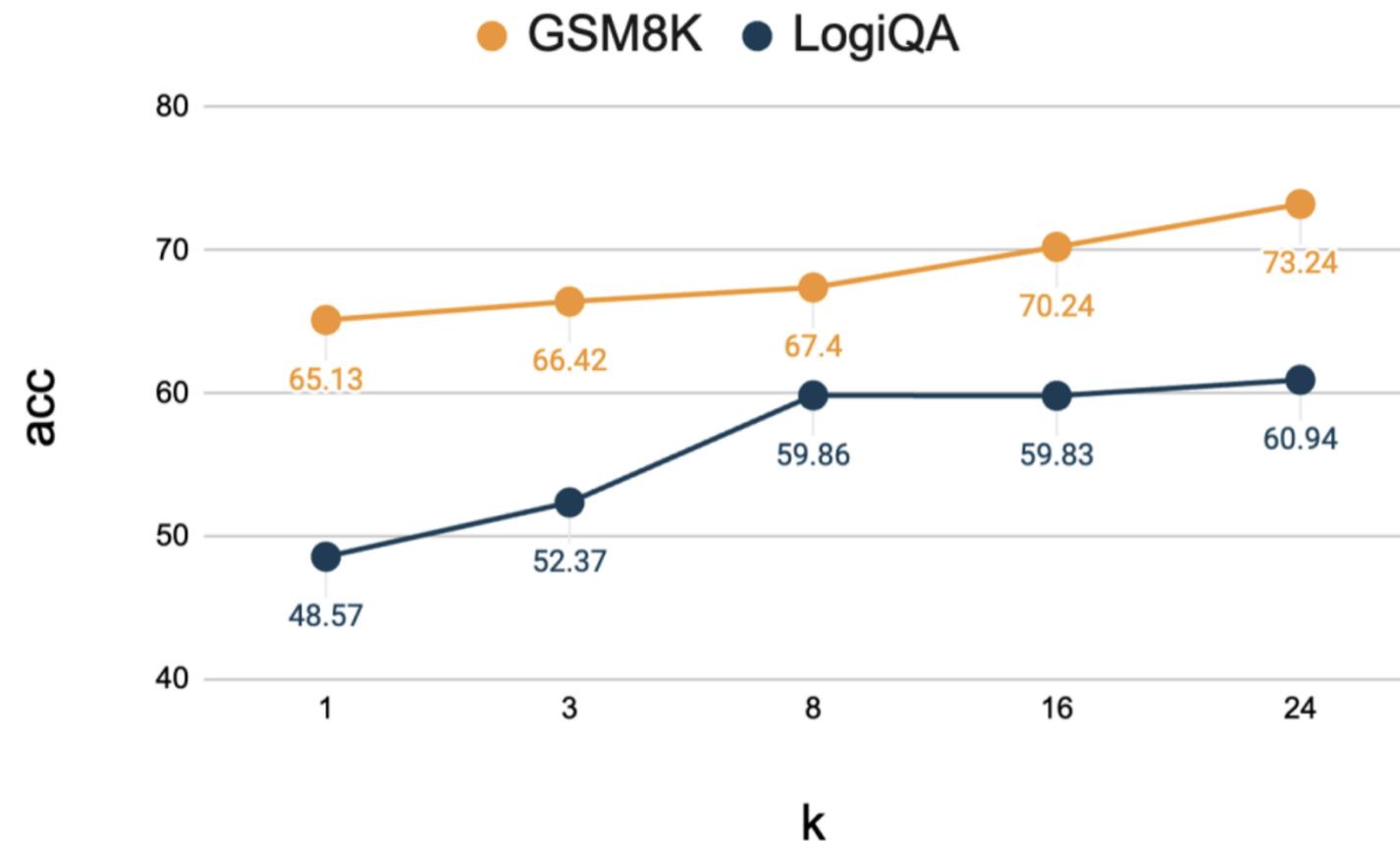
→反復を行うことでAccuracyは向上する

(→当たり前という印象も受ける)



反復回数による性能差

- Self-rethinkingの過程を反復することでAccuracyがどう変化するかを調査
 - 反復を行うことでAccuracyは向上する
 - (→当たり前という印象も受ける)



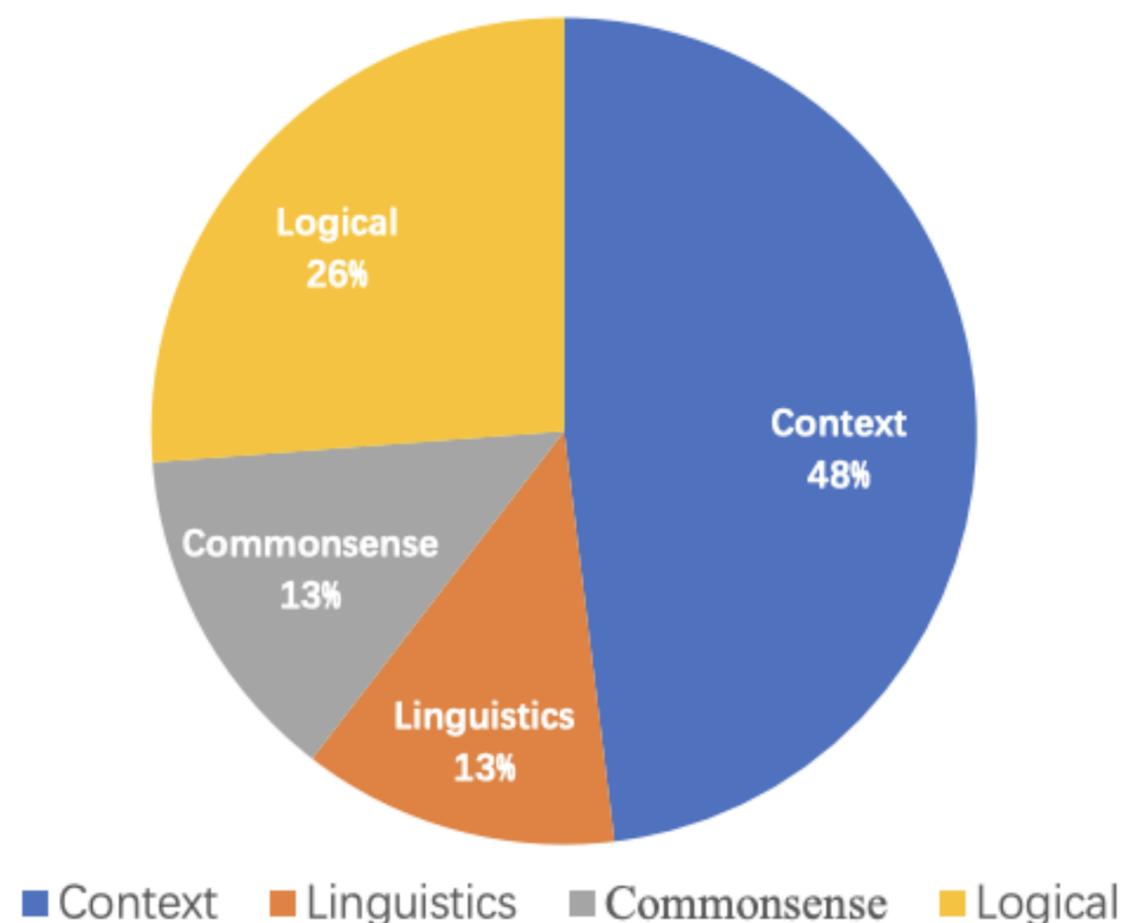
LLMによる推論の誤りの種類：誤りの種類

- 誤りの種類は6種類に分類される
 1. Calculation：計算を行う際の誤り
 2. Numeric：数値の表現の誤り（桁数ミスなど）
 3. Logical：回答への論理の誤り
 4. Linguistics：問題の誤解による誤り
 5. Commonsense：常識に反する推論をした誤り
 6. Context：文脈を誤解した誤り

考察②：LLMによる推論の誤りの種類

■ 常識推論タスク

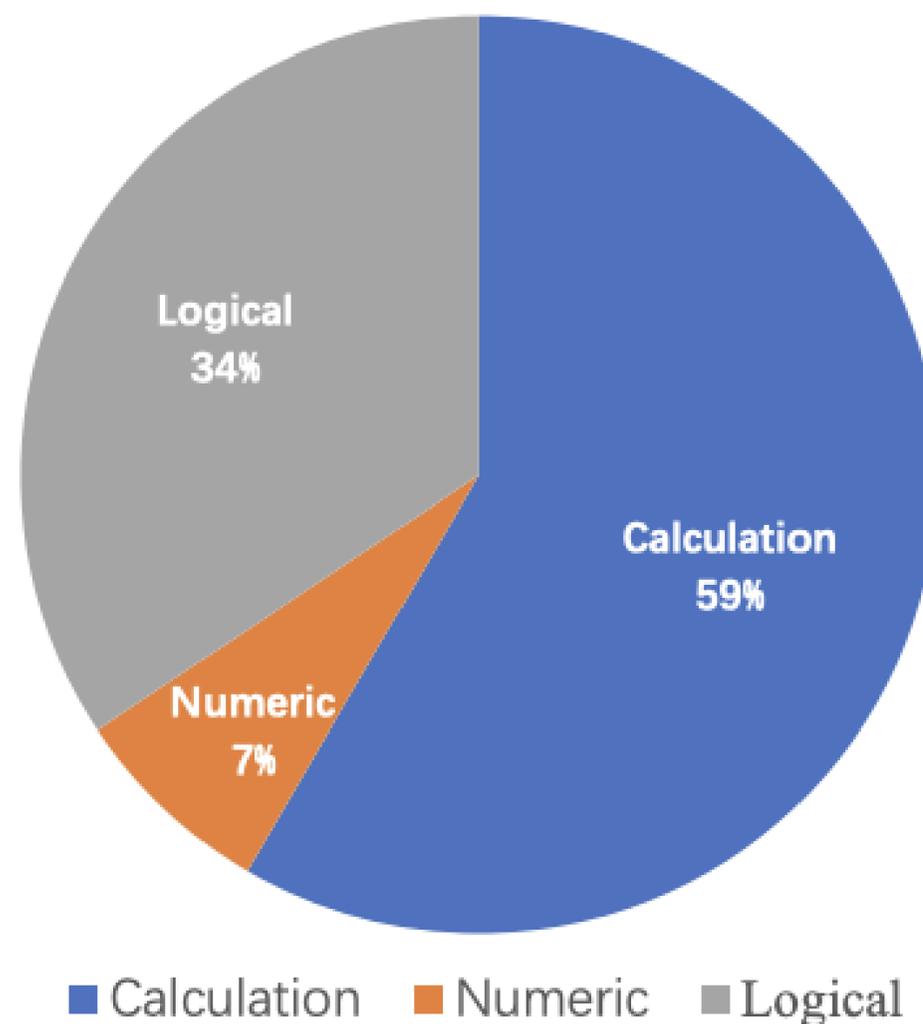
- PaLM2では与えられた文脈から正確な事実知識を想起する能力が不十分である
→RAG^[2]のような手法による改善が必要



考察②：LLMによる推論の誤りの種類

■ 算術推論タスク

- PaLM2では計算ミスや論理的エラーが多く生じている



まとめと今後の展望、疑問点

■ まとめ

- LLMが自身の誤りから推論を修正可能であることを示した
 - PromptingとTuningの両方の観点で推論能力が向上
- PaLM2において推論タスクでどのような誤りが多いかを示した
 - 常識推論では文脈把握の誤り、算術推論では計算ミスが多く生じることを示した

■ 疑問点

- 誤りの種類は言語モデルに依存する可能性がある
- 回答してから誤りを修正せずに、最初から推論の誤りを与えてもよさそう