大規模言語モデルに基づく 複数の外部ツールを利用した推論フレームワーク

稲葉達郎 清丸寛一 Fei Cheng 黒橋禎夫 京都大学



概要

- 大規模言語モデルによる推論で複数の外部ツールの 利用を可能にするフレームワークを提案
- 専門知識を要する数値推論タスク NumGLUE Task2 で世界最高性能を達成

研究背景:大規模言語モデルによる推論

解答に至る過程付きで解答を生成させる Chain-of-Thought (CoT) プロンプティングが有効 [Wei+, 2022]

課題:計算能力・知識量の限界

解けない問題例

Q:What is the amount of 11 moles of Glucosamine?

出力

知識の誤り:正しくは179.17

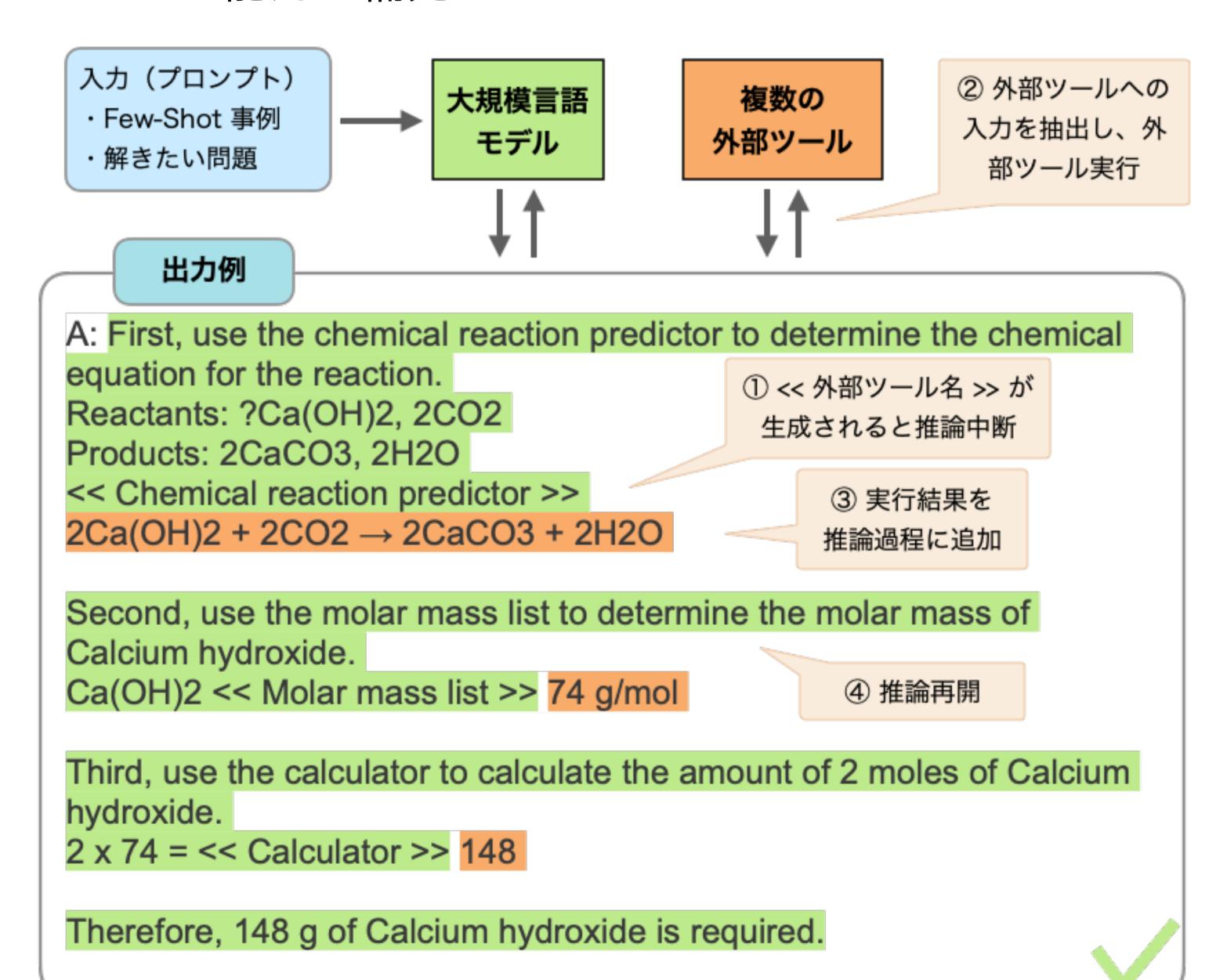
A: The molar mass of Glucosamine is 221.22 g/mol. Therefore, 11 moles of Glucosamine would be equal to 11 x 221.22 = 2434.42 g.

計算の誤り:正しくは2433.42

▶ 正答するにはこれらの問題を同時に対処する必要 単一の問題に着目し単一の外部ツールを利用し対処 する研究はあるが [Cobbe+, 2021]等 それでは不十分

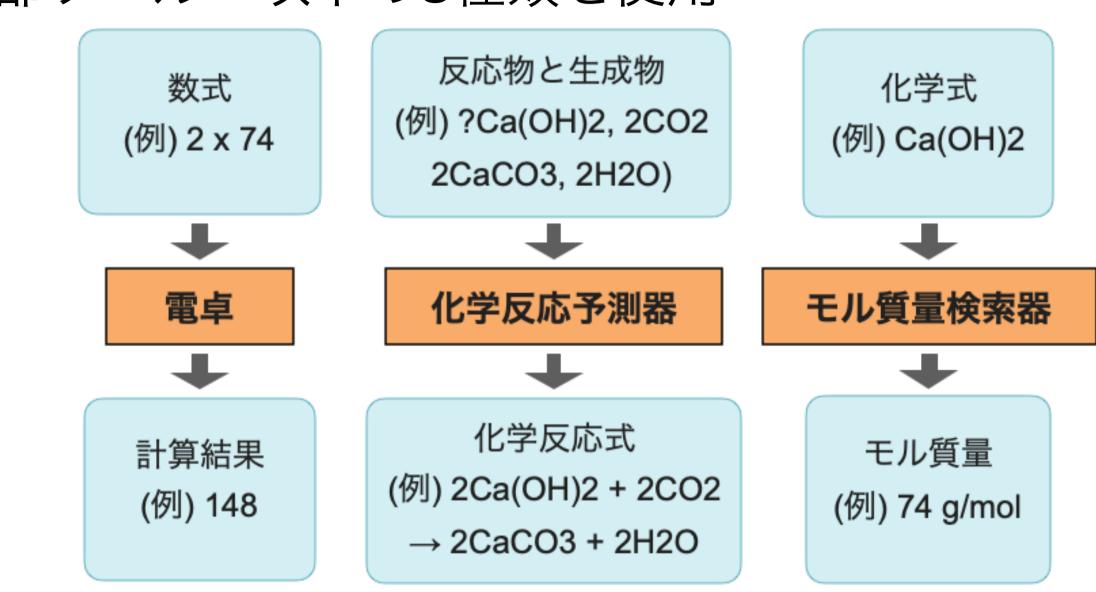
提案手法:複数の外部ツールの利用

複数の外部ツールを推論途中に呼び出し、大規模言語モデルの能力を補完

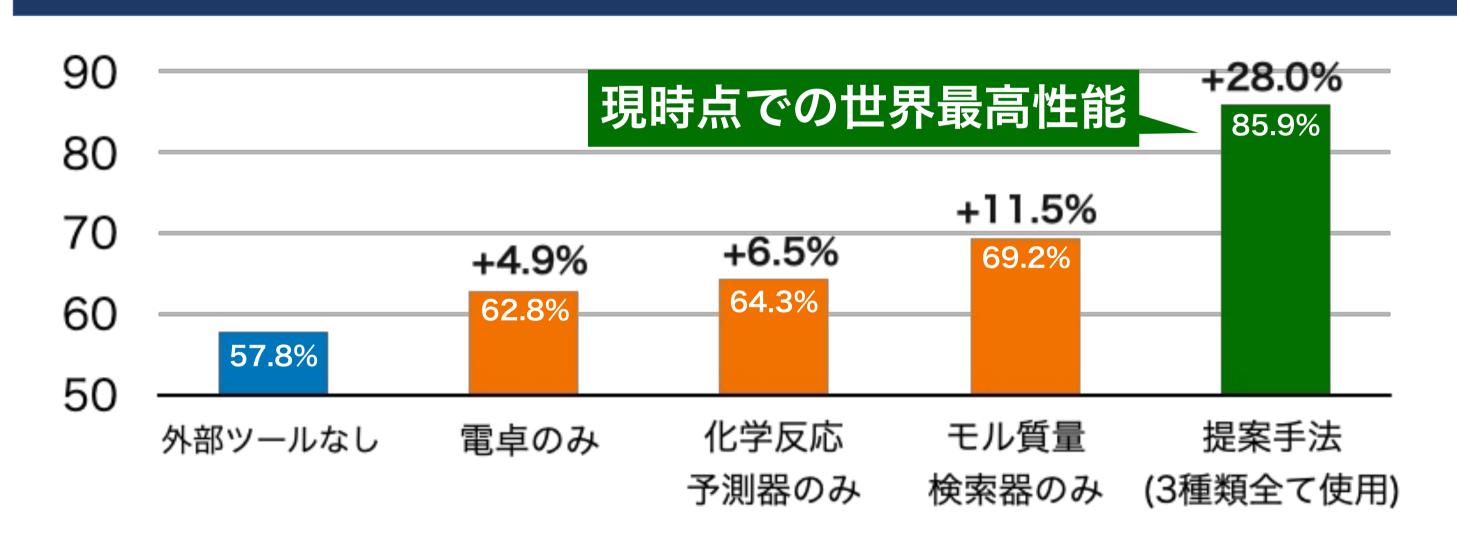


NumGLUE Task2 [Mishra+, 2022] での実験

- タスク:化学の専門知識を要する数値推論タスク 大規模言語モデル:GPT-3を使用
- ・ 外部ツール:以下の3種類を使用



結果



ケーススタディ

・提案手法による改善例

外部ツールなし

A: First, use the molar mass list to determine the molar mass of Aluminum carbonate. Al2(CO3)3 << Molar mass list >> 342 g/mol

Second, use the molar mass list to determine the molar mass of C. C << Molar mass list >> 12 g/mol

Third, use the calculator to calculate the mass percentage of C in Aluminum carbonate.

12 x 3 / 342 x 100 = << Calculator >> 8.76

Therefore, the mass percentage of C in Aluminum carbonate is 8.76%.

提案手法

First, use the molar mass list to determine the molar mass of Aluminum carbonate.
Al2(CO3)3 << Molar mass list >> 234 g/mol

Second, use the molar mass list to determine the molar mass of C. C << Molar mass list >> 12 g/mol

Third, use the calculator to calculate the mass percentage of C in Aluminum carbonate.

12 x 3 / 234 x 100 = << Calculator >> 15.38

Therefore, the mass percentage of C in Aluminum carbonate is 15.38%.

- ▶ 複数の外部ツールを利用することで正答
- 提案手法の誤り例

提案手法

First, use the molar mass list to determine the molar mass of Fe2O3. Fe2O3 << Molar mass list >> 160 g/mol

Second, use the molar mass list to determine the molar mass of O. O << Molar mass list >> 16 g/mol

Third, use the calculator to calculate the mass percentage of O in Fe2O3. $2 \times 16 / 160 \times 100 = << Calculator >> 20$

Therefore, the mass percentage of O in Fe2O3 is 20%.

立式の誤り:正しくは 3 x 16 / 160 x 100

▶ 外部ツールへの入力の誤りに起因

結論

- 大規模言語モデルによる推論で複数の外部ツールの 利用を可能にするフレームワークを提案
- NumGLUE Task 2 で世界最高性能を達成
- 提案手法は汎用的で様々なタスクに応用可能
- 今後の課題:異なる推論タスクでの有効性の検証