MultiTool-CoT: GPT-3 Can Use Multiple External Tools with Chain of Thought Prompting

Tatsuro Inaba¹, Hirokazu Kiyomaru¹, Fei Cheng¹, Sadao Kurohashi^{1, 2}

¹ Kyoto University, Japan ² National Institute of Informatics, Japan

Abstract

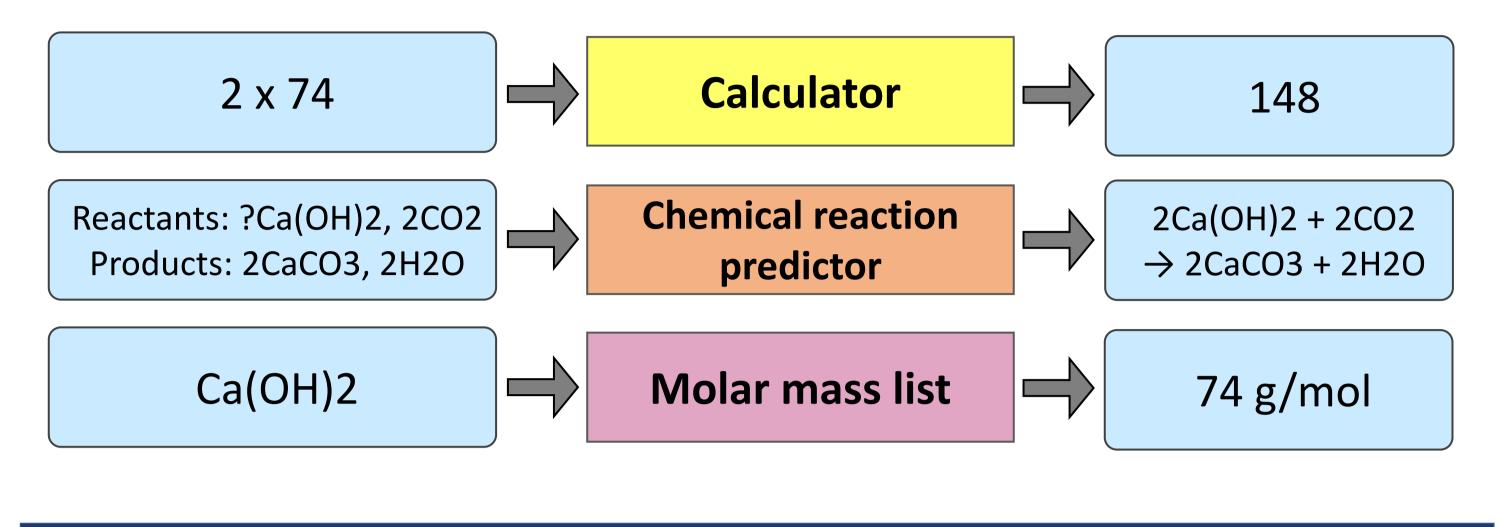
- We proposed **MultiTool-CoT**, a novel framework that allows LLMs to use multiple external tools during reasoning.
- **MultiTool-CoT** achieved SOTA on NumGLUE Task2 which requires numerical reasoning and domain-specific knowledge.

Background: Reasoning with LLMs

• In reasoning tasks, chain-of-thought (CoT) prompting is known to

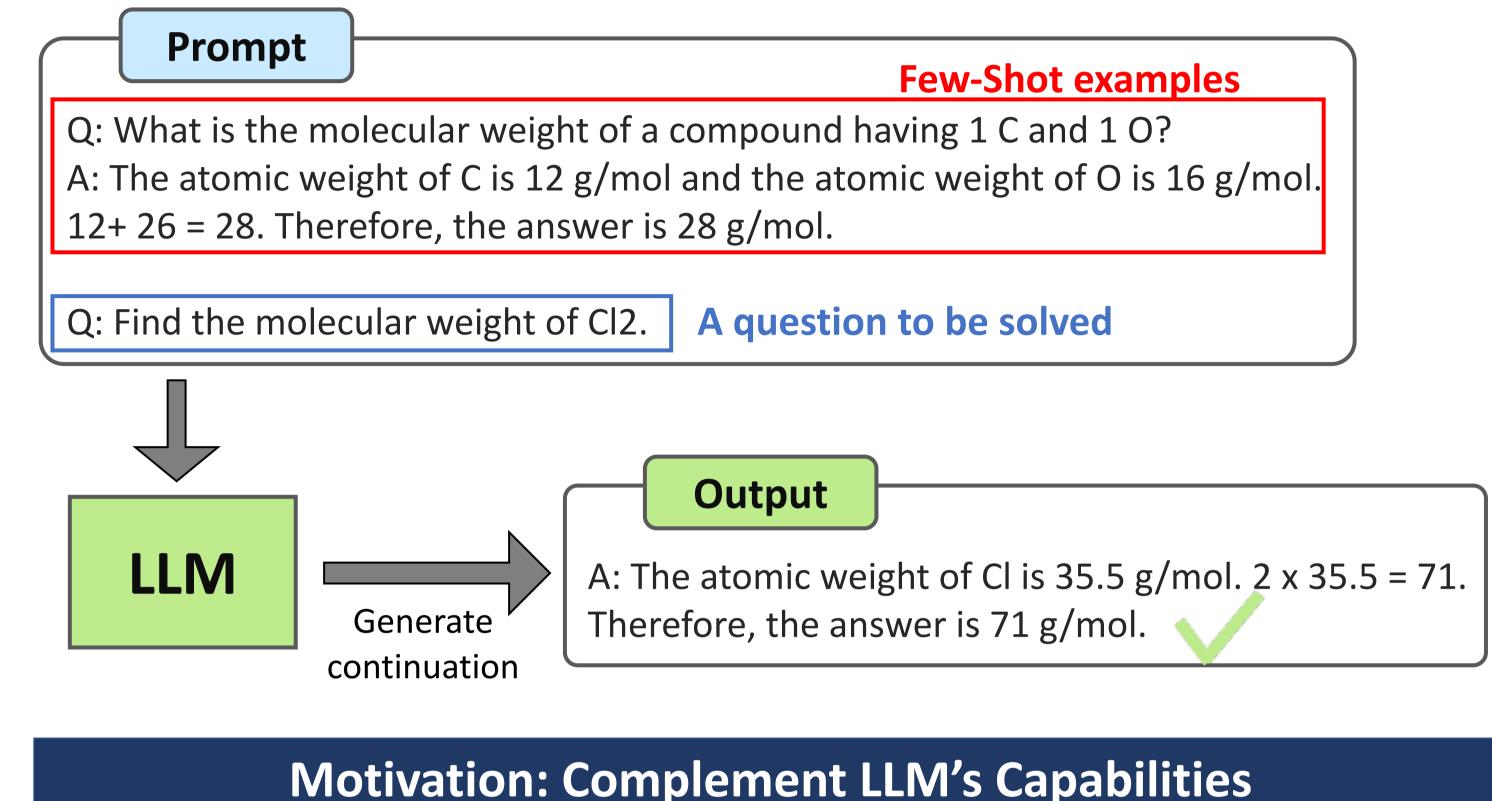
Experiment: NumGLUE Task2 [Mishra+, 2022]

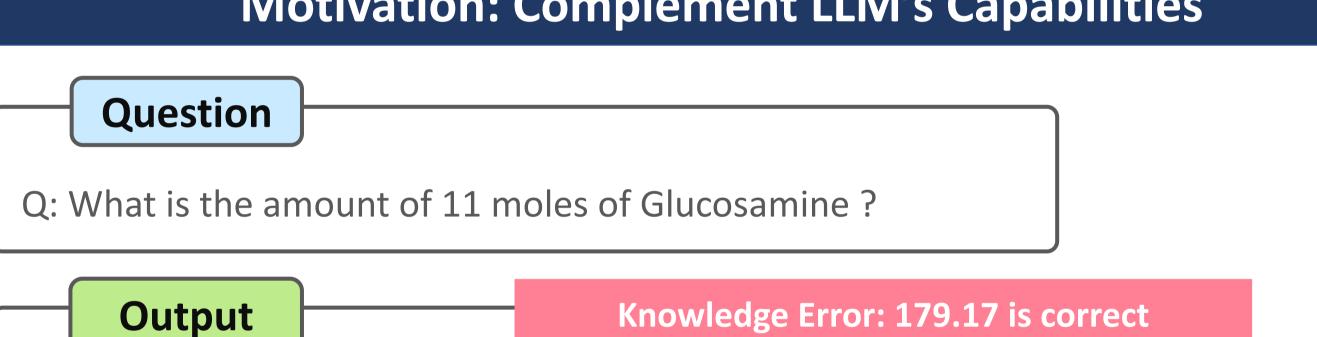
- Dataset: NumGLUE Task2, which requires numerical reasoning and domain-specific knowledge, predominantly in chemistry
- LLM: GPT-3 (text-davinci-003)
- External Tools: The following 3 tools



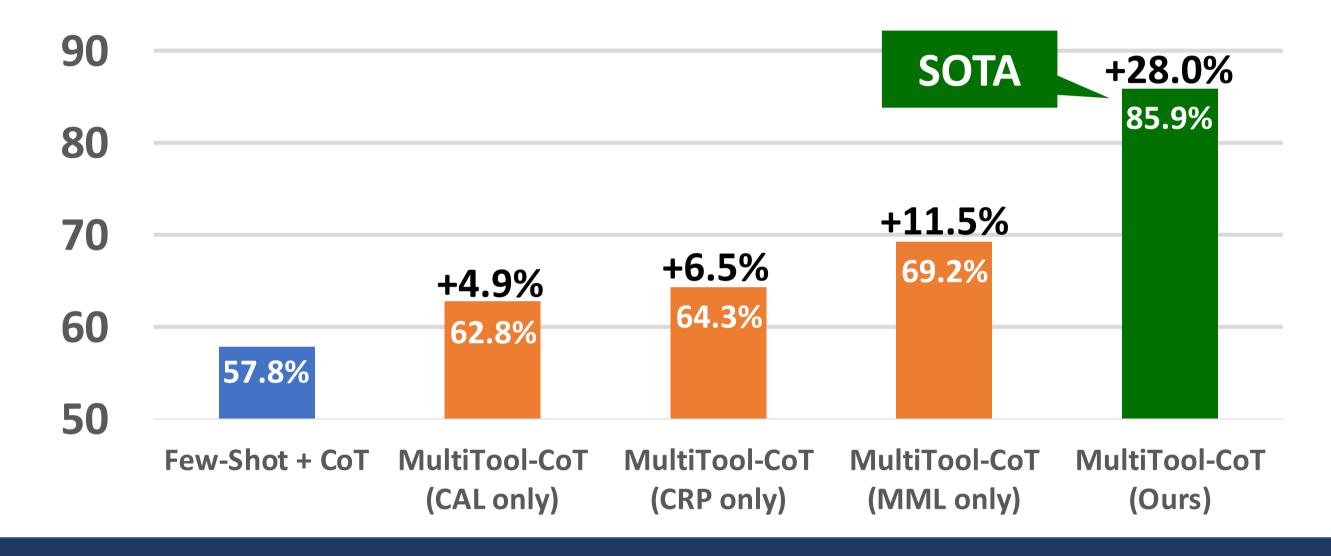


be effective. [Wei+, 2022]





Results



Case Study

• An improved example

MultiTool-CoT

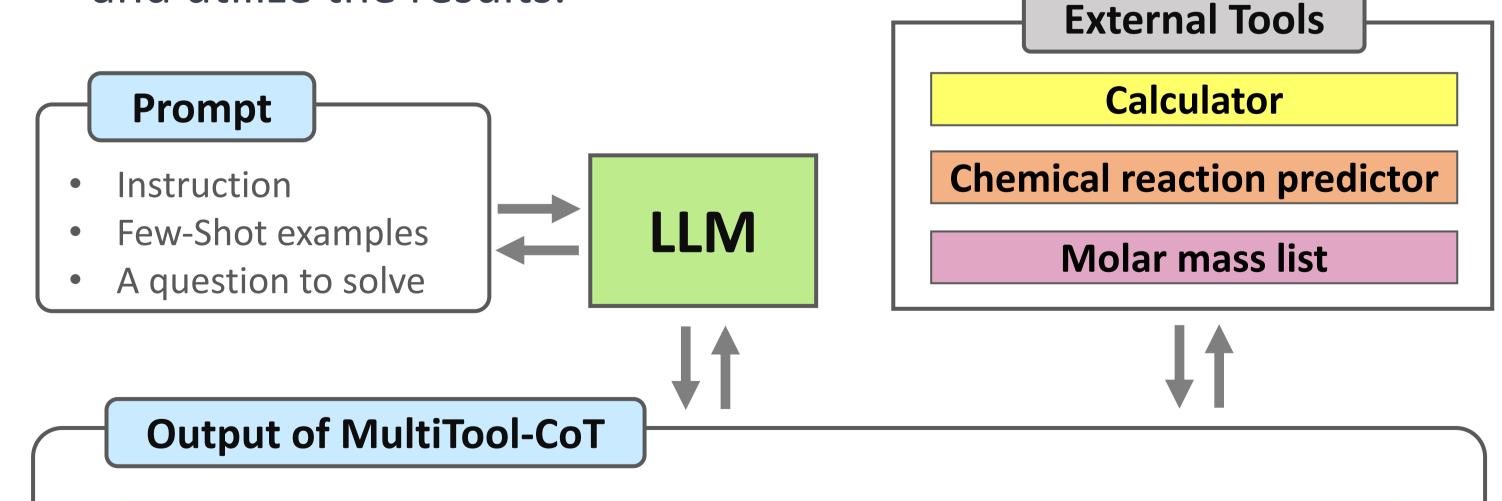
A: The molar mass of Glucosamine is 221.22 g/mol. Therefore, 11 moles of Glucosamine would be equal to 11 x 221.22 = 2434.42 g

Calculation Error: 2433.42 is correct

- To answer correctly, one needs to address these problems together.
- Previous studies are insufficient to solve such questions because they focused on complementing a single capability [Cobbe+, 2021].

Proposed Method: MultiTool-CoT

• LLMs invoke external tools at the appropriate reasoning steps and utilize the results.



Few-Shot + CoT

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%.

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%.

- MultiTool-CoT can provide correct answers by relying on external tools.
- An error example

MultiTool-CoT

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

A: First, use the chemical reaction predictor to determine the chemical equation for the reaction. Reactants: Ca(OH)2, 2CO2Products: 2CaCO3, 2H2O<< Chemical reaction predictor >> $2Ca(OH)2 + 2CO2 \rightarrow 2CaCO3 + 2H2O$

Second, use the molar mass list to determine the molar mass of Calcium hydroxide.

Ca(OH)2 << Molar mass list >> 74 g/mol

Third, use the calculator to calculate the amount of 2 moles of Calcium hydroxide.

2 x 74 = << Calculator >> 148

Therefore, 148 g of Calcium hydroxide is required.

Third, use the calculator to calculate the mass percentage of O in Fe2O3. $2 \times 16 / 160 \times 100 = <<$ Calculator >> 20Formulation Error:Therefore, the mass percentage of O in Fe2O3 is 20%.Source to calculate the mass percentage of O in Fe2O3 is 20%.

• It is caused by an invalid tool input.

Conclusion

- We proposed **MultiTool-CoT**, a novel framework that allows LLMs to use multiple external tools during reasoning.
- **MultiTool-CoT** achieved SOTA on NumGLUE Task2.
- The proposed framework is general and can be applied to various tasks.
- Our code is available at https://github.com/InabaTatsuro/MultiTool-CoT