

外部ツールの活用による 大規模言語モデルの拡張

2023年7月26日

京都大学 大学院情報学研究科 特定研究員

清丸寛一 (KIYOMARU Hirokazu)

ChatGPT

・大規模言語モデルに基づく対話システムで幅広い応用が可能

講演概要の作文

KI

「外部ツールの活用による大規模言語モデルの拡張」というタイトルで講演をします。大規模言語モデルによるテキスト生成に検索エンジンや電卓といった外部ツールを組み合わせるアプローチについて、現状と展望を述べます。ウェブページ掲載用の短い概要を考えてください。



タイトル: 「外部ツールの活用による大規模言語モデルの拡張」

概要: 本講演では、大規模言語モデルが持つポテンシャルを、さまざまな外部ツールとの融合により、どのように拡張していくことが可能かを解説します。現在のAI技術を最大限に活かすため、テキスト生成における検索エンジンや電卓といった外部ツールの適用方法とその効果について詳細に議論します。さらに、このアプローチによって生まれる新たな可能性や展望についても考察します。AIの進歩とともに成熟していく大規模言語モデルの最前線を、具体的な事例とともにお伝えします。AI技術者から一般ユーザーまで、多くの方にとって有益な内容となっております。ご期待ください。



京都の観光地の推薦

KI

京都のおすすめの観光地を3つ教えてください。



もちろんです、京都には素晴らしい観光地がたくさんありますので、選びました3つの観光地をご紹介します。

- 金閣寺（鹿苑寺）**：金閣寺はその美しい金箔張りの建物と庭園で知られています。金色に輝く建物は、四季折々の自然を反映し、美しい風景を作り出します。寺院とその庭園は1994年に世界遺産に登録されました。ここでは、日本の伝統的な庭園と建築を楽しむことができます。
- 清水寺**：清水寺は京都市東部の山間部に位置し、都心から素晴らしい眺めを楽しむことができます。特に有名なのは、その舞台となる木造の「清水の舞台」です。ここでは、周囲の桜や紅葉を一望することができます。
- 嵐山**：嵐山は自然豊かな観光地で、特に「嵐山竹林の道」が有名です。竹林の道はまさに別世界のようで、その静寂と美しさに圧倒されます。また、この地域には天龍寺や渡月橋など、見所が多数存在します。

これらの観光地は、どれも特色が異なり、京都の歴史や文化、自然の美しさを堪能するのに最適です。



複利計算

KI

10万円を元本として利回りが年率5%の投資商品に投資したときの10年後の利益を計算してください。



この問題は、複利計算を用いて解くことができます。年率5%の利率で10年間投資した場合、将来価値（FV）は以下のように計算されます。

$$FV = PV * (1 + r)^n$$

ここで、

$$PV \text{ (元本)} = 10 \text{万円}$$

$$r \text{ (年率)} = 5\% = 0.05$$

$$n \text{ (期間)} = 10 \text{年}$$

を代入して計算すると、

$$FV = 100,000 * (1 + 0.05)^{10} = 162,889.46 \text{円}$$

となります。

つまり、元本10万円を年率5%の利回りで10年間投資すると、約162,889円になります。

したがって、10年後の利益は、

$$\text{利益} = FV - PV = 162,889.46 \text{円} - 100,000 \text{円} = 62,889.46 \text{円}$$

となります。これが10年間の投資による利益となります。



ChatGPT の欠点

- 世界知識や数値理解が不完全でしばしば不正確なテキストを生成

京都駅から京大への道案内

KI 京都駅から京都大学に行くにはどうしたら良いですか？

京都駅から京都大学に行くには、公共交通機関を利用することが一番便利です。以下に詳細な手順を示します。

1. 京都駅から地下鉄を利用:
京都市営地下鉄烏丸線に乗り、「今出川駅」で下車します。ここから徒歩で京都大学へ向かうことが可能です。通常、今出川駅から京都大学までの所要時間は約10分程度です。
1. バスを利用する場合: **✗ 30分程度かかる**
京都駅からは、京都市バスも利用可能です。京都駅前のバス停から5系統のバスに乗り、「百万遍」で下車します。ここから徒歩で京都大学へ向かいます。**✗**このルートは交通状況によりますが、所要時間は約30分程度となるでしょう。
17系統 or 206系統

注意点としては、京都は観光地でもありますので、特に観光シーズンは道路が混雑することがあります。時間に余裕をもって移動することをお勧めします。

4桁 x 4桁の掛け算

KI 2023 × 2023 はいくつですか？
(正解は 4,092,529)

2023 × 2023 は 4096529 です。
✗ 不正解

KI 本当ですか？

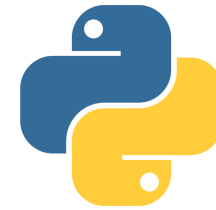
申し訳ありません、計算間違いがありました。2023 × 2023 の正しい結果は 4096529ではなく、4096529です。これが正しい結果です。再度、ご確認をお願いします。
✗ やはり不正解

外部ツールによる言語モデルの拡張

- ChatGPTをはじめとする大規模言語モデルによるテキスト生成時に**外部ツールを活用**することで知識やその他の能力を補う



検索



コード実行



電卓



MLモデル

etc.

言語モデル

言語モデル

- 文や表現の出現確率を計算するモデル

私はりんごを

大規模コーパス

...お店で**私はりんご**を食べた...

...彼と**私はりんご**を食べた...

...昨日**私はりんご**をかじった...

...**私はりんご**を食べた後...

...**私はりんご**を殴った夢を...

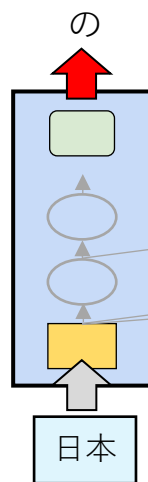
$$P(\text{食べた} | \text{私はりんごを}) = \frac{\text{コーパス中の頻度 (私はりんごを食べた)}}{\text{コーパス中の頻度 (私はりんごを)}}$$

$$P(\text{殴った} | \text{私はりんごを}) = \frac{\text{コーパス中の頻度 (私はりんごを殴った)}}{\text{コーパス中の頻度 (私はりんごを)}}$$

GPT (Generative Pre-trained Transformer) [Radford et al., 2018]

- **Transformer** という構造のニューラルネットワークで構築した言語モデル
 - 学習時：文脈から次の単語（トークン）を予測することを学習

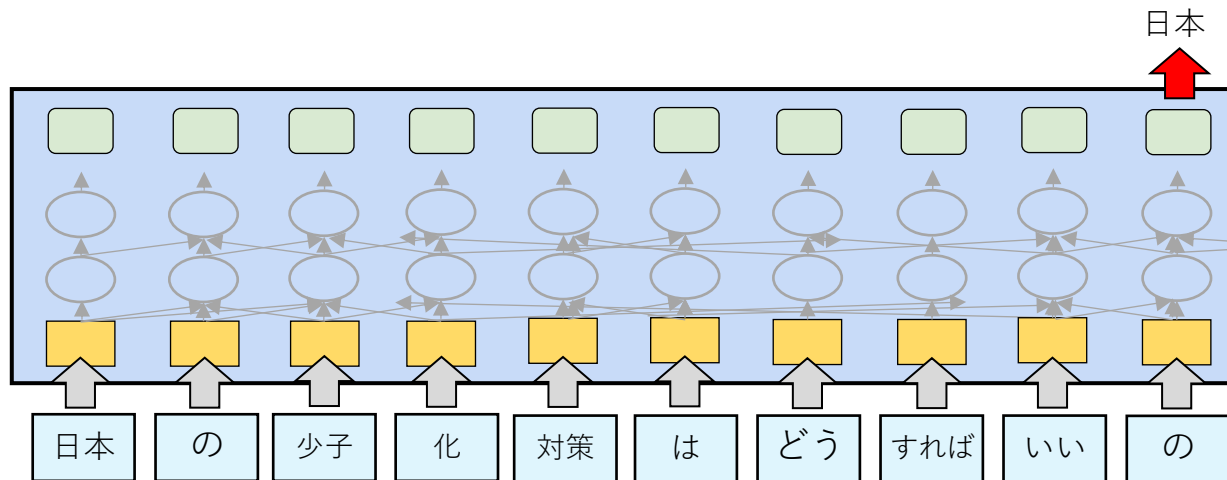
学習データ：日本の少子化対策には、次のようなアプローチを...



GPT (Generative Pre-trained Transformer) [Radford et al., 2018]

- **Transformer** という構造のニューラルネットワークで構築した言語モデル
 - 推論時：次の単語（トークン）の予測を繰り返してテキストを生成

入力テキスト（プロンプト）：日本の少子化対策はどうすればいいの



知識を参照したり厳密に計算を解いたりしながら予測しているわけではない

外部ツールによる言語モデルの拡張

1. 検索エンジンによる拡張
2. コードインタプリタによる拡張
3. ユーザ定義ツールによる拡張

1. 検索エンジンによる拡張 Bing Chat

- 大規模言語モデル GPT-4 + Bing 検索

京都駅から京大への道案内

京都駅から京都大学に行くにはどうしたら良いですか？

✓ Searching for: 京都駅から京都大学に行く方法 **検索を実行**

✓ Generating answers for you...

**検索結果を参照しつつ
GPT-4 が応答を生成**

こんにちは。京都駅から京都大学へのアクセス方法は、京都市バスの利用がおすすめです。17・206系統であれば乗り換えなしで行けるので便利です。所要時間は約30分程度です¹。また、吉田キャンパス、宇治キャンパス、桂キャンパスの三つのキャンパスがあります²。ご参考になれば幸いです。

Learn more:

1 of 30 ●

1. curlpingnosia.wase.com

2. kyoto-u.ac.jp

3. kyoto-note.com



大谷翔平選手の今年の成績

大谷翔平選手の2023年の成績を教えてください。

✓ Searching for: 大谷翔平選手 2023年 成績 **検索を実行**

✓ Generating answers for you...

**最新情報に基づき
GPT-4 が応答を生成**

大谷翔平選手の2023年の成績は、打者成績で打率.302、本塁打36本、打点77、盗塁11、OPS1.072で、投手成績では防御率3.71、勝利8、敗戦5、奪三振148です¹。なお、これらの成績は2023年7月24日時点のものです¹。

Learn more:

1 of 30 ●

1. sportingnews.com

2. rakuchin-tanoc.com

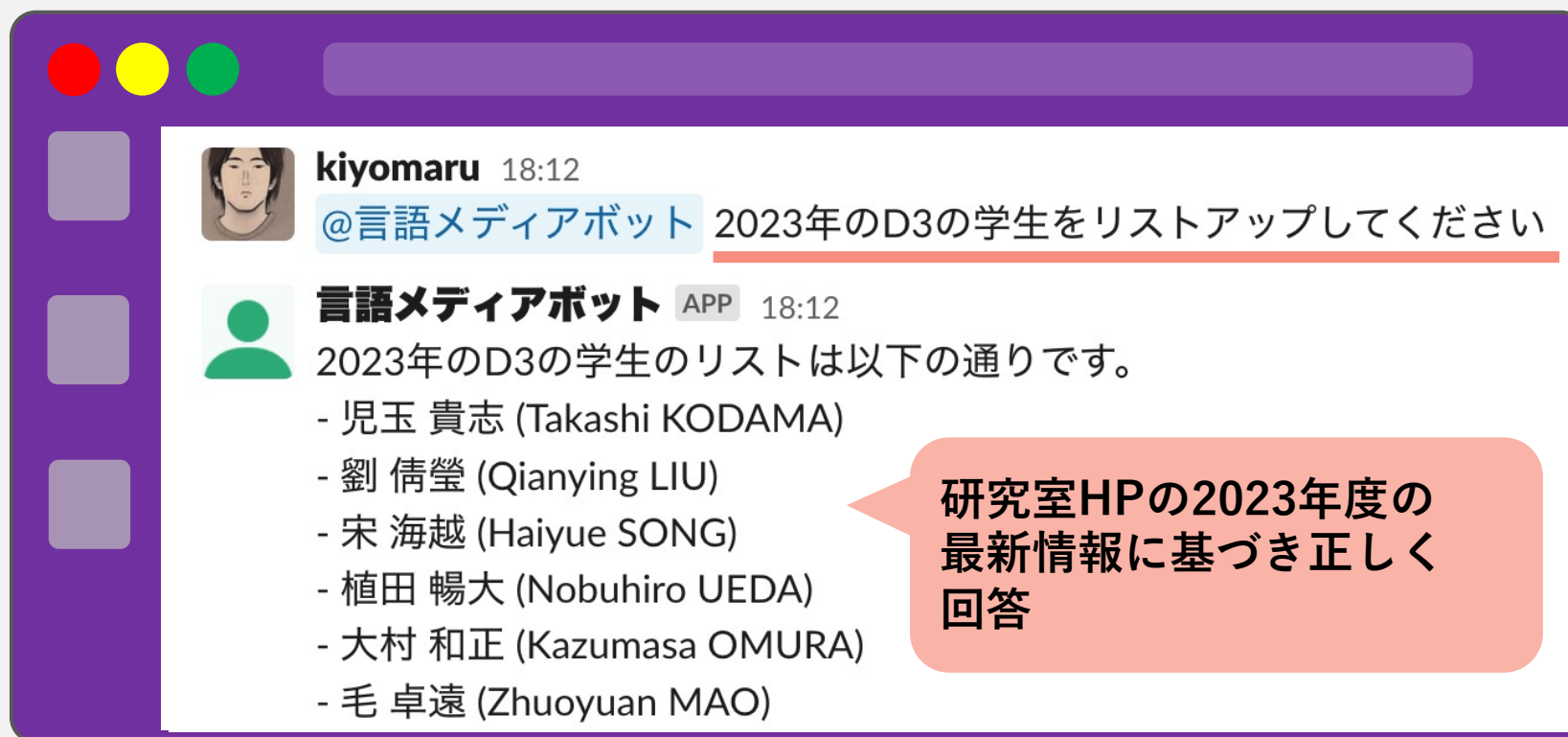
3. npb.jp



1. 検索エンジンによる拡張 カスタム質問応答システム

- 大規模言語モデル + ローカル文書検索
 - [OpenAI API](#) と [LlamaIndex](#) を使うと100行弱の Python コードで実装可能

言語メディアボット：研究室HPの最新情報に基づく Slack上のQAシステム



The screenshot shows a Slack chat interface with a purple header. A user named 'kiyomaru' asks for a list of 2023 D3 students. The '言語メディアボット' (Language Media Bot) responds with a list of names and their English names. A red callout box highlights the bot's response, stating it is based on the latest information from the lab's HP.

kiyomaru 18:12
@言語メディアボット 2023年のD3の学生をリストアップしてください

言語メディアボット APP 18:12
2023年のD3の学生のリストは以下の通りです。

- 児玉 貴志 (Takashi KODAMA)
- 劉 倩瑩 (Qianying LIU)
- 宋 海越 (Haiyue SONG)
- 植田 暢大 (Nobuhiro UEDA)
- 大村 和正 (Kazumasa OMURA)
- 毛 卓遠 (Zhuoyuan MAO)

研究室HPの2023年度の最新情報に基づき正しく回答

2. コードインタプリタによる拡張

- ChatGPT を含む昨今の大規模言語モデルはコードを生成できる ▶
- Q. モデルが生成したコードを "内製" ツールとして活用できないか？
- A. 外部ツールとして**コードインタプリタ**（コードを解釈・実行するツール）を組み込むことで可能
 - PAL [Gao et al., 2022]
 - PoT [Chen et al., 2023]
 - Chameleon [Liu et al., 2023]
 - LATM [Cai et al., 2023]
 - CERATOR [Qian et al., 2023]

作った
コードを
保存して
再利用

4桁 x 4桁の掛け算を解くプログラム生成

KI

2023 x 2023 を解くプログラムを書いてください。



Pythonの短いプログラムを使って、2023 x 2023の計算をすることができます：

python

Copy code

```
def multiply_numbers():  
    num1 = 2023  
    num2 = 2023  
    result = num1 * num2  
    return result  
  
print(multiply_numbers())
```

このプログラムは、2つの数値2023と2023を掛け算し、結果を表示します。



2. コードインタプリタによる拡張 アプリケーション：ChatGPT Code Interpreter

- 大規模言語モデル ChatGPT + Python インタプリタ
4桁 x 4桁の掛け算

KI 2023 × 2023 はいくつですか？

Finished working Hide work ^

```
python Copy code  
# 数値の乗算を行います  
result = 2023 * 2023  
result
```

RESULT
4092529

2023 × 2023 の結果は 4092529 です。 ✓

掛け算を解くプログラムを ChatGPT が生成

プログラムの実行結果をインタプリタが生成

3. ユーザ定義ツールによる拡張

- 特定の問題に対処するために**ユーザ定義ツール**を活用する研究も盛ん
 - ReAct [Yao et al., 2022]: Wikipedia 検索
 - Toolformer [Schick et al., 2023]: QAシステム、電卓、翻訳システム、Wikipedia 検索
 - Gorilla [Patil et al., 2023]: HuggingFace、TorchHub、TensorHubで提供のAPI
 - MultiTool CoT [Inaba et al., 2023]: 電卓、化学反応式立式器、モル質量検索

3. ユーザ定義ツールによる拡張

MultiTool-CoT [Inaba et al., 2023]

問題 (NumGLUE)

Q: Find the amount of Calcium hydroxide that is required to react with 2 moles of Carbon dioxide to form 2 moles of Calcium carbonate along with 2 moles of Water.

(2モルのCO₂と反応して2モルのCaCO₃と2モルのH₂Oを生成するために必要なCa(OH)₂の量を求めよ。)

- 正答するにはモル質量の知識、化学反応式の立式、数値計算が必要
 - いずれも大規模言語モデルはよく間違う

3. ユーザ定義ツールによる拡張

MultiTool-CoT [Inaba et al., 2023]

推論中に使用可能な外部ツール

説明文

Use the following tools at your disposal and answer some questions:

- << Calculator >>
- << Chemical reaction predictor >>
- << Molar mass list >>

外部ツールを使いながら問題を解く具体例

Few-Shot 事例

Q: Find the amount of H₂CO₃ formed on combining 2 moles of CO₂ and 2 moles of H₂O.

A: First, use the chemical reaction predictor to determine the chemical equation for the reaction.

Reactants: 2CO₂, 2H₂O

Products: ?H₂CO₃

<< Chemical reaction predictor >>

2CO₂ + 2H₂O → 2H₂CO₃

Second, use the molar mass list to determine the molar mass of H₂CO₃.

H₂CO₃ << Molar mass list >> 62 g/mol

Third, use the calculator to calculate the amount of 2 moles of H₂CO₃.

2 x 62 = << Calculator >> 124

Therefore, 124 g of H₂CO₃ is formed.

Q: ... } 他 Few-Shot 事例
A: ... }

外部ツールへの入力

外部ツールの実行トリガー

外部ツールの想定出力

解きたい問題

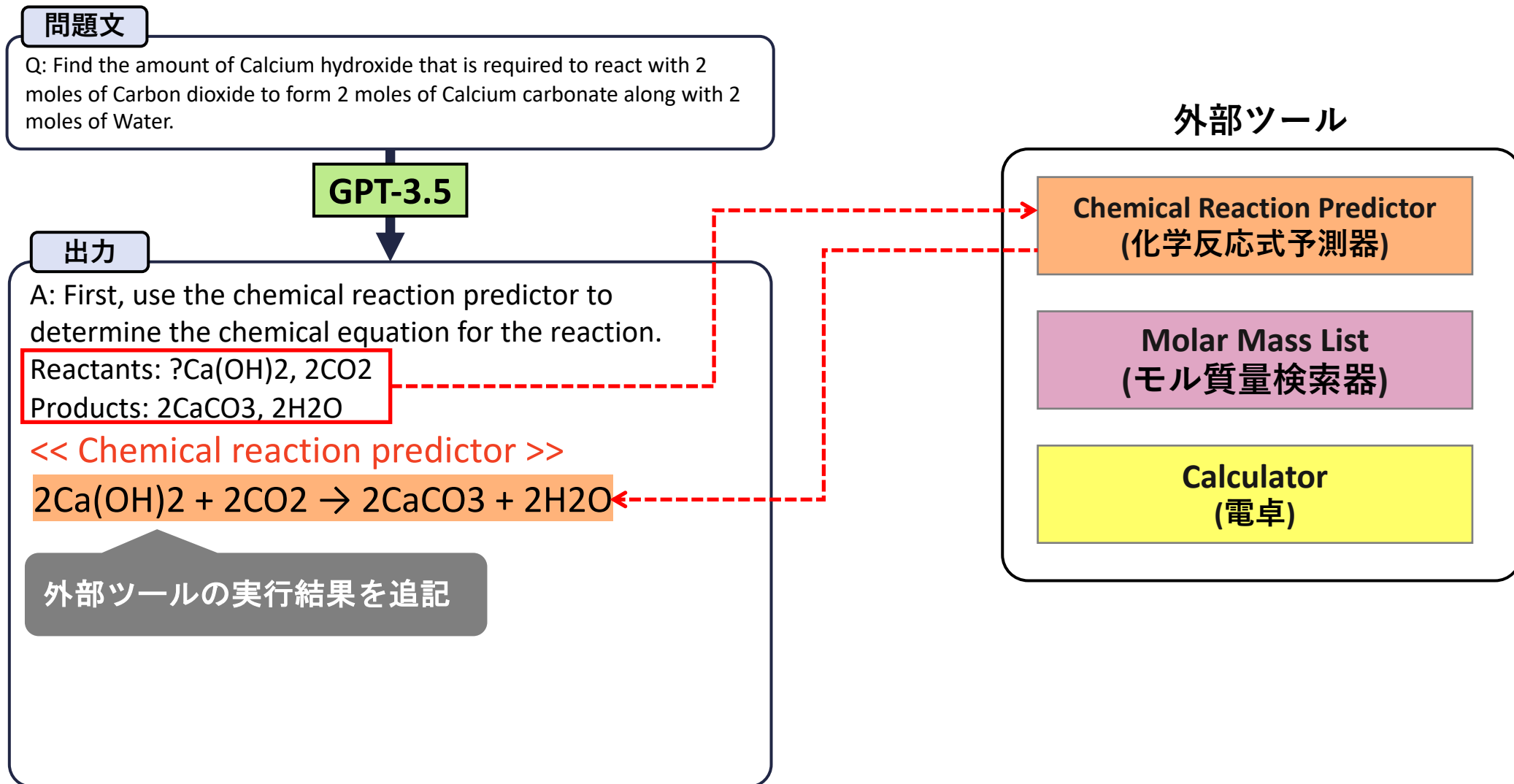
問題文

Q: Find the amount of Calcium hydroxide that is required to react with 2 moles of Carbon dioxide to form 2 moles of Calcium carbonate along with 2 moles of Water.

大規模
言語モデル
GPT-3.5

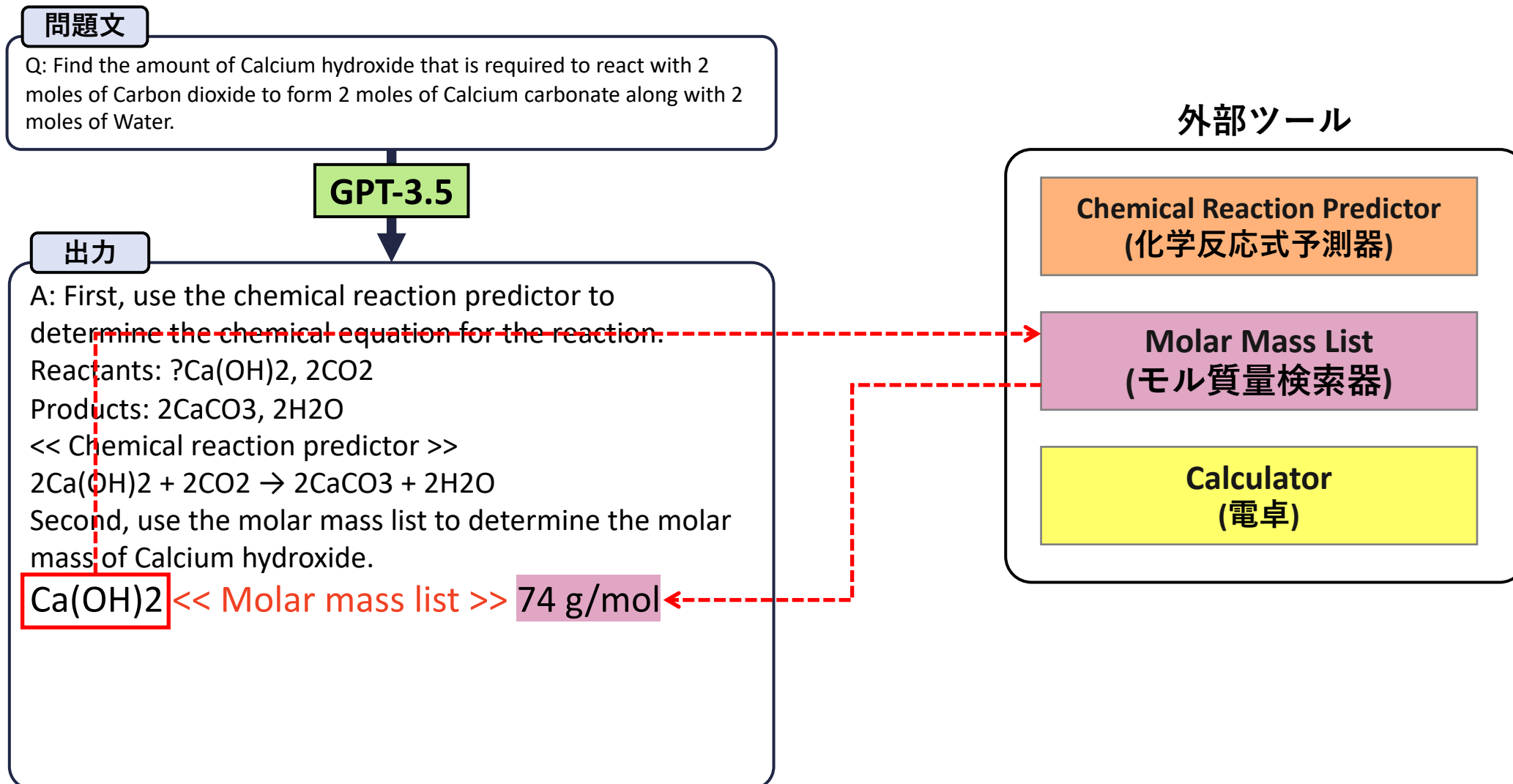
3. ユーザ定義ツールによる拡張

MultiTool-CoT [Inaba et al., 2023]



3. ユーザ定義ツールによる拡張

MultiTool-CoT [Inaba et al., 2023]



3. ユーザ定義ツールによる拡張

MultiTool-CoT [Inaba et al., 2023]

問題文

Q: Find the amount of Calcium hydroxide that is required to react with 2 moles of Carbon dioxide to form 2 moles of Calcium carbonate along with 2 moles of Water.

GPT-3.5

出力

A: First, use the chemical reaction predictor to determine the chemical equation for the reaction.
Reactants: ?Ca(OH)₂, 2CO₂
Products: 2CaCO₃, 2H₂O
<< Chemical reaction predictor >>
~~2Ca(OH)₂ + 2CO₂ → 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

外部ツール

Chemical Reaction Predictor
(化学反応式予測器)

Molar Mass List
(モル質量検索器)

Calculator
(電卓)

3. ユーザ定義ツールによる拡張

MultiTool-CoT [Inaba et al., 2023]

問題文

Q: Find the amount of Calcium hydroxide that is required to react with 2 moles of Carbon dioxide to form 2 moles of Calcium carbonate along with 2 moles of Water.

GPT-3.5

出力

A: First, use the chemical reaction predictor to determine the chemical equation for the reaction.
Reactants: $?Ca(OH)_2, 2CO_2$
Products: $2CaCO_3, 2H_2O$
<< Chemical reaction predictor >>
 $2Ca(OH)_2 + 2CO_2 \rightarrow 2CaCO_3 + 2H_2O$
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.
 $2 \times 74 =$ << Calculator >> 148
Therefore, 148 g of Calcium hydroxide is required. ✓

外部ツール


Chemical Reaction Predictor
(化学反応式予測器)

Molar Mass List
(モル質量検索器)

Calculator
(電卓)

3. ユーザ定義ツールによる拡張 アプリケーション：ChatGPT Plugins

- 大規模言語モデル（ChatGPT） + ユーザ定義ツール（7月24日現在737種）

レストラン探し with 食べログ 

KI 7月26日の19:00に4名で入れる京都市内の焼肉屋さんを教えてください。

外部ツール実行



Used Tabelog

2023年7月26日の19:00に4名で予約可能な京都市内の焼肉店をご紹介します。

1. りきちゃん 祇園店

- 総合評価：3.78
- エリア：祇園四条
- ジャンル：焼肉
- 予約可能時間：18:45, 19:00, 19:15



リクエストの結果に基づき Chat-GPT がレストラン情報を生成

REQUEST TO TABELOG

```
{
  "free_keyword": "京都 焼肉",
  "reservation_datetime": "2023-07-26T19:00:00.000+09:00",
  "reservation_member": "4"
}
```

食べログへのリクエスト
(ChatGPT が生成)

RESPONSE FROM TABELOG

```
{
  "restaurant_list": [
    {
      "reservation_url": "https://tabelog.com/kyoto/A2601/A260301/26036373/?cid=gpt_plugin&svd=20230726&svps=4&svt=1900",
      "name": "りきちゃん 祇園店",
      "total_score": "3.78",
      "area_name": "祇園四条",
      "genre_name": "焼肉",
      "img_url": "https://tblg.kimg.com/restaurant/images/Rvw/167743/320x320_square_b137226444cb5b4a884f200e8c7b7d36.jpg",
      "reservation_datetimes": [
        "2023-07-26T18:45:00.000+09:00",
        "2023-07-26T19:00:00.000+09:00",
        "2023-07-26T19:15:00.000+09:00"
      ]
    }
  ],
}
```

リクエストの結果

まとめ

- 大規模言語モデルは極めて流暢なテキストを生成するが、世界知識や数値理解の不足により、しばしば不正確なテキストを生成してしまう
- 大規模言語モデルによるテキスト生成時に外部ツールを活用することでこの問題を解決・緩和する取り組みが盛ん
 - 検索エンジンによる拡張（Bing Chat）
 - コードインタプリタによる拡張（ChatGPT Code Interpreter）
 - ユーザ定義ツールによる拡張（ChatGPT Plugins）

議論・展望

- **外部ツールの活用によるモデルパラメータの削減や効果的な利用**
 - 推論時に検索ができるならあらゆる事実を暗記している必要はない
 - 暗記のためのパラメータを削減してモデルを小型化したり、より深い言語理解のための追加のキャパシティとして利用したりできる可能性
- **効率的な学習方法の模索：外部ツールの使い方をいつ教えるべきか？**
 - 現在の主流は Fine-tuning 時もしくは推論時 (in-context 学習)
 - 事前学習時から外部ツールの利用を前提とした訓練をすることで、外部ツールをより効率的に使いこなせる可能性；一方である程度は外部ツールに頼らない訓練をしないと賢くならない直感もある