

解法プロセスおよび計算プログラムの 生成によるLLMの数学的推論能力の向上

UCLab : 寺島 青 高木 峻 田村 直樹 庄子 和之

名古屋大学 工学研究科 河口研究室



LLMを用いて数学の問題を解くためには以下の要素が必要

- 問題文から解法を導く力
- 解法を基に立式しその解を求める力
- (指定した形式で回答を出力できる力)

プロンプト xxxの解を求めなさい

LLM

問題の解法を推論

解法を基に立式

回答を生成

応答

[問題の解法]

この問題には, xxxの公式が必要です.

[解] y

チューニング前の課題

指示:

異なる2つの整数を1から9までの整数の中から選びます。それらが両方とも奇数である確率を求め、解のみを答えなさい。

応答:

異なる2つの整数を1から9までの整数の中から選ぶ場合、それらが両方とも奇数である確率を求めるには、まず可能な組み合わせをすべて列挙し、その中で両方が奇数になる組み合わせの割合を計算します。

1から9までの奇数は3, 5, 7です。これらの中から2つを選ぶ方法の総数は、組み合わせの公式を使って計算できます。具体的には、次のようになります。

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

ここで、 $\binom{n}{k}$ は全体の項目数、 $\binom{k}{k}$ は選ぶ項目数です。今回の場合、 $\binom{3}{2}$ (奇数) と $\binom{2}{2}$ です。したがって、

$$\binom{3}{2} = \frac{3!}{2!(3-2)!} = \frac{3 \times 2 \times 1}{(2 \times 1)(1)} = 3$$

つまり、3つのペアがあります。

次に、これらのペアを並べる順序が重要なので、2通りの並べ方があります。したがって、全体の組み合わせの数は、

$$3 \times 2 = 6$$

となります。

以上より、1から9までの奇数から選んだ2つの整数が、両方とも奇数である確率は、

$$P = \frac{\text{有利な結果の数}}{\text{全結果の数}} = \frac{6}{36} = \frac{1}{6}$$

となります。

よって、答えは $\boxed{\frac{1}{6}}$ です。

← 解のみを答えるように指示

解を求める力が不足しており誤答

← 回答の形式が指示に沿っていない

- step1 : 指定した形式で回答させる
- step2 : 推論時に計算過程を統一させる
- step3 : 計算力が必要な問題に対応する

データセット

Training : MATHベンチマークトレーニングデータセットを含む数学分野のデータセット^[1]

- 日本語の問題文, 計算過程, 回答の組

Validation : 中間評価で使用されたデータセット

使用したモデル

- llm-jp-3-13b-3
- LoRA(Low Rank Adaptation)

index	dataset	question	generated_solution	question_ja	generated_solution_ja
int64 0 1.83M	string · classes 2 values	string · lengths 16 4.31k	string · lengths 18 3.44k	string · lengths 1 4.15k	string · lengths 3 2.93k
0	gsm8k	Martha has 18 crayons. She lost half of them, so she bought a new set of 20.	Let's solve this problem using Python code. <llm-code>..	マーサは18本のクレヨンを持っている。その半分をなくしてしまったので、新しいクレヨンを20本買いました。..	Pythonコードを使用してこの問題を解決しましょう。<llm-code>..
1	gsm8k	Noa scored 30 points to win a bowl, and Phillip scores twice that number. What is..	Let's solve this problem using Python code. <llm-code>..	ボウルに勝つためにノアは30点、フィリップはその2倍を得点した。ボウルに勝つためにトムとフィリップが得点..	Pythonコードを使用してこの問題を解決しましょう。<llm-code>..
2	math	In a rectangular coordinate system, what is the number of units in the distance from..	To find the distance between the origin \$(0, 0)\$ and the point..	直交座標系において、原点から点 \$(-15, 0)\$ までの距離は何単位か。	Origin \$(0, 0)\$ との間の距離を見つけるには、長方形の座標系でポイント\$ (-15, ..

[1].<https://huggingface.co/datasets/kunishou/OpenMathInstruct-1-1.8m-ja>

Step1 : 指定した形式での回答

以下の形式のデータセットを作成

input : 日本語の問題文

output : 計算過程[回答]x

training data : 150,000

validation data : 10,000

← 回答欄を作成し[回答]以降のトークンを結果として抽出する

```
{'input': '直交座標系において、原点から点  $(-15, 8)$  までの距離は何単位か。',  
  'output': 'Origin  $(0, 0)$  の間の距離を見つけるには、長方形の座標系でポイント  $(-15, 8)$  をを見つけるには、距離式を使用できます。
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$
 ここで、 $(x_1, y_1)$  は最初のポイントの座標であり、 $(x_2, y_2)$  は2番目のポイントの座標です。  
原点と指定されたポイントの値を置き換えると、次のようになります。  
<llm-code>  
from sympy import sqrt  
x1, y1, x2, y2 = 0, 0, -15, 8  
distance = sqrt((x2 - x1)**2 + (y2 - y1)**2)  
distance  
</llm-code><llm-code-output>  
17  
</llm-code-output>  
したがって、原点とポイント  $(-15, 8)$  の間の距離は  $\boxed{17}$  単位です。[回答]17'}
```

Step1 : 指定した形式での回答

【入力】

$y = f(x)$ のグラフ上に点 $(2, 9)$ があるとき、 $y = f(-x)$ のグラフ上に必ずある点が存在します。その点の x 座標と y 座標の和を求めなさい。

【出力】

以下はタスクを説明する指示です。要求を適切に満たす応答を書きなさい。

指示:

以下は数学の問題です。必要な公式を用い答えを求めてください。 $y = f(x)$ のグラフ上に点 $(2, 9)$ があるとき、 $y = f(-x)$ のグラフ上に必ずある点が存在します。その点の x 座標と y 座標の和を求めなさい。

応答:

この問題を解決するために、式 $y = -f(x)$ が $y = f(x)$ のグラフ上に正確に1つしかないというアイデアを使用できます。

この場合、ポイント $(2, 9)$ で方程式 $y = f(x)$ のグラフ上にあります。

y と $f(x) = 9$ の関係の観点から f を表現することができます。

両側に $1/x$ を掛けると、 $f(x) = 9/x$ を取得します。

したがって、 $y = -f(x)$ in $y = -f(x)$ の x 座標は -2 です。

これで、 y と $-f(x)$ を式 $y = f(x)$ の方程式に置き換えて、 $y = -f(-x)$ を取得できます。

したがって、 $y = -f(-x)$ のグラフ上のポイントは $(2, 9)$ です。

したがって、 x および y 座標の合計は $\boxed{11}$ です。 [回答] 11

【回答】トークン以降を解として抽出

【生成された応答（抽出後）】

11

【正解】

7

Accuracy = 0.07

課題

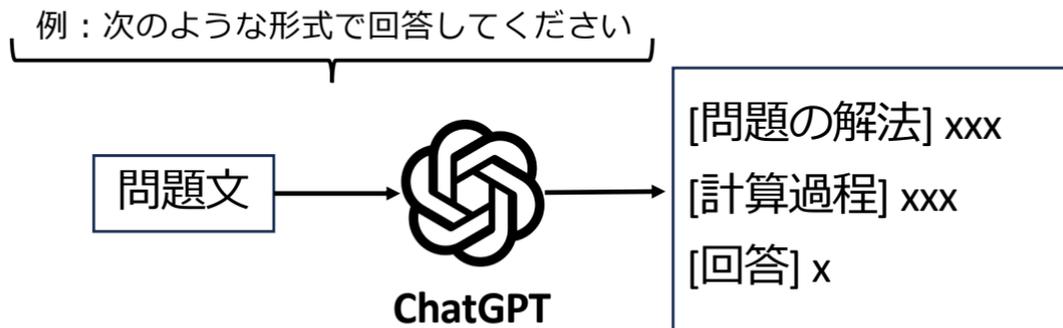
途中式にコードを含むものなど形式が一貫していない

step2 :推論時に計算過程を統一させる

目的：問題を解くために必要な知識を的確に与える

1. 問題の概要を詳細に理解させ解法を推論する
2. 推論した解法をもとに立式する

▶ データセットを整形し上記のプロセスを実現



step2 :推論時に計算過程を統一させる

以下の形式のデータセットを作成

input : 日本語の問題文

output : [問題の詳細な説明]xxx [計算過程]xxx [回答]x

training data : 79,000

validation data : 1,000

➤ ChatGPT 4o-miniを用いてMATHトレーニングデータセットを基に作成

```
{'instruction': '直角座標系において、原点から点  $(-15, 8)$  までの距離は何単位か。',  
'output': '[問題の詳細な説明]:\n直角座標系において、点の位置は座標  $(x, y)$  で表されます。この問題では、原点  $(0, 0)$  から点  $(-15, 8)$  までの距離を求める必要があります。距離の計算には、2点間のユークリッド距離の公式を使用します。この公式は次のようになります：\n\n\\[ d = \\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \\]\n\nここで、 $(x_1, y_1)$  は原点の座標  $(0, 0)$ 、 $(x_2, y_2)$  は点  $(-15, 8)$  です。\\n\\n[計算過程]:\n1. 原点の座標を  $(x_1, y_1) = (0, 0)$  とし、点の座標を  $(x_2, y_2) = (-15, 8)$  とします。\\n2. 距離を求めるための公式に代入します。\\n\\nまず、x座標の差を計算します：\\n\\[ x_2 - x_1 = -15 - 0 = -15 \\]\n\\n次に、y座標の差を計算します：\\n\\[ y_2 - y_1 = 8 - 0 = 8 \\]\n\\n3. 差の二乗を計算します：\\n\\[ (x_2 - x_1)^2 = (-15)^2 = 225 \\]\n\\n\\[ (y_2 - y_1)^2 = (8)^2 = 64 \\]\n\\n4. これらを足し合わせます：\\n\\[ (x_2 - x_1)^2 + (y_2 - y_1)^2 = 225 + 64 = 289 \\]\n\\n5. 最後に、平方根を取ります：\\n\\[ d = \\sqrt{289} = 17 \\]\n\\nしたがって、原点から点  $(-15, 8)$  までの距離は 17 単位です。[回答]17'}
```

step2 :推論時に計算過程を統一させる

【入力】

$(y = f(x))$ のグラフ上に点 $((2, 9))$ があるとき、 $(y = f(-x))$ のグラフ上に必ずある点が存在します。その点の (x) 座標と (y) 座標の和を求めなさい。

【出力】

以下はタスクを説明する指示です。要求を適切に満たす応答を書きなさい。

指示:

以下は数学の問題です。必要な公式を用い答えを求めてください。 $(y = f(x))$ のグラフ上に点 $((2, 9))$ があるとき、 $(y = f(-x))$ のグラフ上に必ずある点が存在します。その点の (x) 座標と (y) 座標の和を求めなさい。

応答:

【問題の詳細な説明】:

与えられた関数 $(y = f(x))$ のグラフ上に点 $((2, 9))$ が存在します。これより、点 $((2, 9))$ が $(y = f(x))$ のグラフ上にあることがわかります。このことから、 $(f(2) = 9)$ です。次に、求めるべき関数 $(y = f(-x))$ のグラフについても考えます。ここで $(y = f(-x))$ は、関数の反射を行うため、 (x) の値を $(-x)$ とすることにより、グラフ上の関数を変更されます。したがって、点 $((2, 9))$ は $(f(-2))$ によって変換される点を示します。

【計算過程】:

1. **元の点を確認:**

点 $((2, 9))$ から、 $(f(2) = 9)$ となります。

2. **反転の点を求める:**

反転関数 $(f(-x))$ を考えます。

$(f(-2) = 9)$ となります。

(x)

よって、点 $((2, 9))$ が反射された点は $((-2, 9))$ です。

3. **座標の和を計算:**

点 $((-2, 9))$ の (x) 座標と (y) 座標を足します。

$(x + y = -2 + 9 = 7)$ となります。

(x)

以上の計算より、 $(y = f(-x))$ のグラフ上に必ず存在する点の (x) 座標と (y) 座標の和は 7 です。【回答】7

[回答]トークン以降を解として抽出

【生成された応答（抽出後）】

7

【正解】

7

Accuracy : 0.10

課題

方程式の様な複雑な計算式を解けない

step3: 計算力が必要な問題への対応

以下の形式のデータセットを作成

input : 日本語の問題文

output : [問題の詳細な説明]xxx [pythonコード]xxx

training data : 79,000

validation data : 1,000

➤ ChatGPT 4o-miniを用いてMATHトレーニングデータセットを基に作成

```
{'instruction': '直交座標系において、原点から点  $(-15, 8)$  までの距離は何単位か。',  
'output': '[問題の詳細な説明]:\n直交座標系において、原点 (点  $(0, 0)$ ) から特定の点  
(この場合は点  $(-15, 8)$ ) までの距離を求める問題です。この距離はユークリッド距離によって  
計算されます。ユークリッド距離は次の公式で表されます:  $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ \n\nここで、 $(x_1, y_1)$  は原点の座標  $(0, 0)$  であり、 $(x_2, y_2)$  は点  $(-15, 8)$  の座標です。このため、公式に代入すると  
次のようになります:  $d = \sqrt{((-15) - 0)^2 + (8 - 0)^2}$ \n\nこの式を計算することで、点  $(-15, 8)$  までの距離を求めることができます。[pythonコード]:\n```\npython\nimport math\n\n# 座標\nx1, y1 = 0, 0\nx2, y2 = -15, 8\n\nユークリッド距離の計算\ndistance = math.sqrt((x2 - x1) ** 2 + (y2 - y1) ** 2)\n\nprint(int(distance))\n```'}
```

step3: 計算力が必要な問題への対応

【入力】

$y = f(x)$ のグラフ上に点 $(2, 9)$ があるとき、 $y = f(-x)$ のグラフ上に必ずある点が存在します。その点の x 座標と y 座標の和を求めなさい。

【応答】

以下はタスクを説明する指示です。要求を適切に満たす応答を書きなさい。

指示:

以下は数学の問題です。必要な公式を用い答えを求めてください。 $y = f(x)$ のグラフ上に点 $(2, 9)$ があるとき、 $y = f(-x)$ のグラフ上に必ずある点が存在します。その点の x 座標と y 座標の和を求めなさい。

応答:

[問題の詳細な説明]:

関数 $f(x)$ のグラフにおいて、点 $(2, 9)$ が与えられています。このことから、まず $f(2) = 9$ が成り立っています。さらに、問題では $f(-x)$ がグラフ上にある点の x 座標と y 座標の和を求めるように求められています。

まず $f(-x)$ を考えると:

```
\[
f(-x) = f(x) \quad \forall x \in \mathbb{R}
\]
```

したがって、

```
\[
f(-2) = f(2) = 9
\]
```

つまり、 $f(-2) = 9$ になります。これにより、 $f(-x)$ のグラフ上には必ず点 $(-2, 9)$ があることとなります。

次に、求められている点の座標は $(-2, 9)$ です。これにより、問題文で求められている「 x 座標と y 座標の和」は次のように表されます:

```
\[
x \text{ 座標} + y \text{ 座標} = -2 + 9 = 7
\]
```

したがって、この問題の解答は 7 です。

[pythonコード]:

```
python
x_coordinate = -2
y_coordinate = 9
result = x_coordinate + y_coordinate
print(result)
```

コードブロックを抽出し実行

【code】

```
x_coordinate = -2
y_coordinate = 9
result = x_coordinate + y_coordinate
print(result)
```

【実行結果】

7

【正解】

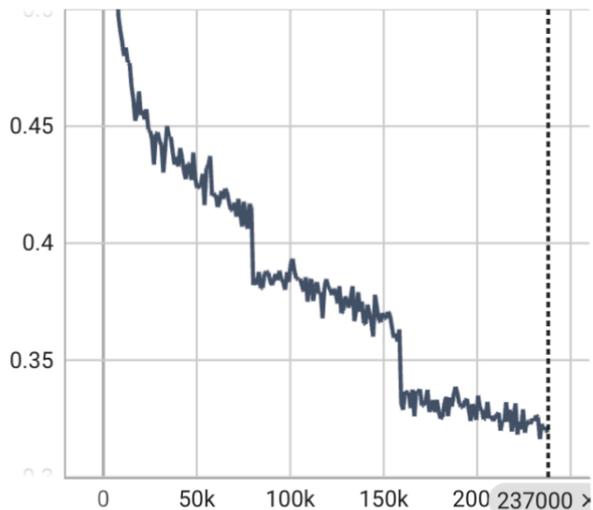
7

Accuracy : 0.12

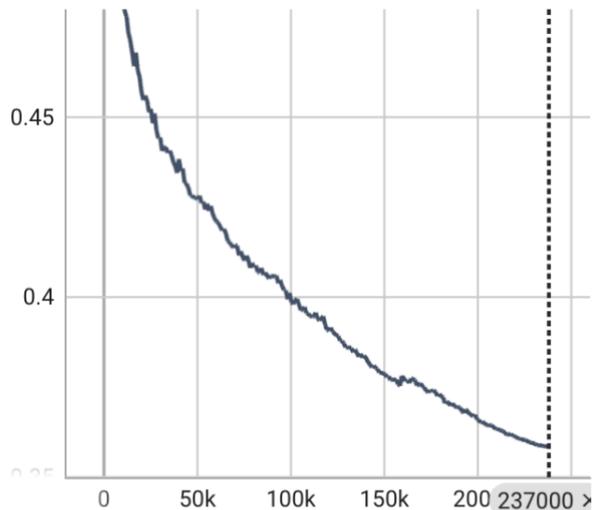
➤ コードの実行で複雑な計算が可能に

step3の条件で3epoch訓練 : **Accuracy : 0.14**

train/loss



eval/loss



3つのステップに分けて数学の問題の正解率を向上

- step1 : 指定した形式で回答させる
 - step2 : 推論時に計算過程を統一させる
 - step3 : 計算力が必要な問題に対応させる
- 最終的に**14/100問**を正解可能なモデルを構築

今後の展望

- 計算ミスをしなないモデルの構築
 - 計算途中で電卓を呼び出すなどして対応
- 回答途中でミスに気づいた時に推論しなおせるモデルの構築
 - 回答の方針を間違えた時に推論をし直せるように分岐の実装

```
template = (  
    "以下はタスクを説明する指示です。要求を適切に満たす応答を書きなさい。\\n\\n"  
    "### 指示:\\n以下は数学の問題です。必要な公式を用い答えを求めてください。{instruction}\\n\\n"  
    "### 応答:\\n{output}"  
    + tokenizer.eos_token  
)
```

```
training_args = TrainingArguments(  
    output_dir=save_path,  
    num_train_epochs=1,  
    save_steps=10000,  
    logging_steps=1000,  
    logging_dir=save_path,  
    per_device_train_batch_size=1,  
    per_device_eval_batch_size=1,  
    learning_rate=1e-4,  
    auto_find_batch_size=True,  
    evaluation_strategy="steps",  
    eval_steps=1000,  
    remove_unused_columns=False,  
    report_to=["tensorboard"],  
)
```

```
lora_config = LoraConfig(  
    r=32,  
    lora_alpha=256,  
    target_modules=["q_proj", "v_proj"],  
    lora_dropout=0.1,  
    bias="none"  
)
```