テクニカルレポート：
英語 + 日本語 数学推論データでファインチューニングした
日本語数学推論能力の検証

小島 亮一
KDDI 総合研究所
ry-kojima@kddi.com

概要

英語で書かれた大規模な数式推論データセット [2] と日本語で書かれた少量の数学推論データセット [3] を結合し，日本語で書かれた数学問題への性能向上を図ったファインチューニング事例を報告する。両者を使うことで推論時に英語が出力に混じる頻度が低減することを確認した。一方で，高度な問題への適用時には依然としてクロスリンガル・リソース面の課題が残っている。今後の展望として日本語の大規模数学データのさらなる整備や強化学習の導入，より大規模な学習環境などが考えられる。

1　はじめに

大規模言語モデルは数学のように厳密な数式操作と論理推論を要するタスクでも盛んに研究されてきている。英語圏ではステップバイステップ推論で解答へ辿り着く高品質なデータセットが豊富に存在する一方で，日本語の数学データは依然として限られている。本稿では，(1) 英語で書かれた学習データ (`open-r1/OpenR1-Math-220k`[2]), (2) 日本語で作成された (`lightblue/distilabel-reasoning-R1-Llama-70B`[3]), これら 2 つを結合してファインチューニングを行い，日本語で書かれた数学問題を解かせる実験を行った結果を報告する。

2　コンペティションについて

言語処理学会 第 31 回年次大会のワークショップ「大規模言語モデルのファインチューニング技術と評価」で開催されるチューニングコンペティション [1] では，日本語に翻訳された MATH データセットによる数学タスクが評価対象となっているが非公開データのため，本稿ではローカルでの評価結果のみを取り上げる。

3　学習手法と環境

3.1　ベースモデル

`llm-jp/llm-jp-3-13b-instruct3` を選択した。日本語の指示応答向けに学習が施されているが，高度な数学推論のデータ量は限られているとされる。

3.2　学習データセット

OpenR1-Math-220k (英語)　英語圏向けに作られた大規模数学データセットである。問題文と解答に加えて推論過程が含まれており，英語による詳細なステップバイステップ推論を学習可能である。英語と日本語の差異はあるが，数式そのものや記号操作は言語に依存しない部分が多いと考えられ，英語で身に着けた数学的な推論能力が日本語で書かれた問題を解く際にも有用か検証したい。

distilabel-reasoning-R1-Llama-70B (日本語) `lightblue/distilabel-reasoning-R1-Llama-70B` は，日本語の数学問題とその回答例・思考過程を含む少量のデータセットである。

本実験では上記の英語・日本語データセットを結合し，シャッフルを行うことでクロスリンガルな混在データを生成した。

3.3　データフィルタリング

トークン長がモデル上限 (4096) を超えるサンプルは除外し，残った約 33k サンプルを使用した。推論過程（`<think>...</think>`）や最終解答（`\boxed{}`）を含めた形式で数学的推論を学習させるかたちになる。

3.4 リソース制約

学習には単一の A100(80GB) 環境を使用し，**4bit 量子化 (QLoRA)** で効率化を図った。一方，同じデータセットで学習した Open R1[4] は H100 を複数ノード使ってフルパラメータでの学習を報告しており，その差が性能向上の度合いに大きく影響した可能性がある。

3.5 学習コード

Listing 1 に，英語 (OpenR1-Math-220k) と日本語 (lightblue/distilabel-reasoning-R1-Llama-70B) を結合して LoRA ファインチューニングした際のコードの主要部を示す。

Listing 1 英語+日本語データを QLoRA で学習するコード

```
[... 省略...]


# ========================================
# 0. 定数
# ========================================
MAX_TOKENS = 4096 # モデルの最大コンテキスト長
MODEL_NAME = "llm-jp/llm-jp-3-13b-instruct3"
EN_DATASET_NAME = "open-r1/OpenR1-Math-220k"
JA_DATASET_NAME = "lightblue/distilabel-reasoning-R1-
     Llama-70B"
OUTPUT_DIR = "./qlora-results"

# ========================================
# 1. トークナイザ& モデルロード
# ========================================
tokenizer = AutoTokenizer.from_pretrained(MODEL_NAME)

quant_config = BitsAndBytesConfig(
    load_in_4bit=True,
    bnb_4bit_compute_dtype=torch.bfloat16,
    bnb_4bit_use_double_quant=True,
    bnb_4bit_quant_type="nf4",
)

model = AutoModelForCausalLM.from_pretrained(
    MODEL_NAME,
    quantization_config=quant_config,
    device_map="auto",
)
model = prepare_model_for_kbit_training(model)

# ========================================
# 2. 設定 LoRA
# ========================================
lora_config = LoraConfig(
    r=16,
    lora_alpha=32,
    target_modules=["q_proj", "k_proj", "v_proj", "o_proj"
         ],
    lora_dropout=0.05,
    bias="none",
    task_type="CAUSAL_LM",
)
model = get_peft_model(model, lora_config)

# ========================================
# 3. データセット読込
# ========================================
en_dataset = load_dataset(EN_DATASET_NAME)
ja_dataset = load_dataset(JA_DATASET_NAME)

# ========================================
# 4. 前処理関数
# ========================================
def build_en_prompt(examples):
    texts = []
    for prob, gens, ans in zip(examples["problem"],
                               examples["generations"],
                               examples["answer"]):
        cot_text = gens[0] if len(gens) > 0 else ""
        if not cot_text.strip().startswith("<think>"):
            cot_text = f"<think>{cot_text}</think>"
        formatted_text = (
            "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"
            "次の数学の問題を解きなさい。"
            "<thinkと></thinkタグの中にステップバイステップ
                で詳しい説明を含めます。>\n\n"
            f"問題：{prob}\n\n"
            "### 応答:\n"
            f"{cot_text}\n\n"
            f"最終的な回答：{ans}"
        )
        texts.append(formatted_text)
    return {"text": texts}

[... 省略...]
```

4 推論と評価

テスト用の日本語版 MATH データは非公開のため，日本語で書かれた数学問題を手動で作成して評価した。

4.1 日本語で書かれた数学問題の例

今回の学習により，日本語のデータセットでもステップバイステップ推論生成できたため，正答が得られるケースが増加した。たとえば，下記のようなシンプルなプロンプトのみで十分であり，日本語形式の解答が生成されやすい。

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:
次の数学の問題を解きなさい。<think>と</think>タグの中にステップバイステップで詳しい説明を含めてください。

問題：x^2 + 2x + 1 を因数分解せよ。

応答:
<think> この問題では、与えられた式をよく知られたパターンで認識することがポイントです。x^2 + 2x + 1 = (x + 1)(x + 1) となります。(x + 1)(x + 1) = (x+1)^2 となります。</think>
最終的な回答：(x+1)^2

ほかにもいくつかの日本語文章問題を与えたところ，英語が出力に混じる頻度は明らかに低下した。とはいえ，複数ステップにわたる難度の高い問題では，途中で解釈が破綻したり不完全な途中計算のまま出力してしまう例も散見される。このことから，日本語データが追加された効果で基本的な日本語応答は自然になったものの，高度な推論（幾何や長い文章題など）では依然として性能が限定的であると言える。

5 考察と今後の方向性

5.1 英語データと日本語データの相補性

日本語の数学コーパスが充実していれば不要になる可能性が高いが，依然として英語のほうが数式推論に関する高品質なデータが豊富であるため，英語データの活用は有意義である。混在データで学習させることにより，日本語タスクでも英語タスクでもある程度対応しやすくなった点はメリットと言える。

5.2 大規模学習リソースとの比較

OpenR1 では H100 複数ノードを用いたフルパラメータ学習を報告している一方で，本研究では A100(80GB)1 枚による QLoRA 方式でしか学習できていない。やはり大規模学習と比較すると性能面での限界は否めない。

5.3 さらなる日本語データ拡充

本研究で取り入れた lightblue/distilabel-reasoning-R1-Llama-70B 以外にも，日本語の数学データセットや証明例などを収集して大規模ファインチューニングを行えば，より自然な日本語数学応答が期待できる。

6 まとめ

英語の数学推論データセット (OpenR1-Math-220k) に加え，日本語で書かれた数学推論データセット (lightblue/distilabel-reasoning-R1-Llama-70B) を混合して llm-jp/llm-jp-3-13b-instruct3 モデルに QLoRA 手法でファインチューニングを行った。結果として英語が混在せずに解答可能なケースが増加した。一方で，リソース制約と日本語数学データの不足から，複雑な推論問題では依然として誤答や不十分な途中経過が散見される。今後の研究では，大規模学習リソースを活用したフルファインチューニングと，高品質な日本語数学データのさらなる整備，強化学習によるアプローチなどが求められると考える。

参考文献

[1] NLP2025 ワークショップ 2. **大規模言語モデルのファインチューニング技術と評価**, https://llm-jp.github.io/tuning-competition/, (アクセス日: 2025 年 2 月 17 日).

[2] open-r1/OpenR1-Math-220k, https://huggingface.co/datasets/open-r1/OpenR1-Math-220k, (アクセス日: 2025 年 2 月 17 日).

[3] lightblue/distilabel-reasoning-R1-Llama-70B, https://huggingface.co/datasets/lightblue/distilabel-reasoning-R1-Llama-70B, (アクセス日: 2025 年 2 月 17 日).

[4] OpenR1, https://huggingface.co/blog/open-r1, (アクセス日: 2025 年 2 月 17 日).