

大規模言語モデルを用いた実世界タスク指向対話における ICL・ファインチューニングの効果の検証

佐々木 裕¹

¹豊田工業大学 知能数理研究室
yutaka.sasaki@toyota-ti.ac.jp

概要

本稿は大規模言語モデル (Large Language Model; LLM) を用いて、時々刻々変化する実世界を考慮したタスク指向対話を実現する手法について述べる。

近年、LLM は人間に近い品質の対話生成ができるまでに発展してきているが、対話モデルは汎用利用を想定して構築されており、個別のユーザの状況に応じた対話ができるわけではない。たとえば、自動走行車との対話を想定したとき、車の位置が時々刻々と変化するにつれ、最寄りのコンビニやガソリンスタンド、道路の混雑状況が変化するため、個別の状況に整合した対話を生成する必要がある。Bing Chat など検索エンジンを組み込んだ LLM も存在するが、個別のユーザが存在している実世界と LLM がリアルタイムにリンクされているわけではない。

そこで本研究では、実世界知識をプロンプトに組み込むことにより実世界対話を実現する手法を提案する。また、KVRET 対話データセットを用いて、GPT-4o や Llama 3 を含む各種 LLM におけるファインチューニングおよび ICL の効果を比較する。

1 はじめに

ChatGPT 等の大規模言語モデルの出現により、商用レベルの対話システムの構築が容易になってきている。対話研究には、タスク指向対話 [1] と非タスク指向対話 [2] の研究が存在する。ここでは、タスク指向型対話は、ホテルの部屋を予約するといった具体的な達成目標を持った対話とし、人間関係を良好にするといった心理的な達成目標は考えない。実用的な用途として、レストラン予約、ホテル予約、航空券予約、天気予報の照会など数多くのアプリケーションが考えられる。一方、非タスク指向型対話は、世間話のような特段の達成目標を持たずに、対話を継続すること自体を目標とする対話である。

現実問題を対象としたタスク指向型対話では、常

識知識以外に、現在時刻や現在位置に対応した実世界知識を使って対話する必要がある。実世界知識に基づくタスク指向対話の研究は、これまでいくつか行われている。DF-Net [3] は、対話の学習の際に表形式で与えられた実世界知識項目に対する表現ベクトルを用いて文を生成することで、実世界対話に対応しようとした。モデルアーキテクチャは、LSTM と GRU を用いたエンコーダ・デコーダモデルである。しかし、表中の別情報を誤って答える場合があるという問題点が存在した。深谷ら [4] は、対話の文脈と参照する表知識の内容の不整合を減らし、表知識の参照をより正確に行うタスク指向型対話システムの構築を目的として、表中の対象タイプを含む応答テンプレートを生成し、その文脈表現を利用して対話文を生成する手法を提案したが、複雑な表参照のケースには対応できていなかった。そこで、本稿では LLM により実世界知識に基づく対話を実現する。

2 タスク設定

表 1 表形式の実世界知識の例

対象	距離	交通状況	タイプ
Shell	3 miles	heavy	gas station
Valero	4 miles	no traffic	gas station
...

本稿で想定するタスク指向対話は、ユーザの持つ目標を達成する対話を実世界知識に整合する形で実現することである。たとえば、表 1 の実世界知識の下で次のような対話を実現する。

User: I need gas.

System: There is the Shell and Valero.

User: Which one is the closest to my position?

System: Valero is the closest, considering the heavy traffic around the Shell.

User: Okay, please navigate me to Valero.

この例のように、単なる表からの項目の抜き出しができる大規模言語モデルを実現しようとしているのではなく、現実の世界の状況に合わせて、常識的な対話することを目標としている。距離的に近い Shell を推薦するのではなく、交通状況と距離の両方を考慮して、ユーザにとってより適切な Valero を推薦することが対話の成功とする。ただし、従来研究の評価法に従い、タスク成功率ではなくシステムの生成文と想定する参照対話文との類似度により性能を評価する。

本研究の貢献は以下の4点である。

- 大規模言語モデルを用いて実世界情報を考慮したタスク指向型対話の実現する手法を提案する。実世界情報をプロンプトに与えることで、実世界知識を参照するタスク指向型対話が可能であることを示した。
- 様々な LLM について、実世界対話の品質を評価指標 BLEU および Sentence Similarity により評価し、性能を定量的に評価し比較した。
- 実世界知識を含むプロンプトにより、大規模言語モデルに対するファインチューニングおよび In-Context Learning (ICL) [5]の効果を明らかにした。特に、GPT 系の LLM は学習データが明記されていないため、実世界情報がなくても想定する対話を生成できる可能性があったが、評価実験でそのようなことは起きていないことを示した。
- 大規模言語モデルに対して、ファインチューニングに用いる対話数の効果を示した。

3 LLM

現時点で、オープンソース、クローズドソースのいくつかの大規模言語モデルが利用可能である。

OpenAI による LLM として、GPT-4o, GPT-4o mini, GPT-3.5-Turbo が利用可能である。パラメータサイズは公開されていないが、GPT-4o は 1.76T 程度、GPT-4o mini が 8B 程度との説もある。これらはソースコードを公開していないが、有料の機能とし、指定した学習データによりファインチューニングを適用したモデルを API を通して利用することができる。

オンプレミスで実行可能な大規模言語モデルとして、Meta の Llama 3, 3.1, 3.2 が利用可能である。パラメータ数が 3B, 8B, 70B など異なるサイズのモデルが存在する。深層状態空間モデルに基づく LLM として Mamba が存在する。これらのモデルはオープ

ンソースであり、直接ファインチューニングを適用することができる。

4 提案手法

LLM を用いた実世界対話を実現するために、以下の手法を提案する。①プロンプトのインストラクションに実世界知識を含め、これを参照するようなプロンプトを設計する。対話毎に実世界知識を変えることでその時点の状況に合わせた実世界対話を実現することができる。②インストラクションを含む対話例を訓練データとして与えて LLM をファインチューニングする。

4.1 プロンプトの設計

実世界知識の表現法としてオントロジーや知識グラフなど様々な様式が想定できるが、本研究では単純に実世界知識が表形式で与えられているとする。このように設定したのは、実世界知識を表形式で与えた対話データセット KVRET [6]が存在することが理由のひとつである。ただし、2 節で説明したように表形式の実世界知識であるからといって、PhoneBook [7]のような単純な表からの項目の抽出問題を対象タスクとしているのではなく、表に書かれている項目間の関連性や比較などを通して、人間が行うのと同様な常識に基づく対話を想定する。

Chat プロンプトのインストラクションである

```
{"role": "system", "content": ...}
```

の content は次のように設計した。

```
"You are a helpful <TARGET SYSTEM>. Please concisely reply based on the following information. <KNOWLEDGE>"
```

<TARGET SYSTEM> は、ナビゲーションタスクでは”car navigation system with an integrated GPS”, 天気予報タスクでは”weather forecast system”, スケジュール管理タスクでは,”personal scheduling system”とした。

これに続き、対話ごとにインストラクション中に実世界知識<KNOWLEDGE>を与える。大量の実世界知識をすべて与えるのではなく、現在の状況にあった実世界知識が選択されていることを想定する。たとえば、カーナビゲーションであれば、近隣の施設やユーザの家の位置が外部システムにより選択され、表形式で与えられていると想定する。スケジュール管理では、数日のユーザのスケジュールが抽出さ

れており、天気予報では近隣の天気情報が選択されているとする。ICL の場合はこれに続いて学習データのうちの 10 件を対話例として与える。

表 1 の実世界知識は、JSON 形式の文字列に変換してインストラクションに組み込む。具体的な例を挙げる。

```
{"distance": "2 miles",
 "traffic_info": "road block nearby",
 "poi_type": "parking garage",
 "address": "550 Alester Ave",
 "poi": "Dish Parking"}
```

インストラクションに続いて、ユーザとアシスタントの対話を交互に与える。アシスタントの発話が生成対象である。

```
[{"role": "user", "content": ...}
 {"role": "assistant", "content": ...}
 {"role": "user", "content": ...}
 {"role": "assistant", "content": ...}
 ...
 {"role": "user", "content": ...}]
```

これで 1 つの対話に対するプロンプトが完結する。なお、対話実験においては、最初のユーザの発話を与えたときのシステムの生成文を評価した次の対話生成では、生成された文ではなく、参照文を生成したときのユーザの発話を与えて文を生成する。つまり、システムが生成すべきアシスタントの対話文の前までの正しい対話を埋め込んだプロンプトを作成し、次の対話文を生成させ評価するというステップをアシスタントの発話の数だけ繰り返す。

4.2 ファインチューニング

大規模言語モデルに対するファインチューニングは対象の大規模言語モデルによって多少異なる。基 GPT-4o, 4o Mini は OpenAI のサーバ上でファインチューニングを行った。Llama は SFT+unsloth を用いて LoRA (Low Rank Adaptation) [8]を適用した。

提案のプロンプトの設計のとおり、各対話に対して、インストラクションから最後の対話文までを含むプロンプトを{10, 100, 2243}件、train set から抽出して使用した。

ファインチューニングを行わなくても、大規模言語モデル自体の学習の過程で評価用のデータセットを学習してしまっている可能性があるため、ファイ

ンチューニング前の性能、および実世界知識を与えない場合の性能とファインチューニング後の性能を比較することで、LLM の事前学習時における dev および test データの漏洩(Leakage)を検証する。

5 実験

5.1 実験設定

大規模言語モデルとして下記を比較する。

- GPT-4o, GPT-4o Mini, Llama 3, Mamba, Claude 3.5 Opus/Sonnet/ Haiku

KVRET データセットは、train, dev, test が 2243, 301, 303 件 (不完全なデータ削除後)。下記の 4 つの実験設定で性能を比較する。再現性を確保するためサンプリングなし、または temperature=0, 最大生成トークン数 40 とした。

- **実世界を含むプロンプトによる対話生成**: 提案手法のプロンプトを用いて実世界知識を与えるが、ファインチューニングは行っていない。
- **実世界知識なし(-WK)**: 単に対話履歴のみから、対話文がどの程度生成できるかを調査する。もし、LLM の学習時に KVRET データセットを記憶していれば、実世界知識がなくても対話を再現できる可能性があるために設定した実験。
- **ファインチューニング(+FT)**: GPT 系は OpenAI のサーバ上でファインチューニングを実施。Llama は SFT+unsloth を用いて LoRA ファインチューニングを実施した。GPT も 2243 対話に対する学習が約 35 分で完了したことから、軽量ファインチューニングであると推察される。
- **In-Context Learning (+ICL)**: 提案手法のプロンプトのインストラクション部分に 10 対話例を追加し、プロンプトチューニングを行った。

評価尺度として BLEU¹および Sentence Similarity²モデルによる文の埋め込みベクトル間の cos 類似度を用いる。BLEU は n-gram の上限を n=1~4 で変えたときの結果をそれぞれ報告する。BLEU では単語の表層の n-gram の一致を取るため、内容的には同じでも表現が異なることでスコアが低く見積もられるため Sentence Similarity も計算した。実験には A6000 (48GB)×2 搭載の GPU サーバを用いた。

¹ torchmetrics.functional.text

² sentence-transformers/all-MiniLM-L6-v2

5.2 実験結果

test set データについての実験結果を表 2 に示す. 各種パラメータは dev set でチューニングしたものを使用. 表中の表記は, GPT-4o, GPT-4o Mini (4o Mini), Llama 3.2 3B Instruct, 3 8B Instruct, 3,1 70B Instruct (Llama32 3b, Llama3 8b, Llama31 70b), Claude 3.5 (Opus, Sonnet, Haiku), Mamba 7b, 実世界情報なし (-WK), 10 対話による In-Context Learning (ICL10), N 対話チューニング (FT N), 1~4-gram BLEU (BLEU1~4), Sentence Similarity (Sim).

従来法 DF-Net の BLEU 14.4 に対し, GPT-4 系にファインチューニングを行うことで BLEU4 が約 23 という高い性能を達成できた. ファインチューニングや ICL を適用していない時点では性能は低く, 世界知識を与えないとさらに性能が低下することから, LLM の事前学習において対話データを記憶しているわけではないと考える. Llama 3.1 70B もファインチューニングにより GPT-4o に迫る性能が得られた. 全般的に ICL の効果はファインチューニングを下回ることが確認できた. Mamba7b は Llama 70b に近い性能を示したが unsloth が対応していないため, GPU メモリの範囲ではファインチューニングを実行できなかった. また, Mamba +ICL の性能が低下したのは, 1000 トークン近い長いプロンプトに弱く, 生成文が空になったり, プロンプトの一部をそのまま生成したりする現象が発生したためである.

開発データに対する結果 (付録 B 表 3) の生成文を目視で分析した結果, ファインチューニングにより, 返答の長さなどのデータセット特有の表現に適応したことと, 実世界知識の参照の正確性が高くなったことの両面が存在した. 特に, ファインチューニングされた GPT-4o の BLEU 値は 23 程度であるが, 内容的には正解の対話文と遜色ない.

6 おわりに

実世界知識に基づいて対話を行う大規模言語モデルを提案し, KVRET データセットにより検証を行った. 実験の結果, GPT-4o にファインチューニングを行うことで, 高い性能が得られることがわかった. Llama 70B モデルもファインチューニングにより, GPT-4o Mini に近い性能を得ることができていることが分かった. 今後の課題として, Mamba 7B や Llama 405B モデルなどをファインチューニングすることで GPT-4o を超えられるどうかを検証していきたい.

表 2 KVRET(test set)での結果

Model	BLEU1	BLEU2	BLEU3	BLEU4	Sim.
GPT-4o	24.61	17.38	12.55	9.14	61.88
-WK	11.41	6.12	3.72	2.49	50.01
+ICL10	32.56	23.95	17.93	13.60	66.94
+FT10	29.77	21.88	16.48	12.59	64.36
+FT100	39.15	29.61	22.85	17.96	70.38
+FT2243	42.24	34.18	28.05	23.28	72.42
4o Mini	21.61	15.10	10.79	7.81	61.04
-WK	12.68	6.93	4.24	2.78	51.58
+ICL10	28.86	20.88	15.49	11.49	66.66
+FT10	32.87	24.39	18.47	14.15	65.95
+FT100	40.81	31.08	24.11	19.01	70.52
+FT2243	42.49	33.83	27.26	22.19	71.71
Llama32 3b	12.98	7.93	5.02	3.32	49.02
-WK	8.39	4.01	2.20	1.27	44.63
+ICL10	16.30	10.50	6.94	4.70	54.66
+FT10	28.83	20.15	14.14	9.99	59.30
+FT100	32.48	23.14	16.86	12.48	65.10
+FT2243	33.22	23.74	17.05	12.43	65.57
Llama3 8b	15.13	9.51	6.25	4.19	50.76
-WK	12.87	7.13	4.24	2.66	57.09
+ICL10	22.76	15.36	10.63	7.55	60.14
+FT10	27.18	18.08	12.35	8.46	59.28
+FT100	34.35	24.47	17.94	13.39	67.35
+FT2243	36.54	27.25	20.46	15.59	67.81
Llama31 70b	17.22	11.41	7.82	5.51	53.60
-WK	9.45	4.50	2.50	1.45	43.09
+ICL10	28.92	20.73	15.20	11.18	66.98
+FT10	29.26	20.43	14.74	10.79	61.95
+FT100	36.77	27.53	20.82	15.90	69.02
+FT2243	40.03	30.52	23.41	18.06	70.37
Claude Opus	15.68	10.55	7.30	5.16	53.95
Sonnet	18.32	12.37	8.44	5.74	62.31
Haiku	17.11	11.66	8.01	5.52	59.73
Mamba 7b	18.47	11.95	7.95	5.45	56.79
-WK	12.89	7.48	4.70	3.10	49.42
+ICL10	14.88	8.86	5.49	3.50	47.84

謝辞

本研究の一部は JSPS 科研費 JP23K11237 の助成を受けた。豊田工大知能数理研究のメンバーに感謝する。

参考文献

1. Diane J. Litman and James F. Allen, Discourse Processing and Commonsense Plans, Chapter 17, in *Intentions in Communication*, eds. Philip R. Cohen, Jerry Morgan, and Martha E. Pllack, MIT Press, pp. 365-388, 1990.,
2. T. Bickmore and J. Cassel, Small Talk and Conversational Storytelling In Embodied Conversational Interface Agents, AAAI Technical Report, FS-99-01, 1990.
3. Libo Qin, Xiao Xu, Wanxiang Che, Yue Zhang, Ting Liu, Dynamic Fusion Network for Multi-Domain End-to-end Task-Oriented Dialog, ACL 2020.
4. 深谷竜暉, 三輪誠, 佐々木裕, タスク指向対話システムの外部表知識の参照能力向上, 言語処理学会第28回年次大会, 2022.
5. Tom B. Brown et al, Language models are few-shot learners, 34th International Conference on Neural Information Processing Systems, 2020.
6. Mihail Eric, Lakshmi Krishnan, Francois Charette, Christopher D. Manning, Key-Value Retrieval Networks for Task-Oriented Dialogue, in Proc. of 18th SIGDIAL, 2017.
7. Roger Waleffe et al., An Empirical Study of Mamba-based Language Models, arXiv:2406.07887, 2204.
8. Edward J Hu et al., LoRA: Low-Rank Adaptation of Large Language Models, International Conference on Learning Representations, 2022.

付録 A

KVRET の (dev set) を開発用データセットとして開発・デバッグ・ファインチューニングを行った。開発用データセットに対する結果を表 3 に示す。

付録 B

KVRET の (dev set) の中から対話例を示す。Generated が生成された対話。それ以外は参照対話。

(GPT-4o Mini +FT2243)

(n0) User: i'm looking for a place to eat can you show me directions to the nearest fast food restaurant

(n1) Generated: The nearest restaurant is Pizza My Heart, 6 miles away.

(n1) Assistant: There is Pizza My Heart 6 miles away: is it okay for you?

(n2) User: Sure, let's go there, just pick a route with no traffic please

(n3) Generated: There is moderate traffic on our way but I sent on your screen another route that will allow you to reach faster

(n3) Assistant: I highlighted its address on the map, drive carefully!

(Llama 3.1 70B +FT2243)

(n0) User: i'm looking for a place to eat can you show me directions to the nearest fast food restaurant

(n1) Generated: We're 6 miles away from Pizza My Heart, would you like directions?

(n1) Assistant: There is Pizza My Heart 6 miles away: is it okay for you?

(n2) User: Sure, let's go there, just pick a route with no traffic please

(n3) Generated: Setting navigation for Pizza My Heart, 528 Anton Ct, with moderate traffic.

(n3) Assistant: I highlighted its address on the map, drive carefully!

付録 C

ファインチューニングのパラメータ。Random seed=3407. GPT: バッチサイズ 4, エポック数 3, 10, LR 乗数 1.8, Llama: LoRA $\alpha=16$, $r=16$, dropout=0, target=q_proj, k_proj, v_proj, o_proj,

表 3 KVRET(dev set)での結果

Model	BLEU1	BLEU2	BLEU3	BLEU4	Sim.
GPT-4o	24.64	17.42	12.16	8.76	61.88
-WK	12.27	7.03	4.44	2.96	50.80
+ICL10	31.97	23.35	17.16	12.76	66.43
+FT10	30.39	22.47	16.82	12.76	64.61
+ FT100	38.18	28.85	22.25	17.40	70.28
+FT2243	40.51	32.35	26.29	21.51	71.55
4o Mini	21.97	15.33	10.87	7.79	60.70
-WK	13.32	7.68	4.87	3.27	52.11
+ICL10	29.27	21.14	15.65	11.73	66.22
+FT10	32.95	23.94	17.80	13.63	64.71
+FT100	38.52	30.31	23.91	19.15	70.62
+FT2243	41.27	33.11	26.95	22.16	70.90
Llama32 3b	12.46	7.31	4.53	2.89	48.54
-WK	10.11	5.08	2.89	1.69	47.56
+ICL10	17.81	11.57	7.77	5.33	55.17
+FT10	28.31	19.74	13.81	9.75	59.39
+FT100	31.93	23.02	16.89	12.55	64.68
+FT2243	32.22	22.99	16.49	12.13	63.91
Llama3 8b	17.95	11.56	7.76	5.49	62.12
-WK	12.53	7.24	4.56	3.04	56.43
+ICL10	22.17	14.99	10.45	7.42	58.78
+FT10	27.75	18.99	13.09	9.02	58.61
+FT100	33.86	24.38	17.84	13.26	67.08
+FT243	33.92	25.11	18.91	14.33	67.00
Llama31 70b	17.01	11.23	7.81	5.58	53.50
-WK	10.44	5.58	3.43	2.26	44.38
+ICL10	28.29	20.00	14.43	10.35	65.92
+FT10	30.67	21.50	15.69	11.50	61.89
+FT100	35.37	26.67	20.56	16.25	67.50
+FT2243	38.68	29.27	22.48	17.45	69.42
Claude Opus	15.27	10.09	6.76	4.62	53.67
Sonnet	17.74	11.65	7.71	5.16	61.12
Haiku	16.44	10.97	7.36	5.02	58.60
Mamba 7b	19.01	12.47	8.40	5.82	56.69
-WK	13.86	8.45	5.41	3.65	51.05
+ICL10	16.08	10.52	7.22	5.24	48.32

gate_proj, up_proj, down_proj, unsloth:
load_in_4bit