

不完全な指示に対する対話型情報補完システムの検討

Consideration of Interactive Information Completion System for Incomplete Instructions

橋本 慧海* 白松 俊
Ekai HASHIMOTO Shun SHIRAMATSU

名古屋工業大学
Nagoya Institute of Technology

Abstract: Efficient communication is indispensable in modern society, with directives playing a central role within it. However, the ambiguity and vagueness of directives can lead to inefficient communication, hindering organizational and individual activities. In this study, we propose a system that utilizes large-scale language models to engage in dialogue with the recipient of a directive, transforming unclear instructions into clear ones. By adopting this approach, we aim to supplement ambiguous directives and enhance communication efficiency. In the prototype implementation, using a fictitious manual for spaceship crews, we confirmed the system's ability to generate effective instructions. Future research will need to consider the gap between experimental and real-world environments.

1 はじめに

現代社会において、効率的なコミュニケーションは不可欠であり、それにより個人や組織の活動が円滑に進行している。特にコミュニケーションの中でも、指示は要求を伝えるために普遍的かつ効果的な手段である。指示は、組織内のトランザクティブメモリへの理解や、豊かな個人間のコミュニケーションを必要とする。これらの要素が不足すると、指示が不明瞭で曖昧になり、非効率的なコミュニケーションを招く。

指示を出す側には、高いメタ認知能力や組織メンバーの知識への理解が求められる。これは出した指示が受け取り手によってどのように理解され、実行されるかを正確に予測する必要があるからだ。同様に、指示を受ける側も、受けた指示が適切であるかを判断し、必要に応じて不完全または不明確な指示を指摘するために高いメタ認知とタスクへの理解が求められる。このような相互の認知プロセスは、緊急性の高いタスクやストレスの下では特に困難となる可能性がある。こういった環境で発生した不完全な指示は誤解や混乱を招き、特に医療や工事現場などでは、重大なリスクとなりえる。

以上のように指示は普遍的な手段にもかかわらず、指示の出し手、受け手のどちらにも高い能力を要求される。ここから日常的に生じる指示の問題に対し、情報

の不足や曖昧さを効果的に補完し、タスクの完遂を支援する新しいアプローチが必要だと考える。

そこで本研究では、大規模言語モデルを用いて指示の受け手との対話によって指示の詳細情報を聞き出し、不完全な指示を完全な指示へと変換するシステムを検討する。このシステムを用いることで、コミュニケーションの不足を補い指示の受け手がタスクを完遂できることを期待する。

2 関連研究

2.1 対話システムと意図推定

従来の対話システムは、ユーザ自身の発話からそのユーザの意図を推定し、それに基づいて情報を提示するものであった。例えば、河原ら [1] や伊藤ら [2] の研究では、ユーザの発話を基にその意図や選好を推定して情報を提供している。しかしながら、本研究では新たなアプローチを採用し、ユーザとの対話を通じて第三者（指示者）の意図を推定し、不完全な指示を補完することを目指している。これにより、ユーザが受け取った指示が不明瞭である場合でも、対話を通じて指示の内容を明確にすることが期待される。

*連絡先：名古屋工業大学
E-mail: e.hashimoto.611@stn.nitech.ac.jp

2.2 GPT-4

GPT-4(Generative Pre-trained Transformer 4) [3] は OpenAI によって提供されている言語モデルである。API を介して利用でき、文章の生成、文章の要約、質問への回答、翻訳などに活用することができる。しかし、学習データは 2022 年 8 月までのデータを用いており、ハルシネーションと呼ばれる間違った情報を生成することがある。

2.3 LlamaIndex

LlamaIndex[4] は大規模言語モデルと外部データを接続するためのインターフェースを提供するフレームワークである。これによって GPT が学習していない最新の情報やユーザが用意したマニュアル等に対応し、GPT が文章を生成することが可能になる。

3 提案手法

3.1 システム概要

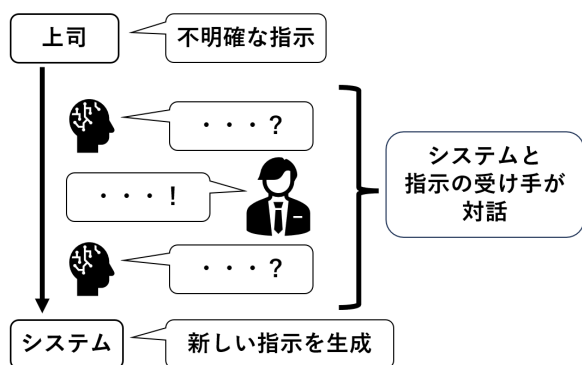


図 1: システム概要図

以下に図 1 の詳細を示す。

- 指示者（上司）がユーザ（部下）に指示を出す
- システムとユーザの対話が始まる
 - ユーザから指示文を聞き出す
「マフラーをチェックしといて」
 - ユーザがその指示を受けた状況を尋ねる
「中古車を査定していた」
 - ユーザが思う上司の指示の目的を尋ねる
「排気系統全体の状態を確認してほしい」
 - ユーザが思う上司の指示の意図を尋ねる
「不備を見つけ安く買い取りたい」
- 抽出した指示、状況、目的、意図を基にマニュアルから適するタスク手順を検索

4. ユーザが理解できる新しい指示を生成

このようにユーザとの対話を通して、指示者の意図を推定し、適切な指示を生成する。ユーザが上司の指示の目的、意図を理解できない場合もあるので、その場合は不明なままタスク手順を検索する。

3.2 マニュアル作成

プロトタイプの実装に際して、架空の宇宙船クルーのためのマニュアルを作成した。このマニュアルには 43 個のタスクが含まれており、すべてのタスクは構造化されている。以下にそのタスク毎の項目とその例を示す。

- タスク名 (例：エアフィルターの交換)
- タスク目的 (例：船内の空気品質を維持する)
- 手順 1 (例：交換が必要なフィルターを特定する)
- 手順 2 (例：T 型レンチと平行クリップを用いて交換する)
- 手順 3 (例：空気品質を確認する)

4 プロトタイプの動作

本システムの有効性を検証するために、3.2 で言及したマニュアルを用いてプロトタイプを作成した。

S	指示を教えてください。
U	道具を持って来いって
S	その指示が出た状況を教えてください
U	エアフィルターを点検しました
S	指示を出した人の目的はわかりますか？
U	たぶんエアフィルターを交換したいんだと思います
S	指示を出した人の意図はわかりますか？
U	分かりません
S	新しい指示を考えているので、お待ちください。(検索中)
S	T 型レンチと平行クリップを持って行ってください。

図 2: 実際の対話 (S=システム, U=ユーザ)

図 2 はプロトタイプでの実際の対話を示した。この対話では 3.2 の例で示したように**エアフィルターの点検**を行っている。マニュアルを検索すると点検には**T 型レンチと平行クリップ**が必要であること分かる。そのためシステムはその**2 つの工具**を持っていくようにユーザに新しく指示した。

5 課題

本プロトタイプの実装にはいくつかの主要な問題点が浮かび上がった。これらの課題を克服するための方法やアプローチをさらに検討する必要がある。

5.1 マニュアルの構造と現実性

使用したマニュアルは明確に構造化されており、43のタスクを持つ。一方、実際の業界のマニュアルはしばしばテキストベースであり、多岐にわたる情報や条件を伴うことが多い。この差異を解消するため、システムがテキストベースのマニュアルも効率的に扱えるように改善する必要がある。

5.2 ユーザーの既存知識の取り扱い

現在のシステムはユーザーの既存の知識や経験を十分に考慮していない。特に、指示を理解し実行する過程でユーザーがもつ前提知識は、指示の解釈やタスクの実行方法に影響を与える可能性がある。この側面をシステムに組み込むことで、よりの確な支援や指示の提供が期待できる。

5.3 実際のタスクの複雑性

実際の業務やタスクは多岐にわたる分岐や条件を伴うことが一般的である。このような複雑さをシステムが取り扱うことができるかどうかは、実用性の観点から考慮すべき課題である。

6 まとめ

本研究では、大規模言語モデルを活用し、ユーザーとの対話を通じて不完全な指示を完全なものへと補完する新しいシステムの検討を行った。実装されたプロトタイプを用いて、架空の宇宙船クルーのマニュアルを基に実験を行った。結果、システムが指示の内容を明確化し、ユーザーがタスクを適切に完遂するための新しい指示を生成することが示された。

ただし、現行のプロトタイプにはまだ課題が残っている。特に、現実のマニュアルとの整合性や、ユーザーの知識データベースを適切に考慮する点での改善が求められる。これらの課題を解決することで、本システムはさらに効果的なコミュニケーションツールとしての可能性を秘めていると考えられる。

今後は指示者の視線 [5][6] から、指示の受け手の推定などを行うことを想定している。

今後の研究では、これらの課題を克服するための手法の検討や、より広範な実験を行うことで、システムの有効性や適用範囲をさらに明らかにしていく予定である。

謝辞

本研究の一部は、JST CREST (JPMJCR20D1) および NEDO (JPNP20006) の支援を受けた。

参考文献

- [1] 達也 河原, 宏彰 川嶋, 高嗣 平山, and 隆司 松山. 対話を通じてユーザの意図・興味を探り情報検索・提示する情報コンシェルジェ. *情報処理*, 49(8):912–918, 08 2008.
- [2] 伊藤亮介, 和範 駒谷, and 達也 河原. 機器操作マニュアルの知識と構造を利用した音声対話ヘルプシステム. *情報処理学会論文誌*, 43(7):2147–2154, jul 2002.
- [3] OpenAI. Gpt-4 technical report, 2023.
- [4] H. Chase. Welcome to langchain.
- [5] 幸大 石正, 秀輝 島田, and 健哉 佐藤. ドライバの視線情報にもとづく運転意図推定. In **第 77 回全国大会講演論文集**, volume 2015, pages 283–284, mar 2015.
- [6] 小川原. 視線運動からの意図推定に基づいたロボットによる行動支援. *インタラクション 2005 情報処理学会シンポジウム*, 2005.