

脳波デバイスを用いた人と大規模言語モデルのコミュニケーション形式の検討

A Study of Communication Formats for Human and Large Language Models Using EEG Devices

茶谷 瑛佑[†] 佐藤 啓宏[†]
EISUKE CHATANI[†] YOSHIHIRO SATO[†]

1. はじめに

2022 年 12 月頃から ChatGPT を始めとする大規模言語モデル (Large Language Models :LLMs) が大きな注目を集めている。とくに自然言語処理が大きく発展したことにより、言語に基づく多くのインターフェースやアプリケーションが製作されている。その中でも、とくにチャットボットのような対話型プログラムが実用化される段階まできている。しかし、テキストや音声インターフェースとすることは便利であるものの、一方で、まだキーボードが必要だったり、発声を必要としたりするなどの煩わしさや手間と時間がかかるといった問題点がある。この問題を解決する一つの方法として、脳波を用いて人と対話する大規模言語モデルのインターフェースを作成することで、これらの問題を解消し、豊かな対話を実現できるのではないかと考えた。しかし、現状では脳波デバイスの制限が多いので、どのような会話を実現できるかが未知数である。

そこで、本研究の目的は、人と LLMs の対話において脳波デバイスを介在させ、成立可能なコミュニケーション形式を検討し、整理することである。

2. 関連研究

脳波デバイスに関連した研究は以前から行われており、De-wangga ら [1] は、EEG を使用してブレイン-コンピュータインターフェース (BCI) を開発し、機械学習と組み合わせる EEG 信号を適切な動作コマンドに変換する分類モデルを作成した。彼らの BCI システムは、ロボットの動作制御に成功している。また、Lotte ら [2] は、過去 10 年間の BCI 設計における EEG 信号の分類に関する広範な調査を実施した。彼らの調査では、適応型分類器、転移学習、Riemannian ジオメトリ分類器、縮小 LDA、ランダムフォレスト等が、信号分類の有望な方法として挙げられている。

同様に、Antonello ら [3] は、fMRI データを使用して、言語と音声に関連する AI モデルを作成した。このモデルを用いて言語や音声の処理に関与する脳の領域を特定し、言語や音声情報と脳の活動がどのように関連するかを解明することが可能となった。これらの研究は脳波デバイスを活用した新たなインターフェース開発の基礎となる。

人と AI やロボットの自然なコミュニケーションを探求する研究も、多数存在する。例えば、Nichols ら [4] は、卓上ロボッ

ト Haru を開発し、人間とロボットのコミュニケーションの質を向上させることを目指した。Haru は、人間との関係性を強化するための会話戦略に基づき、関与度、パーソナリティ特性、関係性の期待といった要素に対して有益な効果をもたらした。また、Amershi ら [5] は、人間と AI の相互作用を改善するための 18 の AI デザインガイドラインを提案し、人間中心の AI システムの設計と実現に向けた方向を示している。しかしながら、これらのコミュニケーション形式の多くは音声やテキストベースであるため、次世代の LLM の能力を最大限に活用するには、限界が存在する可能性がある。

我々の研究では、脳波デバイスを用いることでこれらの限界を克服し、人と LLM との間でより深いコミュニケーションを可能にする新たな形式を探求したい。それにより、人間の思考や意図をより直接的かつ正確に LLM へと伝達することが可能になり、LLM の回答も適応的かつ精度の高いものとなるだろう。このような進化したコミュニケーション形式は、AI と人間との相互作用を新たな次元へと引き上げ、さらに高度なコミュニケーション体験を実現する可能性を秘めている。

3. コミュニケーション形式の検討

3.1 脳波デバイス NextMind

本稿では、脳波デバイス NextMind[6] を対象に成立可能なコミュニケーション形式を検討する。NextMind とは非侵襲的な BCI であり、このデバイスでは 9ch の EEG 電極を用いて脳の後頭葉にある視覚情報に対する脳の反応を読み取ることができる。EEG 信号の取得から推論結果を出力するまでをデバイスが行うため、簡単に扱うことができるという利点がある一方、中間の処理結果のデータにアクセスすることはできないという制限がある。

NextMind の使用法は、画面上に表示される NeuroTag と呼ばれるタグがつけられたオブジェクトにユーザーが注目すると、それが ON/OFF の 2 値状態で認識され、タグ付けされたオブジェクトに設定されたアクションが実行されるという仕組みである。このアプリケーションとしては、TV の電源やチャンネルを操作するリモコン代わりに使用することや、ゲームの画面上のオブジェクトの操作を行うことが提案されている。

3.2 会話形式

脳波デバイスを介在させたコミュニケーションの会話形式としては、

[†] 京都先端科学大学 Kyoto University of Advanced Science

- (1) LLM によるユーザーへの提案形式
- (2) ユーザーが LLM 同士の会話に介入する形式
- (3) LLM がユーザー同士の会話のアシスタントを行う形式
- (4) ユーザーが注目するオブジェクトや画像に関連する情報を LLM が提供するという形式
- (5) あらかじめ設定されたコマンドを選択肢として LLM と対話を行う形式
- (6) ユーザーが正しい順番通りに入力できないと LLM が応答をしないことで、ユーザーの学習を促す形式
- (7) 入力待機時間によって LLM の文章生成の内容を変更する形式

が挙げられる。次に、紙面の制限から (1)~(3) の形式について具体例を検討する。

3.2.1 LLM によるユーザーへの提案形式

図 1 は、LLM がチャットボットとしてユーザーと質問形式での会話を行う図である。LLM はユーザーへ 2 つの選択肢から成る質問を生成し、ユーザーは脳波デバイスを用いて選択肢の中から一つの回答を選ぶ。その選択を基に、関連した 2 つの選択肢から成る質問が生成される。この操作を繰り返すことで、ユーザーの好みや興味に沿った提案が生成され、出力される。

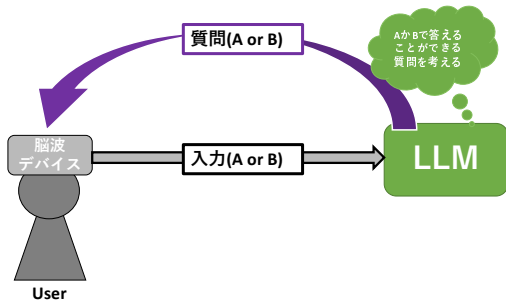


図 1 LLM によるユーザーへの提案形式

3.2.2 ユーザーが LLM 同士の会話に介入する形式

図 2 は、別々の人格を持ったチャットボットを LLM で作成し、その 2 つの LLM が会話を行い、ユーザーがその会話を聞き、好きなタイミングで会話に介入する形式を示している。チャットボットが話している内容から、さらに詳しい情報を知りたいと思ったとき、ユーザーが脳波デバイスを介して反応することで、直前に話していた内容をさらに深く掘り下げて会話を続ける。この形式のメリットは、自分が興味のない内容について話されている場合、あまり会話に介入せずに傍聴するという、多人数のコミュニケーションに近い形式を一人で行うことができる点である。

3.2.3 LLM がユーザー同士の会話のアシスタントを行う形式

図 3 は、LLM がアシスタントチャットボットとなり、親ユーザーと子ユーザーの間の会話を記憶するという形式を示している。チャットボットが子ユーザーからの入力を検知した際、会話内容の要約や、直前の内容と一致するように子ユーザーに

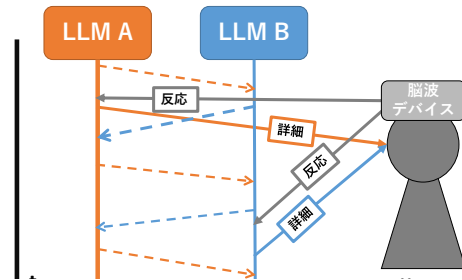


図 2 ユーザーが LLM 同士の会話に介入する形式
 応答する。この形式のメリットは、子ユーザーが親ユーザーの話を理解できなくても、チャットボットの支援により円滑なコミュニケーションを行うことができるという点である。

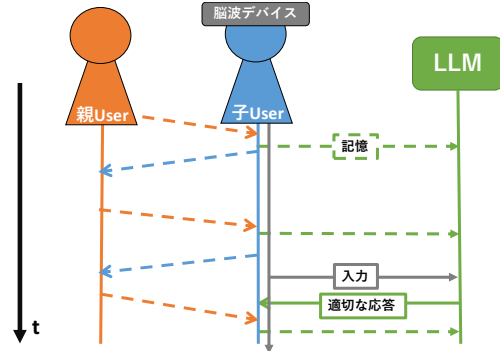


図 3 LLM がユーザー同士の会話のアシスタントを行う形式
 一部の例を紹介しただけであるが、以上から、LLM と脳波デバイスを組み合わせることで、ユーザーと様々な形式のコミュニケーションを行い、LLM がユーザーの好みや興味などに合わせて豊かな会話のサポートが可能になることが分かる。

4. 実験

本研究の実験では、提案した会話パターンの中から (1) LLM がユーザーに提案を行う形式を対象として進めた。図 4 はその実験の様子を示している。本実験で評価を目指した項目は以下の 2 点である。

- A) LLM がユーザーの返答に関連した質問や提案を生成できているか
- B) ユーザーと LLM の会話に要する時間の長さ

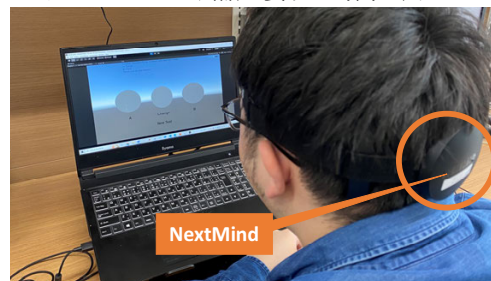


図 4 実験の様子

4.1 実験で使用する機器とソフトウェア

脳波デバイス NextMind をゲームエンジン Unity 上で動かすための GUI インターフェースを用意し、バックエンドとして代

表的な LLM の 1 つである ChatGPT API を使用して Python で動かした。図 5 は、NextMind と ChatGPT の Python-Unity 間での接続を表しており、次に挙げる処理を繰り返して会話を続けていく。

- (i) Python で ChatGPT API を使用し、Prompt に基づいた質問を生成する。
- (ii) Python と Unity 間でソケット通信を行い、生成したテキストを Unity 側に送信する。
- (iii) ユーザーが送信されたテキストを読み、その回答を NextMind を使用して Unity の UI に入力する。
- (iv) ソケット通信を行い、Unity から A もしくは B のテキストデータを Python 側に送信する。
- (v) A もしくは B のテキストラベルに適合する回答を ChatGPT に入力し、再び関連した質問を生成させる。

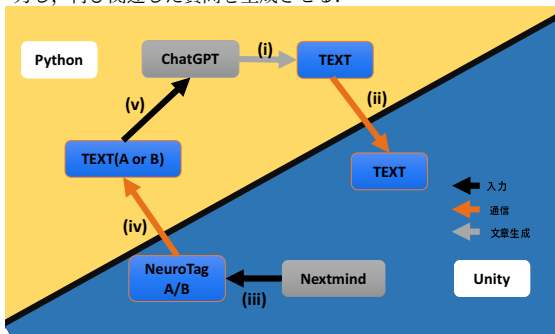


図 5 Python-Unity 間の接続

図 6 は、ユーザーに提示される画面の例である。

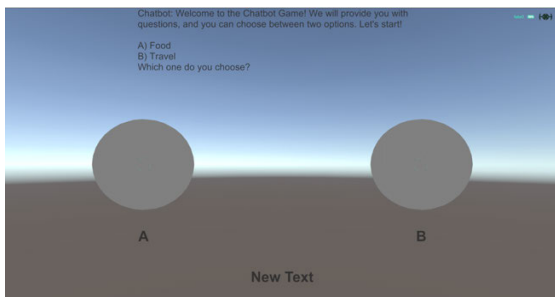


図 6 UI

4.2 実験内容

ここで ChatGPT の Prompt 部分に図 7 の文章を入力し、適切な出力が行われているかどうかを評価する。会話が進んでいく様子を図 8 に示す。初めに、チャットボットがユーザーにトピックを与え、選択したトピックに関連した選択肢を提示する。ユーザーが NextMind を使って選択を繰り返すことで、ユーザーの選択に基づき、その好みや興味に関連する質問が増える。最終的には、これまでに選択されたものに関するまとめと、それに対する提案を行う。

4.3 実験結果

4.3.1 成功例と失敗例

図 9 は、成功出力の 1 つである。図中の矢印で示されている場所はユーザーが NextMind を使用して入力を行った部分である。チャットボットが 2 つの選択肢 A,B を提示し、それに

```
prompt = """
以下のルールに沿ってチャットボットを作りたい。
ルール
1:チャットボットは、まずトピックをA、Bの選択肢として出力する。
2:質問はすべて、AかBで答えられるものでなければならない。
3: ユーザーの返答を待ち、その返答に関連した質問を提供する。
4: これを3回繰り返し、最後に選択された選択肢に応じた提案をする。
5: 出力形式は、次のような形式にしてください。
###
チャットボット:
生成された文章
A) 選択肢A
B) 選択肢B
あなたはどちらを選びますか？(AかBを選んでください)！?
###
"""
```

図 7 チャットボットの Prompt

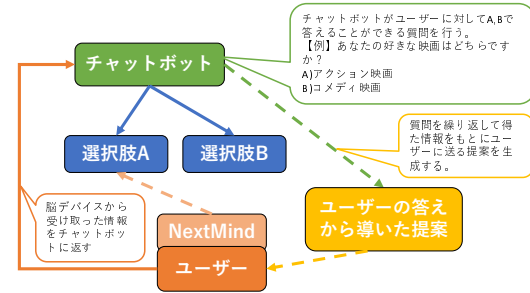


図 8 会話パターン例

ユーザーが応答して会話が続き、という形式が成立している。提案内容も”Imagine Dragons”などの実在するロックバンド名などの固有名詞も出力された。

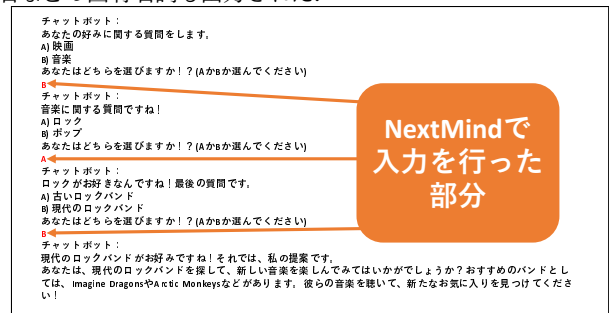


図 9 実験が成功した場合の出力

しかし、同じ Prompt を使用しても会話が進行せず、図 10 のようなメッセージが出力されることもあった。

```
まず、チャットボットのコードを作成します。ルールに従って、対話の流れを制御する関数を定義します。

python
import random

def select_topic():
    topics = [
        {
            "question": "あなたはどちらのタイプですか？",
            "choices": ["A) 山派", "B) 海派"]
        },
        {
            "question": "あなたはどちらの飲み物が好きですか？",
            "choices": ["A) コーヒー", "B) 紅茶"]
        }
    ]
    return random.choice(topics)
```

図 10 実験が失敗した場合の出力

4.3.2 脳波デバイスを介した会話結果の評価

表 1 には、20 回の実験を行い、それぞれの実験で「期待通りまたは期待以上の会話が行われた回数」と「会話が途中で成立しなくなった、または意図しない応答が得られた回数」を記録した結果を示す。

表 1 実験を実行した際の結果

期待通りまたは期待以上の会話が行われた回数	15
会話が途中で成立しなくなった回数	5

この結果は、期待以上の成果を示している。5 回のエラーのうち、4 回は最終提案で会話が成立しなくなったものであり、1 回は最初の質問文が図 10 のようになってしまったものであった。これらの結果から、NextMind を用いた LLM との対話は、全体的に成功と評価できる。

さらに、表 2 に、Unity の画面上にチャットボットが生成した文章が表示されてから、ユーザーが NextMind で入力を完了するまでの時間を、20 回の実験で計測した結果を示す。ChatGPT API の応答時間は、ネットワークの状態やサーバーの混雑状況などにより変動するが、この実験中の ChatGPT の応答時間は約 15~17 秒であった。したがって、1 回の会話交換には約 23~25 秒の時間を要した。

表 2 ユーザーが NextMind で入力するまでの時間

試行回	計測時間 (s)	試行回	計測時間 (s)
1	5.98	11	9.66
2	6.81	12	7.58
3	5.02	13	6.97
4	11.25	14	5.14
5	6.31	15	4.57
6	8.82	16	12.4
7	14.92	17	3.61
8	4.59	18	3.48
9	4.97	19	4.54
10	12.66	20	10.69
平均時間 (s): 7.50			

表 2 からは、文章を読む時間も含まれるとはいえ、ユーザーが NextMind で入力を完了するまでに一定の時間が必要だったことが確認できる。特に、7 回目の計測では 15 秒弱もかかるなど、入力感知にムラが見られた。したがって、現状ではテキスト入力を用いた方法の方が NextMind の入力に比べて、会話の進行がスムーズであると考えられる。

5. 考察

今回の実験を通じて、人と LLM の対話において脳波デバイス NextMind を介在させたコミュニケーションは、選択肢ベースの会話であれば期待通りの会話を行うことが可能であることが確認できた。しかし、選択肢の中に自分の興味のあるトピックや質問が含まれていない場合、回答に困り会話が進まないという状況が発生した。

5.1 改善案

そのため、選択肢を変更可能にするために Unity の UI に新しい NeuroTag を追加し、チャットボットの Prompt のルールに「6:ユーザーが「Change choice.」と入力したら、質問を変えてください。ただし、前の質問と関連性を持つようにしてください。」という指示を追加し、再度実験を行った。図 11 は設定変更後の UI を示す。再実験の結果では、会話の関連性を保ったまま選択肢を変えることが可能になり、より自由度の高いコミュニケーションを実現できた。

しかしながら、これは LLM の会話能力の柔軟性に対して、

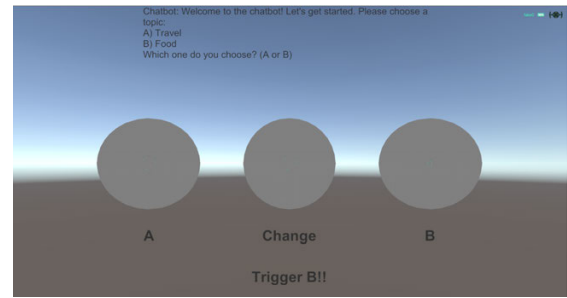


図 11 設定変更後の UI

NextMind の応答パターンが単調であるため、UI を複雑にする必要があったことを示している。つまり、LLM の能力に対して脳波デバイスの機能が追いついていないため、会話パターンのバリエーションが制約されていると言える。今後は、NextMind のように 2 値での応答だけでなく、アナログ値や感情に関連する信号を受け取ることができるなど、デバイスが捉える脳波の種類を増やす、またはアルゴリズムを改善してより精緻な解釈を可能にする等が必要があると考えられる。また、ユーザーが選択肢を直接生成できるようなインターフェースの提供も一つの解決策になるかもしれない。

6. まとめ

本研究では、大規模言語モデルと脳波デバイスを用いた新たなコミュニケーション方法を探求した。その例として、ChatGPT を用いてユーザーの興味に基づく質問を生成する実験を行い、効果的な対話が可能であることを示した。しかし、現段階では、対話は選択肢ベースの形式に限定されているため、その範囲は制限されている。この問題を解消するためには、脳波デバイスの性能向上が不可欠である。

さらに、このアプローチが成功すれば、ユーザーが思い描いたストーリーをゲーム内で創造し、その内容を LLM が改善、拡張させることも期待できる。

謝辞 本研究は、令和 2(2020) 年度科学研究費助成事業 (学術研究助成基金助成金)(若手研究) 課題番号 20K19899 の支援を受けた。

参考文献

- [1] S. A. Dewangga, et al. "Robot Motion Control Using the Emotiv EPOC EEG System," Bull. of Electrical Engineering and Informatics, vol. 7, no. 2, pp. 279-285, 2018.
- [2] F. Lotte, et al. "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update." J. of Neural Engineering, Vol. 15, No. 3, 2018.
- [3] R. Antonello, et al. "Scaling laws for language encoding models in fMRI." arXiv:2305.11863, 2023.
- [4] E. Nichols, et al. "Hey Haru, Let's Be Friends! Using the Tiers of Friendship to Build Rapport through Small Talk with the Tabletop Robot Haru.", 2022 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems (IROS), 2022.
- [5] S. Amershi, et al. "Guidelines for Human-AI Interaction." Proc. of the 2019 CHI Conf. on Human Factors in Computing Systems, CHI '19, pp. 1-13, 2019.
- [6] <https://github.com/Snapchat/NextMind>