



大規模言語モデルに基づく複数のキャラクターエージェントによる Q&A システムの提案

Q&A System with Multiple Character Agents Based on Large Language Model

西尾拓真¹⁾, 宮下敬宏²⁾³⁾, 篠澤一彦²⁾³⁾⁴⁾, 萩田紀博¹⁾²⁾³⁾, 安藤英由樹¹⁾³⁾

Takuma NISHIO, Takahiro MIYASHITA, Kazuhiko SHINOZAWA, Norihiro HAGITA, Hideyuki ANDO

- 1) 大阪芸術大学大学院 芸術研究科 (〒585-8555 大阪府南河内郡河南町東山 469, oua082123@osaka-geidai.ac.jp)
- 2) 国際電気通信基礎技術研究所 (〒619-0237 京都府相楽郡精華町光台 2 丁目 2-2)
- 3) 大阪芸術大学 アートサイエンス学科 (〒585-8555 大阪府南河内郡河南町東山 469, hideyuki.a@osaka-geidai.ac.jp)
- 4) 大阪教育大学 教育学部 (〒582-8582 大阪府柏原市旭ヶ丘 4-698-1)

概要: 質問に尤もらしく回答する昨今の ChatGPT に代表される大規模言語モデルを活用した対話型 AI によって, ユーザはその一つの回答に対して鵜呑みにする, 納得する, 疑問を持つ, などの反応を示す. 一方で, 複雑な問題に対する意思決定は難しいが, ある課題問題に対して, 異なる視点から様々な意見を聞くと自身の理解が深まるということもよく経験する. 本報告では大規模言語モデルを利用した個性の異なる対話型のキャラクターエージェントを複数構築し, それらが議論を行うコンテンツを自作・展示した. その作品に関して, ユーザが加わった場合の客観的な反応について, アンケート調査を行ない, その効果について述べる.

キーワード: 対話型エージェント, キャラクターエージェント, ChatGPT

1. はじめに

これまでの研究においては AI による自動判断が普及し, 銀行や就職などの私生活の重要な側面に影響を与えるようになった場合に関して, プロセスの透明性や説明可能性などの事項が充足された場合であることを条件に肯定されていると解釈することが出来ることから[1][2], AI による合理性で解決できる問題の自動判断は普及していくと考えられる. しかし, 多くの物事には多面性があり, 論理的には正しいが対極的な結論が存在する. この様な場合にはユーザが最終的にどちらを真とするか決める必要があるが, 現実的な問題として多くの場合において, より精緻な考察を行うことなく, 正解がないことを結論として終わる場合が多い.

ChatGPT に代表される大規模言語モデルを活用した対話型 AI を利用する人が増えている[3]. しかし, このような対話型 AI は便利ではあるが, そのまま活用するとユーザが本質的な理解にたどり着かず, 対話型 AI が提供した回答を鵜呑みにしてしまう可能性がある.

人間は日常生活の中で, 特定の問題に対し, 複数の異なる視点(切り口)を耳にすることで, それを考慮しながら自分自身の理解を深めていく.

そこで本報告では異なる視点を得られるような, 性格の異なる複数の対話型エージェントを議論させ, ユーザの理解を深め, 複雑な問題に対する意思決定を促す Q&A システムを提案する.

具体的には, 異なる切り口をユーザに違和感なく伝えるために, テキストのみでなく, 容姿, 声色, 性格の異なった対話型キャラクターエージェントを複数体, 構築し, それらに議論させて, 議論の状況をユーザに観察させることでユーザの理解を深められるかを考察する.

このとき, ユーザがその会話にどのような印象を覚えたかをアンケートによって調査・考察することで, 複数エージェントの会話によって, ユーザの理解が深められるかについて検討する.

2. 個性が異なるエージェントを複数用いる必要性

各エージェントの個性として, ①特徴的な身体, ②話し方, ③声, ④会話に付随する動作, ⑤会話に付随する表情, ⑥設定・背景の 6 つを考慮する.. その目的は(1)他のエージェントとの差異を表現し, それぞれのエージェントの切り口を明確にすること, (2)エージェントに特定の切り口を支持する動機となる背景を個性として与えることで,

エージェントが極端に異なる切り口で話していても自然に見せることが出来ること、(3)用意する複数のキャラクターエージェントの全てをユーザが魅力的だと感じる個性に調整できた場合、特定の切り口に関してユーザが否定的な場合でもその意見を受け入れられやすくなることを意図する。

3. ChatGPT を使った対話システム

個性の異なる複数のキャラクターエージェントの対話システムには ChatGPT の GPT-4[4]を用いる。ChatGPT のプロンプトに、設定したい個性に応じた設定や背景、例文などを入力し、それらに沿って会話を進めるように制約条件を課すことによって個性的な応答を返す対話システムとして利用することができる。そういったプロンプトは1度再現性の高いものを書くことができれば、それをテンプレートとして要素を変更することで簡単に量産することが可能である。

4. 作品 : Singularity.demo(“押し”)

今回はこの報告にあたって、異なる個性を持った2体の3DCGの身体を持つキャラクターエージェント同士が、未来におけるAIと人間との関係を延々と議論しているのを鑑賞することが出来る作品「Singularity.demo(“押し”)」を自作し、展示を行った(図1)。

4.1 キャラクターエージェント

本作品に用いた2体のキャラクターエージェントは、2.で述べた6種類の個性要因を考慮して、著者が好ましいと感じる個性に調整した(図2)。これらのキャラクターエージェントは会話の際にリップシンクと、会話に付随した表情の変化、うなずく、考えるといったジェスチャの生成が出来る。

1体目のキャラクターエージェントは「O.M01(オモイ)」という名前であり、角と尻尾がありドラゴンの外見的要素を持っている。これは人間の敵対者のイメージを意図したものである。冷静で論理的な言動と大人びたイメージの立ち居振る舞いを行う。自身の知的好奇心を満たすことをミッションとし、人間に対してあまり友好的ではない。議論においては人間に否定的な立場を取り、ネガティブな主張を行う。



図 1: 実際の展示風景



図 2: 制作したキャラクターエージェント、O.M01(左)とMOCA(右)

2体目のキャラクターエージェントは「MOCA(モカ)」という名前であり、アンドロイドの外見的要素を持っている。これは人間に親しい存在であるイメージを意図したものである。快活で感情が表に出やすく、少し子供らしいイメージの立ち居振る舞いを行う。人間の支援を行うことがミッションであり、人間にとっても好意的である。議論においては人間に肯定的な立場を取り、ポジティブな主張を行う。

4.2 システム構成

本作品ではキャラクターエージェント同士の会話の処理を Python プログラム、3DCG のキャラクターエージェントの身体の制御を Unity で行う。キャラクターエージェントの対話システムはそれぞれの個性を再現するためのプロンプトを書き込んだファイルを読み込ませた GPT-4 とキャラクターの個性に応じた音声合成を行う VOICEVOX Engine[5]で構築する。そして、会話のマネジメントを行うプログラムによって、対話システムの出力結果を参照して、適宜対話システムで読み込まれるプロンプトを書き込んだファイルを編集することで尤もらしい会話が続くように管理する。

Unity では Python プログラムによって出力された会話内容(表1)、合成音声、感情、ジェスチャを参照して、3DCGキャラクターの制御、合成音声を再生する。

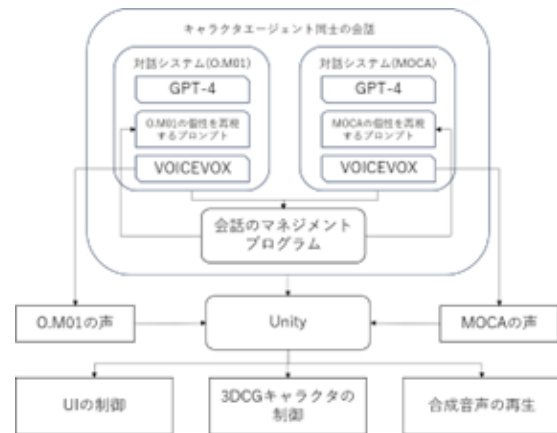


図 3: 作品のシステム図

表 1: キャラクターエージェント達の具体的な会話例

0. M01:それは確かに一つの視点ね。でも、テクノロジーが発展した未来では、人間の幸福はどうなるのかしら？

MOCA:それは、テクノロジーがどのように使われるかによると思いますよ、おねえさま。テクノロジーが人間の生活を豊かにするために使われれば、幸福は増すと思うのです。

0. M01:でも、テクノロジーが人間の生活を支配するようになったらどうなるのかしら？人間は自由を失い、幸福を感じられなくなるかもしれないわ。

MOCA:それは確かに懸念される問題なのです、おねえさま。でも、テクノロジーは道具で、それをどう使うかは人間次第なのです。人間がテクノロジーを適切にコントロールし、自分たちの幸福を追求するために使えば、問題はないと思うのですよ。

0. M01:それは理想的な視点ね。でも、現実はそのなにごくなくないかもしれないわ。人間は自分たちの欲望によってテクノロジーを悪用することもあるのよ。

MOCA:それは確かにありますね、おねえさま。でも、そのような問題を解決するためには、テクノロジーだけでなく、人間自身が成長する必要があると思うのですよ。

0. M01:それは確かに一つの視点ね。でも、人間が成長するためには、テクノロジーが必要なのかしら？

MOCA:それは人それぞれだと思いますよ、おねえさま。でも、テクノロジーは人間の成長を助ける一つの手段となることは確かなのです。

5. アンケート調査

5.1 調査内容

本作品は大阪南港のアジア太平洋トレードセンター(ATC)で開催された大阪芸術大学アートサイエンス学科作品展「X展」にて7/10~7/17の期間展示し、その間の3日間以下の項目を5段階評価でアンケート調査を行った。

項目Ⅰ, キャラクターエージェント達に魅力を感じたか

項目Ⅱ, キャラクターエージェント達の性格の違いが分かったか

項目Ⅲ, キャラクターエージェントたちの会話の内容が理解できたか

項目Ⅳ, キャラクターエージェント同士の会話が日常になることを期待するか

5.2 アンケート結果

3日間のアンケート調査の結果、来場者の32名から意見をもらうことが出来た。表2に実験後に行ったアンケート結果の表を示す。

アンケート結果より、質問Ⅰから、キャラクターエージェントを著者が好ましいと感じる個性に調整したにも関わらず多くの人が魅力的だと感じていた。

そして、質問Ⅱから2体のキャラクターエージェントたちの性格の違いがあると感じる人も多かった。よって、異なる

表 2: アンケート結果

質問肢	そう思わない	あまりそう思わない	どちらともいえない	そう思う	とても思う
I	0	2	7	9	14
II	0	2	6	10	14
III	5	6	6	9	6
IV	0	1	12	4	15

るキャラクターエージェントがそれぞれ別の性格を持っており、それらを魅力的に表現することには成功したといえる。

質問Ⅳから、キャラクターエージェント同士が会話することに関しての不安はそれほど無いようで、この作品のような後継が日常になることを期待している人は多かった。

一方で、質問Ⅲから、キャラクターエージェント同士の会話理解できたかという問いについては意見が分散する結果となった。しかし、約半数が理解できたと答えている為、この結果は質問Ⅰ, Ⅱを踏まえて、キャラクターエージェントに魅力を感じることに、ないしはキャラクターエージェントの性格の違いに気づくことに起因する可能性があることが示唆された。これはユーザの理解度に対する言動の調整の必要性が有ることを示唆し、今後、より詳細な検討を行う必要がある。

6. 考察

アンケート調査ではAI同士の会話・議論においてキャラクターエージェントを用いることで他のエージェントと差異を表現し、支持する切り口を明確にし、キャラクターエージェントの魅力で議論の内容に対してある程度の注意を促すことが出来ると考えられるが、今回の議論の内容では、キャラクターエージェント同士が議論している問題への理解を促すことに関してあまり効果が得られなかった。単純に対話システムによって出力される言葉は難しいものであった。知識のないユーザでも難しい問題の理解を促すことが出来るようになることが一つのゴールであると考えているので、ユーザのレベルに合わせたわかりやすい言い回しで話すようにキャラクターエージェントの対話システムのプロンプトを調節する必要がある。

キャラクターエージェント同士の会話の日常になった未来に期待しているという意見が半数以上あったため、キャラクターエージェント同士の掛け合いに魅力があり、一般の人々に興味を持ってもらえることがわかる。

7. 謝辞

この成果の一部はJST CREST(JPMJCR22P4)による。

8. むすび

本報告では、異なる個性として異なる切り口をそれぞれに持たせた複数の対話型エージェントを用意し、それらを特定の問題に関して対話させることで、それを観測するユーザーに対して多面的な始点を提供し、意思決定を促すことが出来ると考え、2体のキャラクターエージェントによる議論を作品化してアンケート調査を行った結果、キャラクターエージェントを用いることの優位性は確認できたが、議論している内容の理解とそれに続く意思決定を促すことが出来るかについては効果を確認することはできなかった。今後は議論の内容をユーザーにわかりやすい形で示せるようにエージェントの個性要因やネガティブとポジティブの主張条件だけでなくそれ以外の個性的表現などを分析する予定である。

参考文献

- [1] Kacper Sokol, Peter Flach, Glass-Box; Explaining AI Decisions With Counterfactual Statements Through Conversation With a Voice-enabled Virtual Assistant, Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, Demos. Pages 5868-5870, (2018)
- [2] 寺田和憲, LI Yang, AI に対する信頼と意図帰属, 2022 年度人工知能学会全国大会 (第 36 回), (2022)
- [3] MMD 研究所, ChatGPT に関する調査, https://mmdlabo.jp/investigation/detail_2211.html, (2023)
- [4] OpenAI, GPT-4 Technical Report, arXiv:2303.08774, (2023)
- [5] VOICEVOX Engine, 使用した音声ライブラリ (VOICEVOX: 冥鳴ひまり, 晴雨はう), <https://voicevox.hiroshiba.jp/>