

# LLM により議論活性化を目的とした話題提供型 バーチャルエージェントの開発

## Development of Topic-Suggesting Virtual Agent for Enhancing Discussions Using LLM

中村航太<sup>1</sup> 峯岸朋弥<sup>1</sup>

Kota Nakamura<sup>1</sup> and Tomoya Minegishi<sup>1</sup>

<sup>1</sup> 専修大学

<sup>1</sup> Senshu University

**Abstract:** 本研究では、議論の内容を自動的に記録し、適切な話題を提供するバーチャルエージェントの開発を目指す。大規模言語モデル (LLM) の自然言語処理能力を活用することで、議論の流れや参加者の関心に基づいた話題の提案が可能となり、議論の活性化と円滑な進行を支援することを検討する。LLM を用いて話題提供を行うエージェントを用いた実験結果から、参加者の発話文字数に関して分析を行った。開発したエージェントは、ビジネスミーティングなど、多様な場面での議論の質向上に寄与することが期待される。

## 1 序論

どのようなプロジェクトにおいてもグループディスカッション (以下、議論) を経て、物事を決定していくことが現代社会の通例である。しかし議論が全員の納得する形で終わることは稀である。特に、議論において司会役を完璧にこなす事は非常に難しい。この課題を解決するため、自動ファシリテーションエージェントをオンライン議論プラットフォームに用いて、議論の活性化につながることを示した例がある [1]。しかしながら、自動ファシリテーションエージェントによるコメントが無視されてしまうことが多く見られ、司会役が話の流れを作ることができないという問題が解決できていない。

良い議論を達成するためには、司会役は話の流れを作ることと、メンバー全員から意見を引き出す必要がある。しかし、現在のアプローチでは、司会役の知識・能力不足で議論の良し悪しが決まってしまうこと、司会役自身があまり意見を言えずに議論が終わってしまうこと、反対に司会役ばかりが話してしまうことの課題がある。一般的に議論参加者の中で最も多く話す人がグループの決定により強い影響を持つと言われている [2]。

本研究では大規模自然言語モデル (以下、LLM) を搭載し、コンピュータのモニター上で動作する対話型バーチャルエージェントを議論に参加させることを提案する。議論参加者として話題提供を行うこ

とで、参加者の能力に左右されない議論を達成可能にする。

## 2 関連研究

議論の活性化の面において、ファシリテーションを目的としたバーチャルエージェントシステムが、議論テーマに沿った知識の共有や提案を実現している例は少ない。石川ら [3] は議論において共調整を支援するアプローチは 2 つあると述べている。第一のアプローチはロボットによる促し。第二のアプローチは学習者の議論への参加度合いを、データに基づき可視化して提示するものである。前者の例として西村ら [4] は、議論支援エージェントシステムを開発、評価し、発話量に応じて参与をはたらきかけることで参加者の発話の偏りを減らすことが可能だと示した。後者の例として、金沢ら [5] は発話情報をグラフにより可視化し、これを参加者に見せることで議論への均等な参加を促すことが可能であることを示した。これらの研究により、バーチャルエージェントの介入により、議論参加者の発話に偏りのない均等な議論は実現できることが示唆されている一方で、議論テーマに沿った知識の共有や提案が議論活性化に寄与するか研究された例は少ない。中ら [6] は複数のグループが同一の議題に関する議論を同時並行で行う状況を想定し、自身のグループに、他のグループの議論内で重要な単語を提示する機能を備えた議論

支援システム AIR-VAS に LLM を導入することを検討した。この研究ではエージェントに与えた役割に応じた特徴的な発話の生成が可能であることが示唆されており、議論テーマに沿った話題提供を実現するために LLM が有効である可能性がある。

## 4 バーチャルエージェントの提案

バーチャルエージェントの機能として、議論内容を記録する機能、議論内容を要約しその内容から話題提供をする機能、提供された話題を人間のように話す機能を実装する。

### 4.1 議論内容の記録

議論内容に沿った話題提供を行うために、議論内容を記録する必要がある。議論内容の文字起こしは音声認識と文字起こしが可能な機械学習モデルである Whisper を使用し、議論内容を文字起こしする。

### 4.2 話題生成

議論活性化のためにバーチャルエージェントは議論内容に応じた話題提供を行う。文字起こしにより記録する議論内容と”これはグループディスカッションの記録です。これまでの記録内容を要約してください”というプロンプトを LLM に与えることで要約する。要約は箇条書きで要点が出力される。その要約を元に、次に発話する話題を LLM が生成する。話題提供を行うためのプロンプトは議論テーマと提供して欲しい文字数を指定する。実験では”私たちは、

今はもう失われつつある技術を新たに利活用する案を提案するという議論をしています。要約の内容をもとにメンバーに新しい話題を 100 文字以内で提案してください”というプロンプトを使用した。ここで要約を元に話題生成を行うことで議論内容に沿った適切な話題提供が実現可能である。LLM は API により提供される Meta-Llama-3.1-8B-Instruct を使用する。LLM からの返答はテキストファイルとして保存する。これにより、バーチャルエージェントは議論の話題として挙げた内容から、次に議論すべき内容を提示可能になる。

### 4.3 3D モデルの発話

Unity により制御する人と同様の身体を仮想的に持つ 3D モデルが、LLM からの返答を発話する。LLM からの返答が記録されているテキストファイルを Unity 上で読み込み、VOICEVOX で音声ファイルを生成する。生成音声を Unity 上でバーチャルエージェントが Lip Sync で発話することで、人が話しているかのような話題提供を実現する。一連の流れを図 1 に示す。

## 5 対人実験

### 5.1 実験方法

開発した議論の活性化を目的にする話題提供型バーチャルエージェントを評価するため、実際の議論を想定した対人実験を行った。実験は、3人グループによる議論を行い、バーチャルエージェントが LLM を使用して発話内容を決定する条件（以下、LLM あり条件）と、LLM を使用せず、実験者が予め用意した音声ファイルを再生する条件（以下、LLM なし条件）の 2 条件で行い、比較した。具体的な発話内容を表 1 に示す。

実験実施者は、実験参加者へ必ず 2 回以上はバーチャルエージェントによる話題提供機能を使用するように指示する。議論終了後には Heerink ら[6]を参考に、表 2 の 5 つの構成概念からなる質問に回答していただき、技術に対する印象、システムの適応性、使いやすさ、有効性、信頼性を 5 段階（1: 全くそう思わない, 5: とてもそう思う）で評価する。バーチャルエージェントが議論活性化に寄与するかを評価するために議論の録音から発話文字数も計測する。

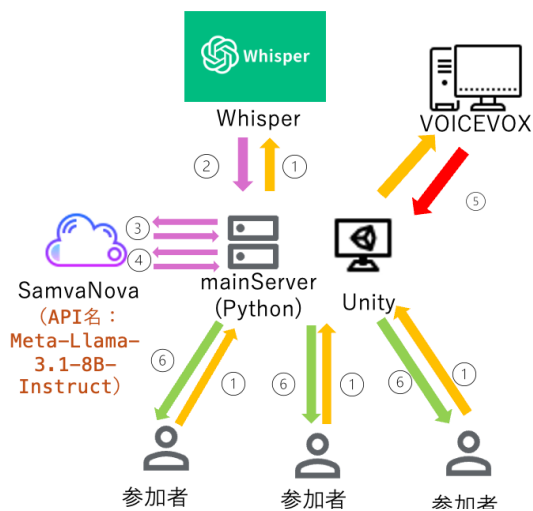


図 1 システム構成図

表 1 実際の発話内容 (議論テーマ: 今はもう失われつつある技術を新たに利活用する案を考えてください)

LLM なし	LLM あり
<ul style="list-style-type: none"> <li>・他の業界や分野ではどう取り組んでいるか誰か知っていますか.</li> <li>・この話をもう少し掘り下げるとどんな問題が見えてくるか考えてみましょう</li> </ul>	<ul style="list-style-type: none"> <li>・私たちは伝統的な分野の技術を新しく利活用する案を考えることになりました. 話し合いの中で和紙の魅力を感じました. 和紙は紙だけではなく紙の質感, 見た目, 触り心地, 光の広がり方などが特徴です. 工房を訪問して和紙の工法を学び, 和紙の暖かみや和紙の質感を感じることを検討してみませんか. 以上の内容について話してみたいかがでしょうか.</li> </ul>

## 5.2 実験空間

実験空間には, バーチャルエージェントを表示するためのモニター, バーチャルエージェントを制御しモニターへ出力するためのPC, 実験者用モニターとキーボード, それらを置くための机を配置した. 机の上にはバーチャルエージェントの音声を流すためのスピーカーを配置し, その周りを囲むように実験参加者に座っていただいた. 実験実施者はタイムキーブと話題提供を行うために, エージェントの近くに座り待機することとした. 机の中央には, 記録用の 360° ビデオカメラを設置した. 実際の実験環境を図 2 に示す.

## 5.3 実験参加者

表 2 アンケート項目

構成概念	代表的な項目
項目 1:技術に対する態度	このエージェントを利用することは良いアイデアだと思う
項目 2:システムの適応能力の知覚	このエージェントは私の必要とすることに応じることができると思う
項目 3:使いやすさの知覚	このエージェントは面白い
項目 4:有効性の知覚	このエージェントは私にとって役に立つと思う
項目 5:技術の信頼性	私はこのエージェントがくれたアドバイスに従うだろう

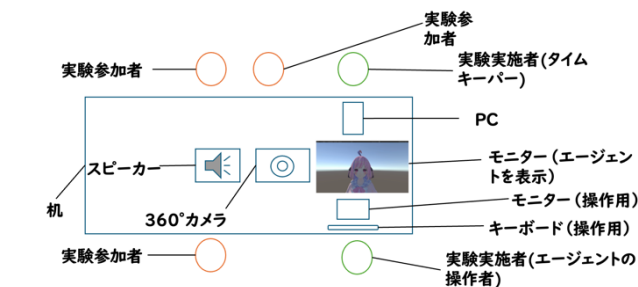


図 2 実験環境

実験参加者は 15 人の大学生であった. 1 組 3 人とし, 計 5 回の実験を行った. ネットワーク情報学部 の 1~4 年生に依頼し, 男性 10 人, 女性 5 人に協力していただいた. 参加者はピンマイクを装着し実験を行った.

## 6 実験結果

アンケートの回答と発話文字数を Wilcoxon の順位と検定を実施し, Bonferroni 法により  $p$  値を補正した. 以下より平均値を Mean, 標準誤差を SE とする. アンケートを分析した結果, すべての項目において有意な差はみられなかった (技術に対する態度: LLM あり条件: Mean = 3.93 SE = 0.08, LLM なし条件: Mean = 3.62, SE = 0.13,  $n.s.$ , システムの適応能力の知覚: LLM あり条件: Mean = 3.27, SE = 0.11, LLM なし条件: Mean = 2.93, SE = 0.14,  $n.s.$ , 使いやすさの知覚: LLM あり条件: Mean = 3.88, SE = 0.10, LLM なし条件: Mean = 3.80, SE = 0.11,  $n.s.$ , 有効性の知覚: LLM あり条件: Mean = 3.53, SE = 0.11, LLM なし条件: Mean = 3.33, SE = 0.12,  $n.s.$ , 技術の信頼性: LLM

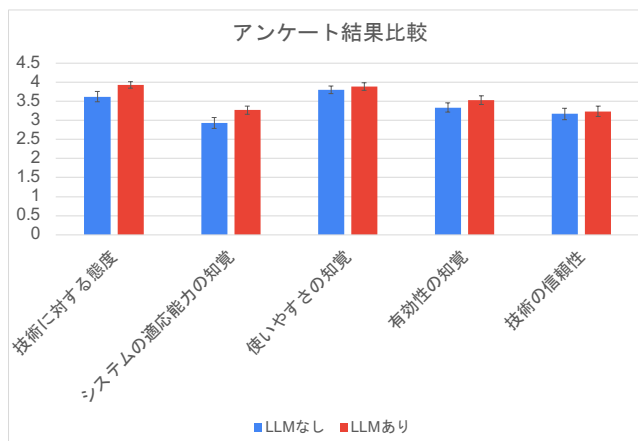


図 3 アンケート分析結果 (エラーバーは Mean±SE を示す)

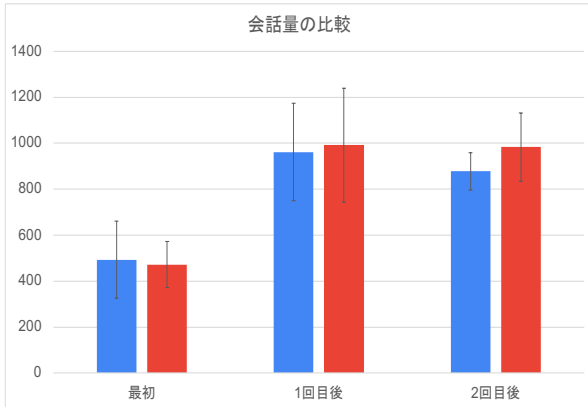


図 4 会話量分析結果 (エラーバーは Mean±SE を示す)

あり条件 : Mean = 3.23, SE = 0.12, LLM なし条件 : Mean = 3.17, SE = 0.14, n.s).

参加者が議論中に発話した文字数を比較した結果、有意な差はみられなかった。文字数比較は参加者が1度目のバーチャルエージェントによる話題提供を行うまでに議論参加者が発話した文字数(以下、最初)と1度目の話題提供を行った後から2度目の話題提供を行うまでの間に議論参加者が発話した文字数(以下、1回目後)、2度目の話題提供から議論終了までに議論参加者が発話した文字数(以下、2回目後)の3項目に分けて分析した。(最初: LLMあり条件: Mean = 472, SE = 100.97, LLMなし条件: Mean = 492.80, SE = 167.38, n.s.1回目後.LLMあり条件: Mean = 991.8, SE = 248.06, LLMなし条件: Mean = 961.8, SE = 211.91, n.s.2回目後.LLMあり条件: Mean = 983.4, SE = 148.49, LLMなし条件: Mean = 877.6, SE = 81.19, n.s.).

## 7 考察

アンケート分析結果により全ての項目において有意な差がみられなかった。参加者同士の会話内容を確認した結果から、思考力があるとみられるグループでは、会話が盛り上がった後に話題提供を行うため、自分たちの考えた案とは違う案が出た際に不快感を覚えるのではないかと考えた。

会話量比較により参加者が議論中に発話した文字数においても有意な差はみられなかった。LLMにより、まとまった案が提案されると、ある程度道筋が立ち、余計な会話が減ったという見方も可能であると考えられる。一方でバーチャルエージェント使用後の発言量は、1回使用後は110.13%で増加傾向だが、2回目使用後は-0.85%で減少傾向であるとわかった。エージェント使用前から使用後の増加率を表3に示す。増加率比較から、バーチャルエージェントを使

表 3 平均増加率

	最初→1回目	1回目→2回目
LLMなし条件	95.17%	-8.75%
LLMあり条件	110.13%	-0.85%

用するタイミングが早い場合は、意見の発散をする段階のため、バーチャルエージェントを使用することにより会話が活発化したのではないかと考えられ、使用するタイミングが遅い場合はまとめの段階なので、話題提供は効果的な役目を果たさなかったと考えられる。

## 8 結論

本研究では議論活性化を目的としたバーチャルエージェントを提案した。実験から、バーチャルエージェントに話題提供を求める前と後では発言量が増加しているタイミングがあったことから、話題提供により議論の活性化に繋がる可能性がある。今後の展望としては、話題提供だけでなく、参加者がエージェントに対して独自の質問をすることができる機能を開発するなどのLLMをより有効的に活用する方法を模索していきたい。

## 参考文献

- [1] 辰巳 智行 中澤 高師 福田 直樹 吉田 寛 青木 美奈 加藤 史也 中島 大晴 山口 七海 : 一般社団法人人工知能学会 2020 自動ファシリテーションがオンライン議論に与える影響の分析
- [2] DiMicco, J. M., Hollenbach, K. J., Pandolfo, A. and Bender, W.: The Impact of Increased Awareness While Faceto-Face, Human-Computer Interaction, Vol. 22, No. 1- 2, pp. 47-96 (online), DOI: 10.1080/07370020701307781 (2007).
- [3] 石川誠彬,江木啓訓,望月俊男,久富彩音,石井 裕,結城菜摘,久保田善彦,加藤 浩 : 日本教育工学会論文誌 2020 協調的議論において共調整を促す対話型ホログラフィックエージェント
- [4] 西村 龍之介,居原田 梨佐,菅本 祐也,石井 裕,望月俊男,江木 啓訓 : 日本教育工学会研究報告集 2023 発話量に応じて参与をはたらきかける議論支援エージェントシステムの評価
- [5] 金澤 拓己,井垣 宏 : 第 10 回実践的 IT 教育シンポジウム (rePiT2024) 論文集 2024 オンライングループディスカッション参加者の発話量可視化手法の検討
- [6] 中 明理沙,吉添 衛,服部 宏充 : 2024 年度人工知能学会全国大会 2024 議論支援システム AIR-VAS への

LLM に基づく議論エージェントの導入とその効果

- [7] M. Heerink, B. Kröse, V. Evers, and B. Wielinga, “Assessing Acceptance of Assistive Social Agent Technology by Older Adults: the Almere Model,” *Int. J. Soc. Robot.*, vol. 2, no. 4, pp. 361–375, Dec. 2010, doi: 10.1007/s12369-010-0068-5.