

校内案内のための音声対話システムに関する研究

Research on a Speech Dialogue System for Campus Guidance

内田 直輝
指導教員 三輪 賢一郎

サレジオ工業高等専門学校 機械電子工学科 情報コミュニケーション研究室

本研究室においては、音声対話による学校案内システムを検討している。本研究では、同システムをさらに詳細な校内案内に対応させるべく、MMDAgent-EX と GPT-2 を用いて独自の対話コーパスでファインチューニングした音声対話システムを構築し、教職員や学生の負担軽減を目指す。

キーワード：音声対話システム, 校内案内, MMDAgent-EX, GPT-2

1. はじめに

最近では、対話型の AI サービスの ChatGPT や、llama を搭載している Meta AI などの、AI 搭載の音声対話が目覚ましい進化を遂げている。中でも GPT は API により様々なサービスが作られ、利用者の生活をサポートしている。

音声対話システムの応用例にキャンパス案内システムがある。例えば奈良先端科学技術大では、ロボットを用いた音声対話システムが検討されており、ジェスチャーや視線をカメラによる画像処理で認識することが可能である。

本校においては、学校案内システムの検討はされているが、研究室や各部屋のより詳細な案内システムは検討されていない。外部の方々の来校時や生徒が学校内を移動する際に、教職員や教室が分からず迷ってしまうことがある。案内する人がいなくても、的確に自分の行きたい場所がわかるような説明が受けられる案内システムを導入するのが望ましいが、掲示板に教職員や教室などの情報をすべて記載することは難しい。しかし、前述の先行研究のようなロボットを製作しようとすると、多大なコストを要する。そこで本研究では、パソコンレベルで実装可能な、サレジオ高専のより詳細な校内案内に特化した AI 搭載音声対話システムを製作し、教職員や学生等の負担を軽減するとともに、誰でも気軽にかつ楽しく利用することができる校内案内システムを構築することを目的とする。

2. 方法

本研究では対話システムのフレームワークとして一般的な、名古屋工業大学で開発の MMDAgent-EX[2]を用いることにした。MMDAgent-EX には、音声認識エンジンである Julius、音声合成ソフトである OpenJTalk が組み込まれており、インストールした時点で最低限の対話が行えるようになっている。本研究ではこれに rinna 社より公開されている GPT-2[3]を搭載し、独自で作成した対話コーパスでファインチューニングを行う。コーパスには日常会話や挨拶などの他に、校内案内システムのための本校教職員、各部屋、教科の情報を紐づけて学習させる。表 1 に関連付けした内容を一部抜粋して示す。また、図 1 にシステムの簡略図を示す。

表 1 対話パターンの一部

教職員	部屋	教科
三輪先生	314	アンテナ工学
吉田先生	303	電気回路 2
米盛先生	304	アナログ電子回路

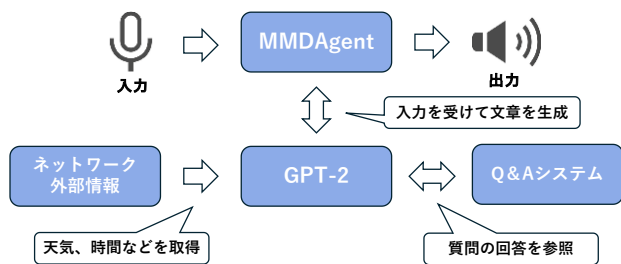


図1 システムの簡略図

3. 評価

本システムの性能評価のための検証実験を、男子学生 8 名、男性教員 1 名を被験者として実施する。具体的には、被験者にはあらかじめ決めた質問を読み上げてもらう形で実施する。この時の音声認識の成功率とハルシネーションの発生確率を(1)式と(2)式で算出して評価を行う。

$$\text{認識成功率} = \frac{\text{認識成功回数}}{\text{対話回数}} \times 100 \dots (1)$$

ハルシネーション発生確率

$$= \frac{\text{ハルシネーション発生回数}}{\text{認識成功回数}} \times 100 \dots (2)$$

ただし、質問文の言い回しは被験者が自由に決めてよいものとし、認識に失敗した場合は一度だけ言い直しを許容して 2 回目に正しく認識すれば成功とみなすこととする。また、被験者に応答の自然さ、ユーザー満足度、会話の多様性、応答速度、内容の一貫性の五つの観点から五段階で評価を行ってもらう。表 2 に五つの観点の詳細を示す。

表 2 五つの観点の詳細

項目	内容
応答の自然さ	不必要に長い応答や、不自然な言い回しがないか
ユーザー満足度	質問への回答が的確で、ユーザーの問題解決に役立つか
会話の多様性	同じ意味合いでも異なる言い回しや表現が使われているか
応答速度	会話のテンポが保たれる程度に、遅延が少ないか
内容の一貫性	一貫した立場や情報を保ち、応答に矛盾がないか

4. まとめ

本研究では MMDAgent-EX に GPT-2 を搭載し、独自で作成した対話コーパスでファインチューニングを行うことで、本校の詳細な校内案内に対応可能な校内案内システムを構築した。

5. 今後の展望

今後は、本システムの評価実験を実施し、その有効性を確認する予定である。

謝辞

本研究には、名古屋工業大学の研究成果物[2]及び rinna 社配布物[3]を利用しております。

文献

- [1] 西村竜一, 内田賢志, 李晃伸, 猿渡洋, 鹿野清宏, “Julius を用いた学内案内ロボット用音声対話システムの作成”, 情報処理学会研究報告, 2001-SLP-39-16, (2001-12)
- [2] Lee, A. (2023). MMDAgent-EX (Version 1.0.0) [Computer software].
- [3] Zhao, T., & Sawada, K. (n.d.). rinna/japanese-gpt2-medium. Hugging Face.