

複数人ユーザの協力的意思決定会話を支援する 情報提供エージェント

Toward an Information Providing Agent who Supports the
Collaborative Decision-making Conversation of Multiple Users

黄 宏軒^{*1, *2} 武田 信也^{*3} 小野 正貴^{*3} 中野 有紀子^{*2}
Hung-Hsuan HUANG Shinya TAKEDA Masaki ONO Yukiko NAKANO

^{*1}立命館大学情報理工学部情報コミュニケーション学科
Department of Information & Communication Science, Ritsumeikan University

^{*2}成蹊大学理工学部情報科学科
Department of Computer and Information Science, Seikei University

^{*3}東京農工大学大学院情報工学専攻
Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology

For an information providing embodied conversational agent (ECA), the situation when there are more than one users is expected. In order to make the agent join the conversation with users, it is necessary to identify the addressee and understand inter-user interactions. This paper presents an ongoing work in trying to estimate the context of the decision making conversation between two customers and an information providing agent. A Wizard-of-Oz (WOZ) experiment is conducted to gather the interaction corpus of human-agent (controlled by a human). The labeled corpus will then be used to develop a model for detecting the addressee and estimate the context of inter-user conversations with machine learning techniques. A prototype agent is being developed utilizing the information to actively intervene the customer's discussion and provide them useful information.

1. はじめに

会話エージェントは、ロボットの物理的制限を無くし、表情の表現や身体の動きに自由度の高いCG描画による擬人化キャラクターを用いることで、ユーザとフェース・ツー・フェースの会話において、リアルさと高度なコミュニケーション能力が追求されてきた。会話エージェントは、一般的なユーザが特に操作の訓練を受けなくても、日常生活で使い慣れた人間同士のコミュニケーション手段で操作できるため、次世代のコンピューターインターフェースのひとつの候補として注目され、世界各国で盛んに研究されるようになった。研究の方向もリアルなキャラクターアニメーション描画やコミュニケーション能力といった基礎研究から、教育訓練や公開展示など、実世界での応用方向へ移行しつつある。

一般来訪者が訪れる展示会や博物館での案内役などに会話エージェントの応用が期待されるが、来訪者グループの複数人が同時にエージェントにインタラクションしようとする場面はしばしば起きると思われるため、複数人のユーザとの会話機能をもつエージェントが望ましい。一対一の会話に比べ、会話参加者が3人以上いた場合、一回一回の発話で、各参加者を、話し手、聞き手、傍参加者の役割に区別する必要性が生じてくる。例えば、エージェントが発話するときにも、傾聴するときにも、その注意を向ける相手をキャラクターの動作として表現する。発話権の受け渡し先の参加者、適切な取得・譲渡方法、タイミングを選択する必要がある。スピードが速く、自由度の高い、エージェント側からは理解しにくいユーザ同士の会話も生じうる。また、これまでの会話エージェント研究では、受動的にユーザからのリクエストに応える、機械的なものが多かったが、エージェントが積極的に会話に参加することによ

り、生命らしさを演出できると期待される。

明示的な言語行動以外に、うなずき、視線方向、身体の向きといった非言語行動は、発話ターンの交替と会話参加者の役割特定に大きく寄与することが、心理学、言語学の分野で前から指摘されている[榎本 03]。本研究は、ユーザの非言語行動を併用することで複数人ユーザとの積極的な会話の実現を目指すものである。そのためには、2つの重要な課題が伴うと考えられる。まず、エージェントがユーザの発話対象を区別し、自分に対する発話に対して的確に回答すること。次に、完全に理解できなくても、ユーザ間の会話の話題を捉え、介入の手がかりとすること。この2つの推定を行うための材料を得るには、情報提供エージェントを題材とし、人間が操作する情報提供エージェントと2人のユーザの間での会話を収録するWizard-of-Oz(WOZ)実験を行った。現在、その分析と予備実験で得られた知見に基づくプロトタイプの開発を進めている。

2. 先行研究

複数ユーザに対応するには、システムの複雑度が格段に高くなってしまったため、これまでの会話エージェント研究の殆どは、一体のエージェント対一人のユーザ、あるいは、複数のエージェント対一人のユーザという比較的シンプルな場面しか想定していない。数少ない複数人ユーザを対象とした研究でも、会話の自由度が限定的なものであり、エージェントを複数ユーザに対応させる研究はまだ十分なされていない。例えば、Rehmらによる3人カードゲームシステムGamble[Rehm 08]では、ラウンドごとに交代するユーザが決まるため、ターンの譲渡先を特定する必要がない。NASAでは、2人の宇宙飛行士が会話しながら音声でロボットを制御するために、飛行士の発話相手を判定する方法を研究[Dowding 06]しているが、ロボットに対する発話はコマンドに限られている。マイクロソフトリサーチでは、複数の来訪者に対応する受付エージェントが

連絡先: 黄 宏軒, 〒 525-8577 滋賀県草津市野路東 1-1-1
立命館大学情報コミュニケーション学科
E-mail: huang@fc.ritsumeai.ac.jp

開発されているが、来訪者の会話参与意思の有無を判定する程度にとどまっている [Bohus 09]。ユーザの発話に対しては、公開展示場などでは困難とされる音声認識と言語理解で全て捉えようとしている。

3. 積極的な情報提供エージェント

会話エージェントの典型的な利用例として、本研究では情報提供エージェントを取り上げた。ここでの情報提供エージェントとは、以下のような会話の場を携わる会話エージェントのことを指す。(1) ユーザは意思決定(選択)のために、該当ドメインのエキスパートであるエージェントに判断材料となる情報を求める。(2) ユーザは、意思決定のタスクの大まかな目標を認識しているが、前もってどんな候補があるか、個別の候補の情報を持っていない。(3) 全てのユーザが自分たちで納得できる決定をした時点で会話の目標が達成できたとみなす。

さらに、積極的に会話に介入する情報提供エージェントは、以下のような行動をする。(1) ユーザの意思決定を助けるために適切なタイミングに適切な情報を提供する。例えば、ユーザの議論が進まない場合、候補としてまだ挙がっていないが、ユーザのニーズに合った別のものを推薦する。既に議論されている候補の補足情報を提供する。(2) 議論のまとめを助ける。例えば、議論が発散していればユーザに注意する。候補の数が減り、結論が早く出るように、候補の問題点を提示する。

4. 会話データ収集実験

積極的な情報提供エージェントの実現にあたって、以下の質問の答えを見つけるために、WOZ 実験を行った。(1) 発話者が意図した聞き手を判定する手がかりがあるか。(2) ユーザが議論すべき課題について相談しているか、議論が発散して雑談になっているかを判定する手がかりがあるか。(3) 情報提供エージェントがユーザを支援するために、どのようなタイミングで、どのようにユーザの会話に介入すべきか。

1人の情報提供エージェントと2人のユーザ間の会話に対し、課題のドメインによってユーザ行動の偏りがないかを検証するために、旅行計画と単位履修登録という2つの課題を設定した。旅行計画の課題では、京都での観光を題材とし、洛東、洛西、洛北、洛中の4つのエリアに分けて14箇所の観光スポットの歴史、見所、グルメ情報、お土産についてのコンテンツを用意した。単位履修登録の課題では、成蹊大学の2回生向け一般教養科目の情報系分野、理工学の技法、言語とコミュニケーション、人間と社会の4つの分類に分けて12科目の授業時間、成績評価方法、担当教員、参考書、授業方法、学生評判、予備知識についてのコンテンツを用意した。課題と関係ないが、起きると想定された会話やエージェントキャラクターそのものに対する質問のコンテンツも用意した。

成蹊大学で2人の同性ペアを条件に、募集した学部生、大学院生21ペア42名(男性28名、女性14名)がユーザとして実験に参加してもらった。パソコンの操作に習熟している必要があるため、これとは別に、情報科学科の学部生4名が一時間ほど練習したあとに、エージェントの操作に当たってもらった。エージェントの操作インターフェースは画面に表示された事前に用意されたコンテンツのメニューをクリックすることで、エージェントがそのまま発話するという簡易なインターフェースであった。また、必要に応じて、キーボード入力のインターフェースも用意した。

ユーザ役の実験参加者は、旅行計画と単位履修登録の2種類のコンテンツにおいて、人間が操作するエージェントから情

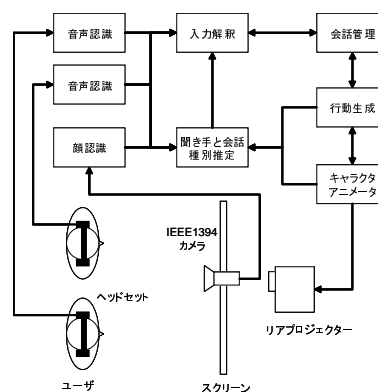


図 1: 提案システムのアーキテクチャー

報を得ながら相談して、一緒に旅をしたい3つの観光スポットと、一緒に履修したい3つの科目を決定するという課題であった。ユーザの慣れによる偏りがないように、ペアごとに課題の順序を交替させた。ユーザ間の交渉を促すために、単位履修の課題では、1人に半日ずつ、都合の悪い時間帯を設けた。また、熱心な議論を促すために、全参加者で統計して、最も人気のある決定をしたペアに報酬が2倍になると教示した。なお、エージェント操作者が隠れており、ユーザに会話しているエージェントが自律システムであると教示した。議論している候補を明確するために、エージェントに覚えてもらう仮登録とその取り消しのルールを導入している。エージェント操作の実験参加者に対しては、第3節のエージェントの定義を説明した上で、積極的にユーザの議論を支援し、会話が始まってから10分程度で議論を収束させるように教示した。

5. 今後の予定

今回の実験では、ビデオデータ、システムのログのほか、マイクの入力で発話区間と顔認識ソフトを用いてユーザ役参加者の顔の動きを記録した。今後は、データの分析を進め、聞き手と会話種別(相談中、雑談中)の推定モデルを構築し、複数のモダリティが補い合うことで、推定精度を高める予定である。さらに、このモデルを活用して、図1で示した自律システムの実現を目指す。

参考文献

- [Bohus 09] Bohus, D. and Horvitz, E.: Open-World Dialog: Challenges, Directions, and Prototype, in *Proceedings of IJCAI Workshop on Knowledge and Reasoning in Practical Dialog Systems*, Pasadena, CA (2009)
- [Dowding 06] Dowding, J., Alena, R., Clancey, W. J., Sierhuis, M., and Graham, J.: Are You Talking To Me? Dialogue Systems Supporting Mixed Teams of Humans and Robots, in *AAAI Fall Symposium* (2006)
- [Rehm 08] Rehm, M.: "She is just stupid"-Analyzing user-agent interactions in emotional game situations, *Interacting with Computers*, Vol. 20, No. 3, pp. 311-325 (2008)
- [榎本 03] 榎本 美香, 伝 康晴: 3人会話における参与役割の交替に関わる非言語的行動の分析, *人工知能学会研究会資料 SIG-SLUD-A301*, pp. 25-30 (2003)