

会話エージェントによる優位性推定に基づくグループ会話への介入 Intervening in Multiparty Conversations by Conversational Agents based on Dominance Estimation

吉野 堯*¹
Takashi Yoshino

八城 美里*¹
Misato Yatsushiro

高瀬 裕*²
Yutaka Takase

中野 有紀子*²
Yukiko Nakano

*¹ 成蹊大学大学院理工学研究科 Seikei University
*² 成蹊大学理工学部 Seikei University

With a goal of contributing to multiparty conversation management, this paper proposes a conversational robot that can facilitate multiparty conversations between three human users. First, we implemented a dominance estimation mechanism using speech signal and head gaze direction recognition. Then, we proposed a method that decides when, to whom, and about what the system should talk to facilitate the conversations based on the dominance level of each user. By implementing the proposed method, we created a communication robot that can provide proper gaze behaviors and facilitate multiparty conversations as a travel guide.

1. はじめに

会話エージェントは、人間と同様の身体的表現を持ち、言語・非言語情報を駆使して人間とコミュニケーションを行う。従来の情報提供型の会話エージェントは、ユーザからの質問に答えることが主な会話機能であった。しかし、グループユーザに対応する会話エージェントの研究では、より積極的にユーザ間の会話に介入し、多人数会話をファシリテートする機能を実現することが新しい課題の1つとなっている。多人数会話に介入する際、重要となるのは、介入対象、介入タイミング、そして介入の内容である。本研究では、会話参加者の顔向きや発話量から会話における優位性(会話を牽引する度合い)を逐次的に推定し、この指標や会話内容から介入対象と介入タイミング、介入内容を決定するシステムを開発した。

2. 優位性決定機構

Nakano et al.[1]は、会話参加者当事者による会話における優位性の判断が、グループ内会話において意思決定を行う上で重要となる情報提供・情報要求行動を行う発話の頻度と強い相関を持つことを報告している。Yatsushiro et al. [2]では、この結果に基づき、情報提供・情報要求の発話頻度を優位性の指標として、会話参加者の言語・非言語情報から優位性を推定する重回帰モデルを提案している。会話を牽引している参加者ほど優位性推定値は高くなり、この値が低い参加者ほど、会話の参加率が低くなる傾向がある。上記重回帰モデルを実装した優位性推定機構の構成図を図1に示す。本機構は、各参加者の音声と顔向きの情報から優位性推定値を算出し、発話ごとにこの推定値を更新する。

3. 介入内容の決定方法

本研究が提案する会話介入システムの介入対象、介入タイミング、介入内容決定手法について述べる。

3.1 介入対象について

多人数対話のファシリテーションを行うエージェントは会話エージェントにも会話に介入する際に特定の意図を持たせ、それに応じて介入対象を決定することが望ましい。本研究では、会

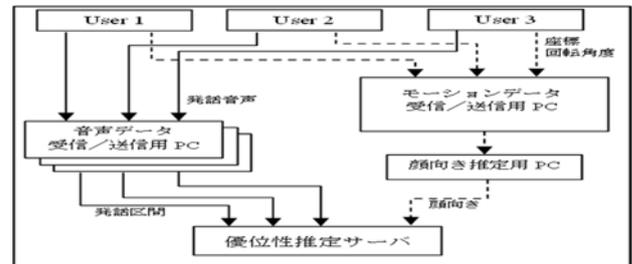


図1. 優位性推定機構

話介入によって期待する効果を3つ定め、それらを達成するために適していると考えられる介入対象を各参加者の優位性から決定する。以下に3つの介入によって期待する効果と、介入対象者、その理由を挙げる。

1) 期待する効果: 目的の早期解決

介入対象: 優位性が最も高い会話参加者

優位性の最も高い会話参加者は意見の重みが他者に比べ重いと考えられる。この対象へ会話介入を行うと、会話が促され、目的の早期解決に繋がると考えた

2) 期待する効果: 介入対象の会話への参加チャンスの増加

介入対象: 優位性が最も低い会話参加者

優位性の最も低い会話参加者は、会話に参加できている度合いが低いので、会話介入によって会話への参加を促せられると考えた

3) 期待する効果: グループ全体の会話参加度合の均一化

介入対象: 優位性が最も高い会話参加者以外

優位性の高い参加者以外に会話のチャンスが生まれ、グループ全体でバランスの取れた会話が行われるのではないかと考えた。

上記の3種類のいずれかを目的とするかは事前に決められ、会話途中での変更は行わない。しかし、介入対象は介入タイミングの都度各参加者の優位性推定値から決定される。

3.2 会話介入タイミングについて

乙木ら[3]は、複数人ユーザと会話エージェントによる会話場面において、エージェントがより適切にユーザ同士の会話に介入することができるタイミングの推定を目的に、コーパス収集 WOZ 実験を行った。発話状況や非言語情報の分析を行い、5種類の会話への介入タイミングを提案して

いる。本研究では、この5種類から、「Stagnation：グループでの会話が停滞している状態」と「Argument：タスクの内容を議論している状態」に着目し、これらのタイミングを検出し、会話介入を行う。

具体的なシステム実装では、会話中に3秒以上の沈黙があれば Stagnation の状態と判断し、1分以内に同じ話題が3回以上現れた場合、Argument の状態とした。また、上記2種類の介入タイミングではあるが、会話の流れを考慮すると、これらの介入が不適切であると判断される場面の為に例外的な介入タイミングとして Other を設けた。例えば、Argument 状態ではあるが、介入対象が現在話題になっている会話に参加できていない場合などに適用される。

3.3 介入内容について

大原本[4]は、会議ファシリテーションに着目し、ファシリテータの行動や発言についての分析を行い、ファシリテータが会話に介入する際に持つ意図を、①「議論を発散させる」②「議論を収束させる」③「意図をわかりやすく具体化する」という3種類に分類した。本研究では、これらのうち、①「議論を発散させる」介入内容に着目する。具体的には、「〇〇さんはどう思いますか？」のようにそれぞれに意見を求めるような介入がこれに当たる。

議論を発散させる介入内容として、以下3つの介入内容の候補群を定めた。

- a) 特定のユーザに意見を求める
例. あなたはどう思いますか?
- b) 他の方向性を提案する
例. 明治神宮はどうでしょうか?
- c) 現在の話題を補強するような意見を求める
例. 明治神宮の何に魅力を感じましたか?

これらは、会話の状況に応じて選択される。また、3.1節で述べた3種類の介入によって期待する効果それぞれが上記3つの介入内容候補群を持つものとし、計9種類の介入内容を用意した。例として、2)「他の方向性を提案する」場合に、期待する効果毎の介入内容は次の通りである。

- 1') 期待する効果: 目的の早期解決
例. (現在の話題とは別の話題)についてはどう思いますか?
- 2') 期待する効果: 介入対象が会話へ参加するチャンスの増加
例. (現在の話題とは別の話題)に行きたいと思いませんか?
- 3') 期待する効果: グループ全体の会話参加度合の均一化
例. (現在の話題とは別の話題)などいかがでしょうか?

3.4 会話介入タイミングと介入内容の関係

最後に、介入タイミングと介入内容の関係について述べる。介入タイミングは、3.2節で述べた3種類があり、これに前節で述べた3種類の介入内容候補群を組み合わせる。

・介入タイミングが Argument の場合

特定の話題で会話が議論状態になっているときには、3.3節で述べた介入内容候補群のc)を使用する。「議論」が行われている場面で、会話参加者に「情報の補強を促す」ことでより具体的な検討を行う手助けになると考えた。

・介入タイミングが Stagnation の場合

議論が停滞状態であると考えられるときには、介入内容候補群のb)を使用する。会話が「停滞」している状態であるため、「他の方向性を提案」することで、議論を再開する手助けになると考えた。

・介入タイミングが Other の場合

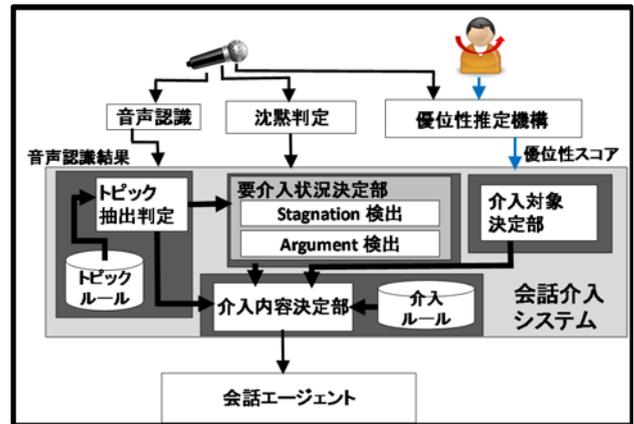


図 2. システムアーキテクチャ

例外的な状況であると判断された場合には介入内容候補群の a) を使用する。「特定の会話参加者へ意見を求める」行動は、状況によらず受け入れられると考えた。

3.3節で述べた通り、介入内容は、介入内容候補群と、介入によって期待する効果の組み合わせになるため、計9通りの介入内容となる。さらに、会話内容に多様性を与えるため、それぞれの介入内容に3種類の会話パターンを設けた。従って、会話内容には27パターンが存在する。

4. システム実装

ここまでで述べた会話介入手法を組み込んだ会話介入システムを構築する。また本研究では、多人数会話を円滑に行うため会話エージェントに優位性と参与役割を考慮した注視行動モデルを実装した。これにより会話エージェントは自然な顔向き動作を実現した。

4.1 優位性推定機構

Yatsushiro et al. [2]の優位性推定機構では、会話参加者の顔向き情報と発話区間を得ることで優位性をリアルタイムで計測していたが、これまでモーションキャプチャを用いていた、顔向き推定のためのモーションデータ取得を、簡便さと十分な精度が得られる、Microsoft社のKinectセンサを利用するように変更した。

4.2 会話介入システム

3章で述べた介入内容決定方法を組み込み、リアルタイムで多人数会話に介入できる会話介入システムを構築する。本システムは複数の候補地から行き先を決定するというドメインで多人数会話が行われることを想定している。提案するシステムの構成を図2に示す。各モジュールについて以降で詳細を述べる。

トピック抽出判定部: 会話における現在のトピックを認識するために、Google音声認識を利用する。今回は行き先候補地名をトピックとした。会話中に予め用意された候補地名が得られた場合、会話の現在のトピックをその候補地名とし、得られた時間や、話者名などを保持する。これらは介入内容を決定する際に用いる。

要介入状況決定部: ここでは、介入タイミングの判定を行う。先に述べたように、グループ間で3秒以上の沈黙を観測すると、Stagnation状態、特定の候補地名が1分以内に3回以上得られた場合、Argument状態とする。これらのタイミングが判定された場合、介入対象決定部へと処理が進む。なお、例外である

表 2. 介入内容候補群選択

介入タイミング	例外判定条件	介入内容候補群
Argument	該当トピックについて介入対象が30秒以内に話している	現在の話題を補強するような意見を求める
	該当トピックについて介入対象が30秒以内に話していない	特定のユーザに意見を求める
Stagnation	1分以内にトピックが5回以上入れ替わっている	特定のユーザに意見を求める
	1分以内にトピックが5回以上入れ替わっていない	他の方向性を提案する

Other 状態の判断は、介入内容を生成する際に行う。

介入対象決定部：3.1 節で述べた介入によって期待する効果と、現在の各参加者の優位性の値から介入対象者を選択する。

介入内容決定部：介入タイミングが観測されると、介入内容決定部が呼び出される。ここまですでに得られた介入タイミング、介入対象、会話トピックに応じて、3章で定めた介入内容決定手法に基づき、介入内容を決定する。具体的な介入内容の例を表2に示す。例外条件に該当した場合には、介入タイミングはOtherとなり、介入内容候補群も変化する。これら介入内容候補群に介入によって期待する効果から決まる介入対象も加味して、介入内容を決定する。

4.3 注視行動制御システム

注視行動は多人数会話を円滑に行うための重要なコミュニケーションシグナルであるため、本研究では、Yoshino et al [5]で提案されている優位性と参与役割を考慮した注視行動モデルを実装した。参与役割とは、発話を行っている「発話者」、その発話が向けられている「受話者」、会話には参加しているが発話者でも受話者でもない「傍参与者」から構成される。図3に注視行動モデルを実装したシステム構成図を示す。

4.3.1 入力処理

Kinect から得られるフェイストラッキング結果とマイクロフォンアレイから得られる音声データを基に、発話者と各参加者の顔向き推定を行う。顔向き認識部はKinectから各実験参加者の顔の回転角度 yaw が入力され、microcone は全方向の音源情報を取得することが出来る。これを用いて発話者認識部ではその音源方向にいる参加者を発話者と推定する。

4.3.2 参与役割推定

発話者と各参加者の顔向き情報から参与役割を推定する。発話者は先ほどの発話者認識部で推定されている。受話者推定では発話者推定直後にその発話者が視線を向けている参加者を受話者とする。そして、発話者と受話者のどちらでもない参加者を傍参与者と推定し、各参加者の参与役割を明確にする。

4.3.3 優位性推定

発話者と各参加者の顔向き情報から各参加者の優位性を計算する。優位性の推定方法として、4.1 で述べた優位性推定機構を使用した。

4.3.4 顔向き生成

注視行動モデルを用いて、会話エージェント次の注視行動を出力する。そのために各参加者の優位性と参与役割の情報から注視行動モデルを用いて次の注視対象を計算し、ロボット動作制御部に送る。最後に、顔向き生成から入力された次の注視

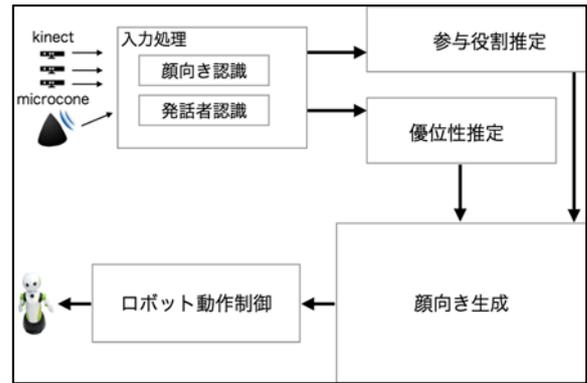


図 3. 注視行動モデルのシステム構成図

対象情報に基づき会話エージェントに顔向きの命令を送信し、ロボットの注視行動が制御される。

5. おわりに

本研究では、多人数会話の介入できるエージェントをめざし、優位性推定モデルを元に会話参加者の優位性を推定する優位性推定機構を実装し、介入対象の決定手法、介入タイミングの決定手法、介入内容の決定手法を提案し、これらを実装した会話介入システムを作成した。最後に、優位性推定機構、会話介入決定システムに加え、ロボットの顔向き制御システムを作成し、リアルタイムで会話に介入することのできる会話介入システムを構築し実装した。今後は、作成したシステムの評価実験を行い、ロボットによる会話への介入の影響や有用性について評価を行う予定である。

謝辞：本研究は、科学技術振興機構(JST)戦略的想像研究推進事業(CREST)「実践知能アプリケーション構築フレームワーク PRINTEPS の開発と社会実践」の支援によって実施した。

参考文献

- [1] Nakano, Y. I. and Fukuhara, Y. Estimating Conversational Dominance in Multiparty Interaction. in 14th ACM International Conference on Multimodal Interaction (ICMI2012). pp. 77-84. 2012.
- [2] Yatsushiro, M., Ikeda, N., Hayashi, Y., and Nakano, Y. I. : A Dominance Estimation Mechanism using Speech and Head Motion Data, 6th Workshop on Eye Gaze in Intelligent Human Machine Interaction (GazeIn'13), 2013.
- [3] 乙木翔地, 堀田怜, 黄宏軒, 馬場直哉, 中野有紀子, 川越恭二: 複数人ユーザ会話におけるエージェントの割り込みタイミングの推定手法の検討, HAI シンポジウム(2012).
- [4] 大本義正, 戸田泰史, 上田一博, 西田豊明: 議論への参加態度と非言語情報に基づくファシリテーションの分析, 情報処理学会論文誌 52(12), 3659-3670(2011)
- [5] Yoshino, T., Takase, Y., and Nakano, Y. I. : Controlling Robot Gaze according to Participation Roles and Dominance in Multiparty Conversations, in 10th ACM/IEEE International Conference on Human-Robot Interaction(HRI2015), 2015