

発話交替シミュレーションシステム ARABAHIKA

- 人間観察に基づく自律的発話交替エージェントの提案 -

湯 浅 将 英[†] 徳 永 弘 子[†] 武 川 直 樹[†]

ARABAHIKA: Autonomous Turn-Taking Agent System based on Bayesian Model

MASAHIDE YUASA,[†] HIROKO TOKUNAGA[†] and NAOKI MUKAWA[†]

1. はじめに

複数の人が会話しているときに、人の視線の動きを視線検出装置などで測定し、発話している人の方向を見たり、適切なタイミングで会話を開始したりするエージェントやロボットが開発されている^{1),2)}。

しかし、これらのエージェントやロボットは、視線の制御にはシンプルなルールを用いているのみで、その動作も単純なものだけに留まっている。人同士の会話では、聞き手は話し手を見る、話し手に見られた人が話し出す、といった単純なルールだけではない。たとえば、話しかけた人と異なる人が話し出す場合や、話し手にたとえ見られていなくても割り込んで話し出す場合がある。また、従来までの研究では言語の内容から話し手と聞き手を推測したり、対象ユーザは一人であったりと、複数人の会話における視線を考慮したものは少ない。

そこで本研究では、実際の複数人の発話交替時における視線の動きを観察し、その結果から視線行動の傾向を算出し、ベイジアンネットワークによる発話交替時の視線モデルを作成した。さらに、このモデルを実装したエージェントを複数作成し、視線によって発話交替をするシミュレーションシステムを作成した。視線入力装置を用いることで、ユーザはエージェント同士の発話交替に参加できる。これにより、「ユーザが視線を向けたエージェントが話し出す」ことだけでなく、「ユーザが視線を向けたエージェントとは別のエー

ジェントが話し出す」ことや「ユーザが視線を向けても、エージェント同士の会話になかなか入れない」などの多様な発話交替状況をユーザは体験できる。

2. 観察による会話モデルの作成

2.1 発話交替と視線

発話交替における視線について、伝らは二人対話の会話コーパスの分析に基づき発話交替規則を明らかにしている³⁾。それによると、ターン(発言)構成単位の終了地点を移行適格場所(transition relevance place: TRP)とし、TRPでは、言語的要素、パラ言語的要素、非言語的要素により予測され、円滑な発話交替がされると考えている。そして、聞き手の予測の誤りなどにより、重複や沈黙などが起こることもあると述べている。

この知見に基づき、我々は親友同士の女性三人の会話を収録したビデオから発話交替を観測した⁴⁾。話中のときとTRP時において、話し手が発話終了時にどこを見ていたか(聞き手あるいは、それ以外)の視線に注目し、その回数を集計した。観察の結果、話し手の視線を受けないにもかかわらず、聞き手が発話を開始する場合も多いこと、TRP時に話し手が他の人を見ていても、聞き手が話を始める場合があることが分かった。表1にそれらの頻度を示す。この結果に基づき、頻度を確率としたベイジアンネットワークを作成し、エージェントに実装した。

3. エージェントへの実装

エージェントへの実装にはTVMLを用いた⁵⁾。外部から制御できるAPIも用いて、確率から制御できる

[†] 東京電機大学 情報環境学部 School of Information Environment, Tokyo Denki University

表 1 参加者の TRP における視線行動と話し出しの頻度

視線行動	頻度
話し終えた話者に見られた聞き手が話し出す確率	51%
話し終えた話者に見られなかった聞き手が話し出す確率	24%
話者が話し終えたときに1秒以上無話になる確率	14%
話者は聞き手を見ないが話し出す確率	10%
その他	1%
計	100%



図 1 (A) 三体のエージェント



図 2 (B) 二体のエージェントとユーザー (動作テストのため、ユーザーの視線位置は赤いマーカーで示される)

ようにした。また、今回はエージェント同士は、実際の人間の言語ではないが人間の言語を話しているように聞こえる音声を合成して作成し(例: ARABAHIKA などの意味の無い言葉)、エージェントに発話をさせる。本研究では、三体のエージェントのもの((A), 図1)と、赤外線を用いた視線検出装置(NAC社製)により、二体のエージェントの発話交替にユーザーが参加できるもの((B), 図2)の体験システムを作成した。

(B)のシステムでは、表1の確率を調整することで、ユーザーは「ユーザーの発話終了時に、視線を向けられたエージェントが話し出す」ことだけでなく、「ユーザーの発話終了時に、視線を向けられなかったエージェ



図 3 (B) におけるユーザーから見た画面(話に入りにくい状況の例)

ントが話し出す状況」「エージェントらの発話終了時に、ユーザーが話に入りにくい状況(図3)」「発話終了時に、ユーザーは発話せず、エージェントも話し出さず、沈黙が起こる状況」などが体験できる。体験した使用者からは、「向けた視線に反応していて人間みたいだった」「視線を向けていないエージェントに割り込まれるときがあり、人同士の会話みたいで驚いた」「友達同士で起こるような、なかなか話に入りにくい状況がありイラついた」などの感想があり、従来までのモデルでは表現できなかった多様な発話交替状況の生成ができていたことがわかった。今後はエージェントに個性を持たせた場合での検証と客観的な評価をしていく予定である。

4. ま と め

実際の人間の会話を観察し、ベイジアンネットを用いた視線モデルをエージェントに実装した。複数のエージェントによって生成される発話交替のシミュレーションシステムを作成した。これにより、ユーザーは多様な発話交替の状況を体験できる。今後は、ユーザーによる体験と評価を収集し、より自律的なエージェントの開発を目指す。

参 考 文 献

- 1) 松坂要佐, 東條剛史, 小林哲則: グループ会話に参与する対話ロボット構築, 電子情報通信学会論文誌 DII, Vol.84, No.6, pp.898-908 (2001).
- 2) 知野哲朗, 福井和広, 山口 修, 鈴木 薫, 田中克己: "GazeToTalk": メタコミュニケーション能力を持つ非言語メッセージ利用インタフェース, インタラクシオン '98 (1998).
- 3) 石崎雅人, 伝 康晴: 談話と対話, 東京大学出版会 (2001).
- 4) 徳永弘子, 湯浅将英, 武川直樹: 3人会話における発話交替時の視線行動分析 - 聞き手の立場から見た発話・非発話戦略 -, 信学技報告 HCS2006-45, pp.23-28 (2006).
- 5) TVML: <http://www.nhk.or.jp/strl/tvml/>.