

# ロボットの意思表示によるインタラクションの円滑化

山形佳祐<sup>†</sup> 大本義正<sup>‡</sup> 西田豊明<sup>‡</sup>

京都大学工学部情報学科<sup>†</sup> 京都大学大学院情報学研究科<sup>‡</sup>

## 1 はじめに

近年、高次の会話機能を実現することができる、身体性を持つ会話エージェント (Embodied Conversational Agent, ECA) が盛んに研究されている [1]. しかし、会話エージェントの多くはユーザが 1 人である場合を想定したものが多く、複数のユーザが同時に会話に参加する場合の会話エージェントに必要な要素はまだ確立されていない.

我々はインタラクションの円滑化を目指し、多人数インタラクションを行う会話エージェントの開発を進めてきた. そのために、我々はまずクイズにおける司会者が場を円滑に進めるケースを想定し、ユーザの個別の状況を識別し、認識結果やユーザのリクエストによってエージェントを動かすシステムを開発した [2]. しかしこのシステムで会話をするエージェントと、ランダムで会話をするエージェントとではあまり差が出なかった. その理由として、ディスプレイに司会をしている 2 次元エージェントを写し、そのエージェントによるポインティングとして、少し離れた場所に置いた 3 次元の矢印のポインタを使用して表していたため、エージェントとのポインティングとの同期がうまくいかず、ユーザの意識が分散してしまった点ということがある. また、ビデオによると、エージェントが会話をするタイミングが不自然であったということも考えられる.

そこで、本研究では司会者をロボットにして、エージェントとそのポインティングを一致させてインタラクションが円滑になるかを調べる.

## 2 システムの構成

各ユーザの識別をするために、Omron 社の開発した顔認識ライブラリである Okao Vision を用いて顔の位置、顔の方向、顔の大きさを取得し

Making Interaction Smooth by Expression of the Will of Robot

<sup>†</sup>Keisuke Yamagata

School of Informatics and Mathematical Science  
Faculty of Engineering, Kyoto University

<sup>‡</sup>Yoshimasa Ohmoto, Toyoaki Nishida

Department of Intelligence Science and Technology,  
Graduate School of Informatics, Kyoto University

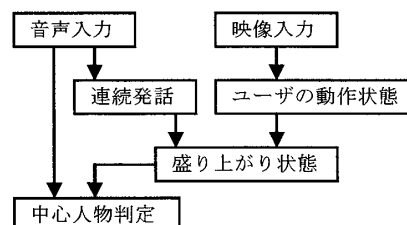


図1 盛り上がり状態及び、中心人物推定におけるデータフロー

た. また各ユーザに付けられたスロートマイクから音声入力を得た. 図1のように、これらの入力からユーザ間から会話の盛り上がり状態の推定をして、音声入力とその盛り上がり状態から会話での中心人物を推定した. その結果をロボットに送り、推定した中心人物の方向にロボットの顔を向けて、ヒントを用いるようにユーザを促したり、早く次の問題に移るように解答を催促したりといった行動や、盛り上がり状態のときには語り掛けないといった行動をした.

## 3 実験

実験の参加者は、3人1組で大学生を募集し、4組12名の参加者を得た.

本実験では状態推定に基づく行動表出を伴ったエージェント (以下、推定エージェントと呼ぶ)、伴わないエージェント (ランダムエージェント) との比較実験を行った. 下の表1は推定エージェントとランダムエージェントとの挙動の違いを比較したものである. それぞれの組に推定エージェントとランダムエージェントが司会をするクイズゲームに参加してもらい、10問のクイズを制限時間なしで答えてもらった. ユーザは問題に対する解答とヒントの提示を求めることをした.

表1

	推定エージェント	ランダムエージェント
ヒントを促すタイミング	盛り上がり状態か推定、タイマー	タイマー
催促のタイミング	盛り上がり状態か推定、タイマー	タイマー
発話の相手	中心人物	ランダム
補足説明	盛り上がり状態か推定	毎回行なう
回答後の沈黙	盛り上がり状態か推定	発生しない

アンケートには 7 段階評価での SD 法を用いて、14 の設問に対して、推定エージェントに対する評価とランダムエージェントに対する評価をしてもらった。

#### 4 結果

アンケートの結果を表 2 に示す。表中の数字は各設問に対し、推定エージェントとランダムエージェントのうち相対的に良いと感じた参加者の数を表している。

表 2

設問	推定	ランダム	引き分け	内容
9	1	7	4	親しみやすさ
10	3	4	5	喋るタイミングの適切さ
11	4	2	6	積極的に喋ったか
12	6	5	1	状況に応じた振る舞い
13	1	3	8	会話が弾んだか
14	2	4	6	一人で考える時間が短い
15	2	5	5	エージェントによって会話が弾んだか
16	1	7	4	エージェントが注意を向けているか
17	3	4	5	進行がスムーズだったか
18	1	8	3	頭の動きに注意を向けた
19	2	3	7	頭による指示がわかった
20	3	4	5	頭の向きが自然だった
21	4	1	7	沈黙を感じた
22	2	5	5	促しに素直に従った

このアンケートから、ユーザは相対的にランダムエージェントのほうが良いとみなしているが、細かくアンケート内容により分析してみたが、どこによって差が出たのかはわからなかった。しかし、ビデオ分析により推定とランダムのエージェントの発話がユーザの会話とかぶるときの被り方が違うことがわかった。そこで、表 3 に被験者が 10 問のクイズを答えるまでに、何回エージェントが催促や促しを行ったかの回数と、エージェントの発話が人の話し始めとかぶった回数を示している。

表 3

グループ (エージェントのタイプ)	エージェントの干渉回数	エージェントの発話が人の話し始めとかぶった回数	気が利いたと感じたエージェントのタイプ
1(ランダム)	7回	1回	ランダム
1(推定)	5回	4回	
2(ランダム)	3回	0回	ランダム
2(推定)	4回	3回	
3(ランダム)	4回	0回	推定
3(推定)	6回	3回	
4(ランダム)	4回	1回	ランダム
4(推定)	0回	0回	

「出題者が解答者(自分達)に注意を向けていると思いませんか」という設問 16 に対し、グループごとでエージェントの印象がどう違うの

かを調べるために、各グループで相対評価を取ったところ、グループ 1, 2, 4 がランダム、グループ 3 が推定のほうが良いと判断した。

この結果と表 3 から、エージェントの干渉回数がユーザの印象に影響していると考えられる。干渉回数が多い時の方が、エージェントは注意を払っているとユーザには感じられた。

また、アンケートでは、ランダムエージェントの方が相対的に好印象であった。ビデオにせると、会話にかぶるタイミングは、推定エージェントは発話しはじめる時、ランダムエージェントは会話最中であった。盛り上がり状態でなくなったときに話し始めようとする、次の会話の最初のほうにかぶってしまい、人は心理的に、話の途中よりも話し始めに会話をかぶせられるほうが印象が悪くなるので、推定エージェントの印象が下がったと考えられる。

ロボットの頭の向きに違和感があったと感じましたか、という設問 20 に対して違和感があったと感じたユーザは少なかった。先行研究でのユーザの意見として、矢印のポインタが動いたことはわかっても矢印の向きを気にしないという意見があったのに対して、今回の実験では、司会者が途中で自分を見ていることに気がついたという意見がはじめて挙がった。

#### 5 まとめ

以上より、司会者をロボットとして視線と動作を一致させた今回の実験では、推定エージェントとランダムエージェントを比較した結果、ランダムエージェントの方が司会者としての印象はよかった。よって、クイズにおいて会話や進行をスムーズに進めるには、不自然な間や、エージェントとユーザの話し始めがかぶったりすることを減らす、ということが大切だとわかった。また、実験順序により GNAT の結果が変わるので、様々な組み合わせで検証していく必要がある。

今後は、司会者であるロボットが会話に干渉する方法を探るため、WOZ での会話、話し始める前に動作によって気を引くこと、顔を向ける人物の名前を呼ぶこと、などの多人数ユーザ会話におけるエージェントの自然な会話介入方式を調べていく。

#### 参考文献

- [1] Toyooki Nishida, Conversational Infomatics An Engineering Approach, WILEY, 2007
- [2] Hung-Hsuan Huang, Human-Agent Interaction Symposium 2009, 2C-5, 2009
- [3] Nosek BA, Banaji MR. 2001. The go/no-go association task. Soc. Cogn. 19:625-66