

## 2 体のエージェントを用いることによる対話活性度の向上

### The Use of Two Agents to Stimulate Conversation Activity

吉岡 駿<sup>1\*</sup>, 松元 崇裕<sup>1</sup>, 岨野 太一<sup>1</sup>, 後藤 充裕<sup>2</sup>, 今井 倫太<sup>1</sup>

Shun Yoshioka<sup>1\*</sup>, Takahiro Matsumoto<sup>1,2</sup>, Taichi Sono<sup>1</sup>, Matsuhiro Goto<sup>2</sup>, Michita Imai<sup>1</sup>

<sup>1</sup> 慶應義塾大学

<sup>1</sup> Keio University

<sup>2</sup> NTT サービスエボリューション研究所

<sup>2</sup> Service Evolution Laboratories, NTT

**Abstract:** ヒューマンエージェントインタラクションにおいて、人とエージェントの対話を扱った研究はあるが、人がエージェントへ話すこと自体を前提としているものが多い。しかし、対話を開始するためには、人をエージェントとの対話へ引き込むことが必要になる。また、引き込みが成功しても、対話を通じて人を楽しませるのは容易ではない。そこで本研究では複数のエージェント間で活性度の高い対話を行うことで、人を対話へ引き込むアプローチの提案を行う。更に、活性度の高い対話によりエージェントとの対話の楽しさ、印象が向上するか評価を行う。

## 1 はじめに

近年、対話ロボットが日常生活に登場する場面が増えている。代表的なものとして Softbank のペッパー君が挙げられる。Softbank の店頭にとどまらず、今や観光に行った先々でお目にかかることができる。他にも、タカラトミーが発売しているオハナスや Robi Jr. など、娯楽商品としても対話ロボットは活躍の場を広げている。本研究では、これらの対話ロボットが人との対話を盛り上げ、対話の印象を向上させることを目的とする。

対話活性度は対話の盛り上がりを示すもので、2章で示すように様々な要素からなっている。対話活性度が高ければ、その場にいる人はその対話をより楽しんでいると言える。しかし、対話ロボットとの対話の多くは一問一答形式であること、ロボットが単なる命令対象であると認識されがちであることから、必ずしも対話活性度が高いとは言えないのが現状である。

多人数対話においてロボットが対話の仲介役となって対話を活性化させる取り組みとして、松山らの対話に参加できていない人の参加を促すもの[4]や瀬島らの仮想空間において対話活性度を可視化するもの[5]などが挙げられる。しかし、これらは受動的に対話の活性を促すものであり、ロボットが積極的に対話を活性化するものではない。

また、ロボット 2 体の間の対話を通して人との相互作用を図る取り組みとして、一条ら[1]は非言語

情報のオーバーラップにより、人がロボットの対話内容を積極的に理解しようとする認知的な引き込みを確認している。しかし、この研究は人の注意誘導システムの構築を目指しており、本研究とは目的が異なっている。

本研究では、2 体のロボットを用い活性度の高い対話を実現するために、音声と動作のオーバーラップに注目する。これらを同時に行うことで、それを聞いているユーザを対話に引き込むことを提案する。ユーザがロボットとの対話に引き込まれることで、対話の印象が向上することが期待できる。

本稿では、まず 2 章で関連研究および対話活性度について述べる。3 章ではオーバーラップの定義およびその実装方法について説明する。4 章で実装に用いたシステムを説明する。5 章では、実装の妥当性を検証する実験について述べ、6 章でまとめる。

## 2 対話活性度

### 2.1 関連研究

多くの対話の盛り上がりに関する研究は、人の対話の活性度を測り、その情報を基にロボットの行動を変更させるアプローチをとっている。例えば、松山ら[4]は発話量を基に対話活性度を測り、それが少ない人は対話に参加できていないものと判断し、対話のロボットはその人物に優先的に話しかける。また、瀬島ら[5]は仮想空間において仮想観客とし

\*連絡先：慶應義塾大学理工学部情報工学科

〒232-8522 神奈川県横浜市港北区日吉 3-14-1

E-mail: yoshioka@ailab.ics.keio.ac.jp

て向日葵のアバタをユーザの対話活性度に応じて動かすことで、場の盛り上がりを反映することを提案している。そうすることで、ユーザ同士の一体感を生み出し、円滑なコミュニケーションが実現されるとしている。

## 2.2 対話活性度の測定

対話活性度を測る研究は多く行われており、大きく分けて2つのアプローチがなされている。ひとつが音声情報に注目した計測方法である。声圧、ピッチなど直感的にわかりやすいものに加え、様々な要素が存在する。松山ら[4]は活性度の指標として、発話量、笑顔頻度、頷き頻度を挙げ、その中でも特に発話量に注目している。西村ら[6]は音声の韻律とオーバーラップと対話者の客観的な対話活性度の高さの関係を、瀬島ら[5]はオーバーラップと主観的な活性度の高さの関係をそれぞれ示した。対話活性度計測のもうひとつのアプローチは、動作に注目したものである。前田ら[7]は手の加速度や頭部の回転回数が活性度に影響を与えることを示した。また、小野ら[3]と喜多[2]はどちらも行動のオーバーラップと対話活性度の関係を示唆している。

## 3 オーバーラップ

オーバーラップの判定は、[1]の研究を今回用いるロボットから取得できる情報に適応するように拡張した。[1]では「閾値以上サーボ値を変化させるコマ間を対話活性度に関与する行動部と定義し」ているが、本研究では対話活性度に関与する行動部を「単位時間におけるサーボの回転角が閾値を超える部分」とした。単位時間におけるサーボの回転角を  $\theta_i$  ( $i = 1, 2, \dots, servo\_num$ )、サーボが閾値以上である判定を  $a_i$  ( $i = 1, 2, \dots, servo\_num$ )、動作が対話活性度に関与する判定を  $A$ 、オーバーラップの判定を  $O$  とすると以下の式で表せる。

$$a_i = \begin{cases} 1 & \text{if } (s_i \geq threshold) \\ 0 & \text{else} \end{cases}$$

$$A = \bigcup_{i=0}^{servo\_num} a_i$$

$$O = A_{sota1} \cap A_{sota2}$$

### 3.1 オーバーラップの実現

ロボットの行うことができる動作を基本動作と指

表1 対話行為と動作の対応

対話行為	動作
自己開示	肯定を表す 両手で頬を叩く 片手を挙げる
質問	疑問を表す
応答(Yes)	肯定を表す
応答(No)	否定を表す
あいづち	肯定を表す 片手を挙げる
フィラー	考え事をする

示動作に分けて処理を行った。基本動作、指示動作ともに、1体のロボットが同種の動作を同時に行うことはできないが、基本動作と指示動作は並列に行うことができるものとする。

#### 3.1.1 基本動作

基本動作は頭部と腕を使った動作で、ロボットの意図や感情を表現するために用いる。単純化するために次の6つの簡単な動作に絞った：頭を上下させ肯定を表す動作、頭を左右に振り否定を表す動作、頭を傾げて疑問を表す動作、右手/左手を挙げる、両手で頬を叩く動作、右下/左下に目を逸らし考え事をする。

これらの動作と発話の対応付けを行うために、一つ一つの発話に対話行為タグを付けた。対話行為タグとは、話者がどのような意図で発話をするのかというのを分類するためのもので、松浦ら[8]によって設計された。表1に用いた対話行為と対応する動作を示す。一つの対話タグに対して複数の動作が対応付けられているものに関しては、そのタグのつけられた発話毎にランダムで一つを選択する。また、もう一方のロボットの発話を聞いている際のオーバーラップさせる動作は、肯定を示す動作、片手を挙げる動作、両手で頬を叩く動作の中からランダムで選択する。また、オーバーラップさせる動作の開始時刻は相手の発話の開始と同時、発話の1/4の時間がたった時、発話の真ん中、開始から発話の3/4時間がたった時の中からランダムで選択する。

#### 3.1.2 指示動作

指示動作は頭部(目線)もしくは腕を使い、発話相手もしくは興味対象を示す。今回使用するロボットでは目線で方向を指し示す際は頭全体を動かす。指

示方向は単純化のため、ユーザ方向、もう一方のソータの2方向に限定する。目線を使った指示動作では、指示対象が移動するまで同じ方向を見る。しかし、腕にはぶらんとおろしている状態があるため、ある時間指し示した後、しべし続けずに下すのが自然である。このような性質の違いから、指示をするために目線を用いるか腕を用いるかは、指示する時間の長さによって決定した。具体的には閾値を設け、閾値より長い時間指示し続ける場合に目線を用い、閾値より短い場合に腕を用いる。

発話に伴って指示対象がユーザからもう一方のロボットに、もしくはもう一方からユーザに移動する際は、発話開始時に発話を行っているロボット(S1)が目線ないし腕で新しい指示対象を示し、その一定時間後にもう片方のロボット(S2)が目線で同じ対象を指示する。S1が腕による指示をした場合、S2はS1が腕を下すと同時に目線を元に戻す。

## 4 実装

### 4.1 システムの構成

本研究では対話ロボットとしてヴィストンのソータ(Social Talker)を2体用いる。ソータの本体は高さ280[mm]、直径160[mm]、重量763[g]である。頭部にはカメラとマイクが取り付けられている。頭部と胴体の取り付け部に3つのサーボモータがあり、3自由度で回転可能である。また、それぞれの腕に2つずつ、胴体と底部の取り付け部に1つサーボモータが付いている。他に、目と口のLEDがあり、これらはソータの内部に搭載されたコンピュータ(Intel Edison)によって制御される。発話時の音声合成では1つの音量、ピッチ、速度でのみ、発話可能である。このことから対話活性度を音声の韻律で表現することはできない。また、表情を変化させることはできない。しかし、これらは動作と発話オーバーラップの実現には必須はでない。

2体のロボットの操作および同期を行う手段としてNTTのR-envを用いた。R-envにおいて、あるシナリオは行動の集合が記述されたノードと遷移条件が記述されたエッジからなる状態遷移図として作成される(図1)。容易に2体のロボットの同期を図る



図1 R-envのシナリオ作成画面

ことができる。

## 4.2 対話シナリオ

複数人が対話するシナリオとして、目的を持った対話を多対多で行う「会議」、目的を持った対話を一対多で行う「講義」、目的を持たない対話を多対多で行う「雑談」などが挙げられる。本研究ではより実生活に近い状況を想定するため、雑談を行う。以下に雑談の例を挙げる。

S1: こんにちは  
 S2: こんにちは  
 S1: お話ししよう  
 S2: うん、しよう  
 S1: 今日はいい天気だね  
 S2: でも天気予報によると、これから雨が降るらしいよ  
 S1: そうなんだ。傘持っていないや  
 S2: 僕もない  
 S1: ねえ、傘持ってる?(ユーザへの問いかけ)

## 5 印象評価

4章で説明したシステムにより、ロボット間での対話活性度が高くなることを示すために実験を行った。比較対象として、発話内容、および発話のタイミングは変わらず、動作のオーバーラップを含まないシナリオを用意してどちらがより対話活性度が高く見えるか、被験者6人(男3人、女3人)による評価実験を行った。それぞれの被験者には3.1章で説明した手順でロボット間のオーバーラップを含む約40秒の動画とオーバーラップを含まない同じく約40秒の動画を観てもらい、それぞれに関して「対話が盛り上がっていると感じたか」「対話が楽しそうであると感じたか」の2点について1-5(そう思わない—そう思う)の5段階で評価してもらった。

### 5.1 結果

実験の結果を表2に示す。対話の盛り上がりも対話の楽しさ、いずれも動作のオーバーラップがある場合の方が高い平均を得ることができた。しかしt検定の結果は「対話の盛り上がり」が0.17で有意差はみられず、「楽しさ」は0.08で有意傾向にあるのみにとどまった。

また自由記述の欄において、被験者#2は二つの動画を観た後で、「一つ目(オーバーラップなし)よりも

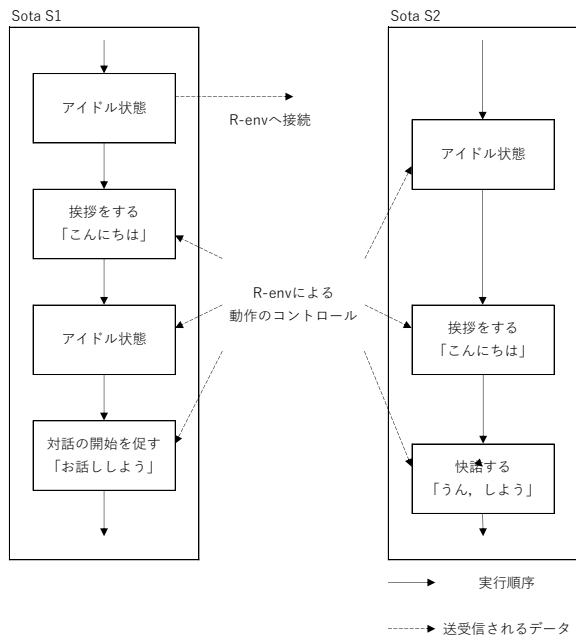


図 2 ロボット同士の対話の例

2 つ目(オーバーラップあり)の方が会話のテンポが良かった」と回答している。

## 5.2 考察

今回、有意差が見られなかった要因として、対話が約 40 秒と短かったことが考えられる。また、実際にロボットを目の前にしての実験ではなくビデオを見ての実験だったことも、有意差が出なかった原因と考えられる。

しかし、被験者#2 による、自由記述は興味深い。先述の通り、2 つの動画の違いは動作のオーバーラップのみで発話のタイミングはどちらも同じにもかかわらず、このように感じたことは、オーバーラッ

表2 実験結果

被験者#	盛り上がり		被験者#	楽しそう	
	オーバーラップあり	なし		オーバーラップあり	なし
1	4	4	1	4	2
2	3	1	2	3	1
3	5	4	3	5	4
4	3	4	4	4	5
5	4	3	5	5	4
6	5	4	6	5	4
平均	4.0	3.3	平均	4.3	3.3

プの効果を示唆しているといえる。

## 6 まとめと今後の展望

発話内容からオーバーラップを含んだ動作を生成し、その動作を行っているロボットが人に与える印象を検証した。

今後の計画として、発話のオーバーラップも含め、より対話活性度の高いロボット間の対話を設計するにはどうすればいいのかを検討する。そして、ロボット2体に人間も加えたシナリオを作成し、ロボット間の対話活性度(オーバーラップ)が人間の対話活性度に影響するか否か、人間に好印象を与えることができるか否かを確認する実験を行う予定である。

## 7 参考文献

- [1] 一条剛志, 棟方渚, 小野哲雄: 複数ロボットの対話活性度を用いた注意誘導システムの研究, HAI シンポジウム, pp. 250-255, (2015)
- [2] 喜多壮太郎: 動作の認知科学 人はなぜ動作をするのか, 認知科学, Vol. 7, No. 1, pp. 9-21, (2000)
- [3] 小野哲雄, 今井倫太, 石黒浩, 中津良平: 身体表現を用いた人とロボットの共創対話, 情報処理学会論文誌, Vol. 42, No. 6, pp. 1348-1358, (2001)
- [4] 松山洋一, 齋藤彰弘, 秋葉巖, 渡邊萌美, 小林哲則: 「置いてけぼり」を救う多人数会話活性化ロボット, HAI シンポジウム, (2012)
- [5] 瀬島吉裕, 渡辺富夫, 石井裕: 仮想観客インターフェースに会話活性推定モデルを付与した場の盛り上がり支援システム, 日本機械学会論文集, Vol. 79, No. 807, pp. 94-106, (2013)
- [6] 西村良太, 北岡教英, 中川聖一: 対話における音律変化・タイミングのモデル化と音声対話システムへの適用, 人工知能学会言語・音声理解と対話処理研究会, No. 48, pp. 37-42, (2012)
- [7] 前田貴司, 高嶋和毅, 梶村康祐, 山口徳郎, 北村喜文, 岸野文郎, 前田奈穂, 大坊郁夫, 林良彦: 3人会話における非言語情報と「場の活性度」に関する検討, 電子情報通信学会技術研究報告, Vol. 109, No. 185, pp. 43-48, (2010)
- [8] 松浦大樹, 徳久雅人, 村上仁一, 池原悟: 心的状態の解析のための対話行為タグ付きテキスト対話コーパスの試作, 電子情報通信学会技術研究報告, Vol. 106, No. 486, pp. 31-36, (2007)