

事例ベースを用いた擬人化エージェントのしぐさ表出

Generation for gesture of embodied conversational agent by using a case base

尾本和也¹ 岡田将吾² 新田克己³

Kazuya Omoto¹, Shogo Okada², and Katsumi Nitta³

東京工業大学(院) 総合理工学研究科 知能システム科学専攻
Department of Computational Intelligence and Systems Science,
Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology

Abstract: Our goal of the research is to implement the embodied conversational agent which recognizes the transition of issues or topics in the discussion by users and generates the appropriate nonverbal behavior corresponding to the issues. In this paper, we analyze the nonverbal behavior of users by observing actual discussions and relationship between the nonverbal behavior and issues in the discussion. Moreover, we proposed the system that output candidates of nonverbal behavior which the agent generates by using issues as the input queries. The embodied conversational agent which works based on the proposed argument system can speak using text and output nonverbal behavior.

1 はじめに

擬人化エージェントは非言語情報をユーザーに伝達出来るため、対話・会話システムに関する分野で盛んに研究されている。例えば、教育分野では数多くの擬人化エージェントが開発され、対象物体を直接示しながら教育できる点や、学生のモチベーションの向上のつながる点や、コンピュータと学生のコミュニケーションの幅を広げる点などが利点として挙げられる[1]。

また、我々は法学部における議論スキルに関する教育支援システムの開発を行ってきた。このシステムは、オンラインでメッセージを交換する、チャット形式の議論環境を提供する。この議論環境のもとで、人間同士の議論の内容を観察し、議論中に割り込んで、次に行うべき発言の助言を与える「助言エージェント」や、人間の代わりに発言を自律的に行う「議論エージェント」の開発を行ってきた[2][3]。これらのエージェントの知識源としては、その議題に関する背景知識（想定される論点）や、その議題に関する過去の議論記録（事例ベース）などを用いていた。このような知識源を利用できるのは、法学部における議論教育においては、教師が課題を決めた段階で、主要な論点があらかじめ想定出来るからであり、また、同じ課題で何組もの学生が議論演習を行うからである。

開発したシステムを法学部の演習に実際に応用し

て評価実験を行ったところ、議論の初級者からは有用であるとの評価をいただいたものの、中級者からは現実の議論教育を行うには必ずしも十分な助言/議論機能を持っていないことが指摘された。また、ユーザインタフェースとして、発言時にアニメーションの表情変化を起こすことで非言語情報を伝達していたが、表情変化が単調であることが指摘された。

このようなことから、今後の議論エージェント研究の方向性として、我々は、(1)事例ベース中の議論記録の分析をあらかじめ行って特徴抽出をすることによって、より高度な議論エージェントを開発すること、(2)事例ベースに非言語情報を加えることにより、より親しみやすいエージェントを開発すること、を将来的な課題としている。

本研究では、このうちの(2)に着目し、今までのテキストベースの事例ベースを拡張して身振りの情報を付け加えた事例ベースの構築と利用について述べる。

一般的な対話の会話エージェントの研究分野では、発言内容に応じた身振り生成の研究はなされているが、議論における身振り生成は一般的な会話よりも複雑になりがちである。それは、議論の中に、対立する場面や協調する場面があり、場面に応じて身振りをきめ細かく制御する必要があるからである。また、同じ場面であっても、人によってその身振りは大きく異なることがある。このような場合に、非言語情報を利用した事例ベースがあれば、それを利用

することによって擬人化エージェントの身振り生成が容易になると考えている。

2 システム概要

本研究では、複数のユーザが議論サーバに接続して、サーバを介してメッセージを交換する議論環境を想定している(図1)。議論サーバには、議論の主題ごとに、その議論に関する背景知識が論点リストとして用意されている。ここで論点は、事実や主張を表す命題と考える。例えば、以下は論点の例である。f1~f4は論点のIDである。この例では、f1はf3の根拠であり、f2はf4の根拠である。f3とf4は対立している。

- f1: 売った商品が故障していた。
- f2: 発送時には故障はなかった。
- f3: 売り主に落ち度がある。
- f4: 売り主に落ち度はない。

議論で交換されたメッセージについて、その中に出現する単語の組から、素意の発言に含まれる論点を抽出し、その論点の推移を観測して議論の推移を把握する。

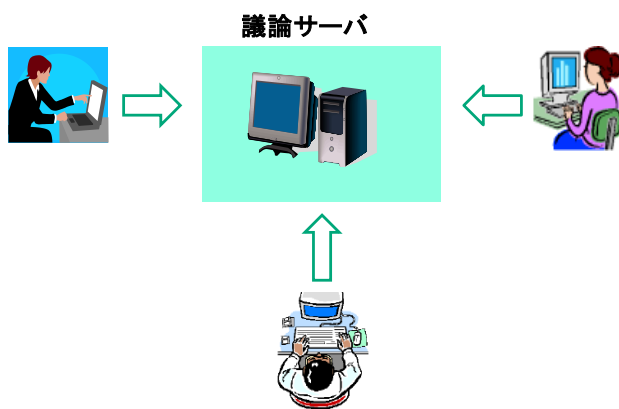


図1 議論サーバを介した議論環境

議論エージェントは、図1の議論環境において、人間の代行として発言を行い、あわせて、アニメーションの身振りを変化させる機能を持つ。

議論エージェントの概要を図2に示す。このエージェントは、「論点リスト」「マルチモーダル事例ベース」「身振り道がデータベース」の3つの知識源を持つ。「論点リスト」は上述したように、議論の主題に関する背景知識を与える。マルチモーダル事例ベースは、その主題に関する議論の記録である。この事例ベースには、テキスト情報だけでなく、そのときの身振りに関するタグ付けがなされている。同一

の主題に関して、複数の議論記録があることを想定している。

相手から発言と身振り情報が届くと論点リストを参照して、論点と身振りラベルの情報に変換される。

これをマルチモーダル事例ベースと比較することによって、過去の類似場面を複数列挙する。マルチモーダル事例ベースには、過去の発言記録の他にそこに含まれる論点や身振りラベルが発言ごとに記録されている。類似検索においては、前回の2回分の発言について、論点と身振りラベルの一致度を考慮して、類似場面を検索する。

一方、マルチモーダルデータベースに含まれている複数の議論記録を比較し、論点の移動確率やそれぞれの場面における各身振りの頻度などの統計情報を分析する。これは、大勢の学生の議論傾向を知るためであり、たとえば、相手が同じ発言を身振りをしたときに、「頭をかきながら謝る」場面と、「頭を下げて謝る」場面と、「上を向いて、いなおる」場面のどれが多いかを統計的に得ることによって、「平均的な応答(発言,身振り)」を推測することができる。

列挙された複数の類似場面について、議論の戦略や統計的な解析に基づき、相手への応答としてふさわしい場面を選択し、その身振りに相当する動画データを身振りデータベースから検索して、それを議論相手に提示する。

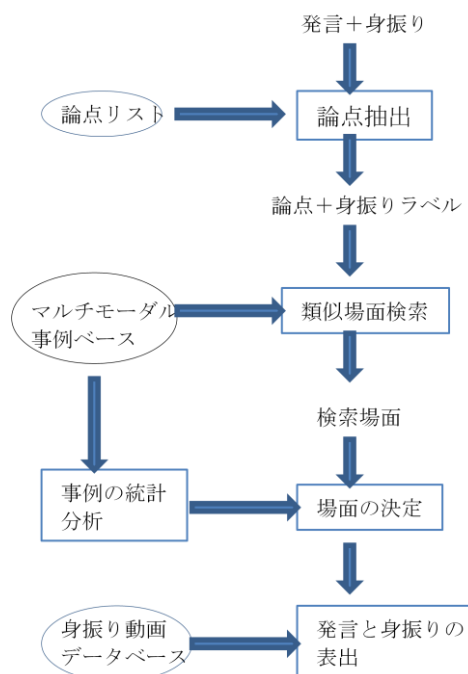


図2 議論エージェントの動作概要

3 システム詳細

議論エージェントの主な構成要素の説明を行う。

3.1 論点リスト

議論は自然言語によるメッセージの交換で行われるが、同じ内容でも発言の言い回しや表現方法は人によって様々である。そこで、われわれは以前から論点ベースの議論解析を行ってきた。

前述したように、ここでは、「論点」は事実や主張を表す命題であると考え、論点間には、一方の論点の成立が他の論点の成立を支援する(根拠となる)関係や、一方の論点の成立が他の論点の成立を妨げる(矛盾する)関係がある。このような関係を、図3に示すような論点グラフとして表現することが出来る。このグラフは議論を開始する前に、議論の主題が決まった段階で大雑把に決定することが出来る。議論を行うと、この議論グラフの上で論点が移動する。同じ議題の議論でも、人によって論点の遷移の様子は大きく異なる。これを観測することによって、議論のスキル評価や類似場面の検索を行っている。

発言のメッセージから論点を抽出するには、メッセージを形態素解析し、出現する単語の組み合わせを利用して論点の推測を行う。われわれは、論点別の発言記録から、その論点を抽出するための単語の組を識別するための機械学習の方式を、別途提案している。

メッセージから複数の論点が抽出された場合、その組み合わせによっては、論証(結論と理由の組)を抽出できる場合がある。たとえば、一つのメッセージ中に f1, f3 という2つの論点が抽出されたとき、論点グラフ上で、f1 と f3 には「f1 が f3 を支持する」という関係にあるとき、これは「f1 だから f3 である」という論証を表していると推測する。

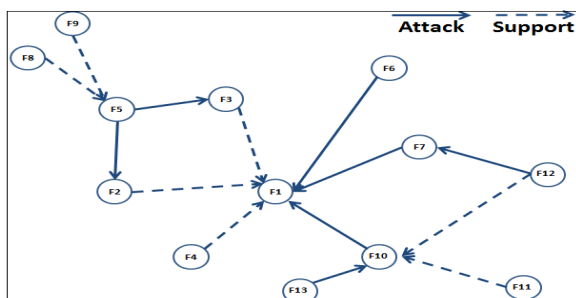


図3 論点間の関係図

3.2 マルチモーダル事例ベース

本研究で用いるマルチモーダル事例ベースは、議論中に撮影した動画を元に身振り情報を頭部、胴体、腕部に分けて抽出し、抽出された情報を発言ごとに iCorpus studio[4]を用いてタグ付けし(図4)、以下のようにXML形式で保存する。

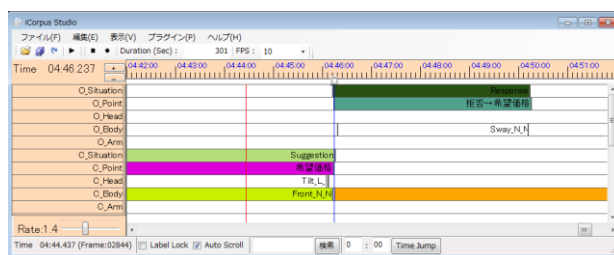
```
<発言 id="15" 発言者="山田" 論点="F3"  
  頭部="傾き" 胴体="真っ直ぐ" 腕部="腕組み">  
  契約の解除は認められない  
</発言>
```

タグ付けは、発言者だけでなく、相手の発言を待っている側の身振りについても行う。

観測する身振り情報を表1に示す。この他に、発言の言い回し(丁寧な口調、乱暴な口調など)や表情変化などのさまざまな情報が考えられるが、今回は身振りの情報のみに限定している。



(a) 議論の動画記録



(b) iCorpus studio

図4 身振りのタグ付け

表 1: 非言語情報の種類

部位	種類
頭部	真っ直ぐ・右傾・左傾・上向き・下向き・傾き・前方に出す・傾げる
胴体	真っ直ぐ・前傾・後傾・右傾・左傾・揺れる
腕部	腕組み・手を前に出す(縦)・手を前に出す(横)・顎を触る・口を隠す・頭を触る

3.3 身振り動画データベース

身振り動画データベースは、頭部、胴体、腕部のそれぞれのラベルに応じて、あらかじめ作成したアニメーションの身振りの動画ファイルのデータベースである。ゲーム用動作センサ Kinect を利用して、人間が行った身振りを識別し、アニメーションの動作生成ソフト MikuMikuDance を利用して蓄積している。

3.4 論点抽出

前述したように、入力されたメッセージは、形態素解析されて、そこに出現した単語の組み合わせによって論点を推測する。論点が複数、推測された場合、2つの論点が並列に存在する場合と、一方の論点他方の論点の根拠になっている場合、一方の論点他方の論点の否定になっている場合がある。このような関係を論点グラフを用いて判断しながら、論理構造を組み上げていく。

身振りの抽出に関しては、現時点では人手による抽出に留まっており、自動抽出は出来ていない。しかし、Kinect による動作認識の実験を始めており、将来的には身振り認識の自動化を目指している。

3.5 類似場面検索

議論の現時点の場面と類似する場面を事例データベースからの検索は以下のように行う。

連続する2発言の組の論点とラベルを考える。

(id=t-1, 自分の発言, 論点, 頭部, 胴体, 腕部)

(id=t, 相手の発言, 論点, 頭部, 胴体, 腕部)

これが一致するとき、この次の自分の発言

(id=t+1, 自分の発言, 論点, 頭部, 胴体, 腕部)

が、次の議論エージェントの発言と身振りとする。

この2発言の組を検索キーとし、事例ベース中の各発言記録中の連続する2発言が一致する場面を検索することになる。ただし、身振りのラベルが完全に一致しなくても、論点一致していれば類似しているものとし、身振りラベルが一致する数が多ければ類似度が高いと判断する。これは、人によって、同じ発言でも身振りの個人差があるからである。

3.6 事例の統計解析

同一課題に関する複数の議論データを比較することによって、全体の平均的な身振りや、個人の特徴的な身振りを測定することが出来る。たとえば、表2は、3つの議論記録の中で同じ論点の発言に関して、特徴的に観測できた身振りを示したものである。このように3人の動作に関して、異なる傾向の動作をしていることが示されている。

表 2: 人による身振りの相違

実験参加者	非言語情報(大きさ,速さ)	出現回数
参加者A	首を傾げる(小,速)	10/12
	揺れる(普,普)	9/12
参加者B	顎を触る(普,普)	5/8
参加者C	前傾(大,普)	5/6
	手を前に出す(縦)(普,普)	5/6

3.7 場面の選択

類似場面検索において、事例ベースから複数の場面が検索される。検索されたある場面は、相手に反論するかもしれないし、ある場面は、相手に妥協するかもしれない。また、同じ発言をする場合も、表2に示したように、異なる身振りで応答することがある。

検索された場面からどれを選択するかという戦略が、この議論エージェントの個性を決定する。これには様々な戦略がある。

(1) 特定の人に合わせた戦略

検索された場面の中で、特定の人(例えばAさん)の場面を選択する戦略である。これは、Aさんと同じような発言と身振りをするエージェントを実現したい場合に有効である。

(2) 平均的な反応に合わせた戦略

事例ベースの統計的分析により、同一場面が一番多い論点と身振りを選択する戦略である。平均的なエージェントを実現したい場合に有効である。

(3) あらかじめ設定された個性に合わせた戦略

たとえば、相手に有利な論点を優先的に選択する

ことによって、協調的なエージェントを実現することができる。反対に、自分に有利な論点を優先することにより、利己的なエージェントを実現することが出来る。また、論証を含む発言を優先的に選択することによって、理屈が多いエージェントを実現することが出来る。さらに、大げさな身振りを優先的に選んだり、大げさな身振りを避けたりすることによって、個性的な身振りの傾向を実現することが出来る。

われわれは被験者を使った予備実験において、相手のエージェントに感じた印象が、あらかじめ設定したエージェントの個性とかなり一致することを確認している[2]。

3.8 動画選択

場面の選択がなされると、相手に応答する発言と、身振りラベル（頭部、胴体、腕部）が決定する。この動作を身振り動画データベースから選択して、相手に表示する。現段階では、論点と身振りのラベルから一意に動画を選択している。

しかし、前述したように動作は Kinect を用いて作成するため、同じ身振りでも人によって異なる印象の身振り動画が生成される。したがって、動画作成者の名前も検索キーに含めて、個性的な身振りを選択することも可能である。

4 エージェント動作例

この章では、実際に開発したエージェントの挙動の例について述べる。

議論の課題として、マンションを借りる際の家賃交渉の問題を設定した。この課題は、大家と借り手の間で家賃の交渉をすることを想定し、交通の便、駅からの距離、騒音、家賃相場、他の類似マンションの存在、などの情報を設定した。

この課題について、交渉記録をビデオ撮影して収集し、人手でテキストと動作を抽出して事例ベースを構築した。

同一課題で、人間が交渉エージェントと新たな交渉を行ったときに、ユーザーが表出するしぐさによって、エージェントが表出するしぐさが変化すること観測するために、検索で用いるユーザーの論点とエージェントが用いる返答用の論点を固定した。

まず、ユーザーが「希望価格の提示」という論点を含む発言を行い、その発言を行う際に、頭部を傾げ、胴体を前傾させている場合を考える。そこで、そのユーザーが表出した情報を入力し、エージェントが返答を行う際に用いる「提示価格の拒否」とし

入力を行った。その結果、エージェントは頭部を傾げ、胴体を揺らしているしぐさを表出した。実際にエージェントが表出したしぐさのスクリーンショットを図5に示す。

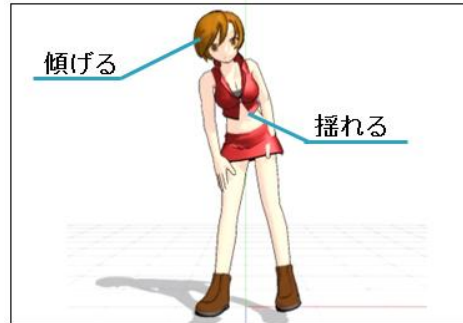


図5 エージェントしぐさ例1

次に、同じくユーザーが「希望価格の提示」という論点を含む発言を行った場合を考える。今回は、その発言を行う際に胴体を揺らしながら、顎を触ったとする。同様に、ユーザーが表出した情報を入力し、エージェントが返答を行う際に用いる「提示価格の拒否」とし入力を行った。その結果、エージェントは頭を触るしぐさを表出した。ここで、エージェントが表出したしぐさのスクリーンショットを図6に示す。

このように議論エージェントは、ユーザの身振りによって異なる応答を示している。



図6 エージェントしぐさ例2

5 おわりに

マルチモーダル事例ベースを用いた議論エージェントの基本的な動作を説明した。議論エージェントの応答の品質を決定するのは、豊富な事例の収集と、類似検索の手法と、場面選択の手法である。現時点では、基本機能が実装された段階であり、これからこれらの手法を洗練化していく。また、アニメーションの品質も議論エージェントの印象に大きく影響

する。今回は MikuMikuDance を使ったが、これには別の手法や別のキャラクターも存在するので、それらを試す必要がある

参考文献

- [1] W. Lewis Johnson, Jeff W. Rickel, and James C. Lester.:
Animated Pedagogical Agents: Face-to-Face Interaction
in Interactive Learning Environments, Artificial
Intelligence, (2000)
- [2] 前田憲生, 新田克己: 調停教育における論争エージェ
ントの開発, インテリジェントシステム・シンポジ
ウム講演論文集, Vol. 16, pp. 199-204, (2006)
- [3] 田中貴紘, 新田克己: 事例に基づいたオンライン調
停支援システムの研究, 人工知能学会誌, Vol. 22,
pp. 145 (2007)
- [4] iCorpus Studio, <http://www.ii.ist.i.kyoto-u.ac.jp/>
iCorpusStudio/,