

ボケて返す対話型エージェントの基礎検討

鈴木 奨¹ 呉 健朗¹ 瀧田 航平¹ 堀越 和¹ 中辻 真² 宮田 章裕^{1,a)}

概要：

発展を続ける情報分野を支える技術の一つである対話型エージェントは、今後もより多くの場面での活躍が期待されている。一方でエージェントとの無機質な対話に親しみを感ぜないユーザには、このような対話型エージェントは受け入れてもらえない可能性が懸念される。この問題を解決するために我々は、ユーザの発言の一部をわざと間違えて聞き返す、ボケて返す対話型エージェントを提案する。このエージェントにより対話にユーモアが生まれ、ユーザが親しみを感ぜることが期待される。我々は、認知科学領域で支持されている不適合-解決モデルに基づいてユーモアのある対話モデルを構築した。このモデルでは、元の単語と音が近く意味が遠い単語の中で、ユーザが理解できると思われる単語を聞き間違え語とする。また、実験により本提案手法の有効性が確認された。

A Study of a Conversational Agent Replying with a Joke

SHO SUZUKI¹ KENRO GO¹ KOHEI TAKITA¹ NAGOMU HORIKOSHI¹ MAKOTO NAKATSUJI²
AKIHIRO MIYATA^{1,a)}

1. はじめに

現代社会において対話型エージェントは、看護やショッピングなど様々な場面でみかけられるようになった。将来的には家庭や介護などの人間との良好なコミュニケーションを求められる場面での活躍も期待されている。しかしエージェントとの無機質な対話に親しみを感ぜないユーザにはこのような対話型エージェントは受け入れられない可能性がある。[1]では親和的な関係の形成に笑いが欠かせないとされており、ユーモアと親しみの関係性が窺える。そこで我々は対話型エージェントにユーモアのある対話をさせることで、ユーザは親しみを持つことができるのではないかと考えた。この発想に基づき、本研究では、対話型エージェントによるユーモアのある対話実現についての基礎検討を行なう。

本稿の貢献は、ユーモアのある対話を実現することで

ユーザに親しみを持たせ、対話型エージェントがユーザに受け入れられないという問題の解消を狙う点である。

2. ユーモア発話を行なうエージェントの研究事例

ロボットや対話システムがユーザに笑いを提供する技術は大きく分けて、一方的に話すエージェントとユーザと対話を行なうエージェントに分けられる。

一方的に話すエージェントの例として、[2], [3], [4], [5]が挙げられる。[2]ではロボットがユーザに笑い感情を誘起させる手段として大喜利が用いられており、[3]は文の感情に着目してボケの生成を行なっている。これらはいずれも、笑いを通してユーザとエージェントのコミュニケーションをより良いものにするという試みである。また[4], [5]では駄洒落や漫才の形式を用い、ある単語を別の単語に置換することでエージェントによる笑いの実現を目指している。このような研究はエージェントによるボケをユーザに見せることで笑いの提供を試みているため、ユーザとエージェントの間で対話などの直接的なコミュニケーションは発生していない。

¹ 日本大学 文理学部
College of Humanities and Sciences, Nihon University

² NTT レゾナント株式会社
NTT Resonant Inc.

^{a)} miyata.akihiro@nihon-u.ac.jp

ユーザと対話を行なうエージェントの例として [6] が挙げられる。[6] は、単語間類似度を用いたユーモア発話の自動生成手法を提案している。

3. 研究課題

今後より普及していくと予想される対話型エージェントは、未だその対話の多くが無機質なものである。無機質な対話に親しみを持てないユーザには、このようなエージェントは受け入れられないという問題が懸念される。こうした背景の中、笑いを通してユーザと良好なコミュニケーションを築こうとしている事例はいくつか存在する [2][3][4][5]。しかし、これらはエージェントが一方向的に話すことでユーモアを表現しており、ユーザと対話は行なっていない。[6] は対話中でのユーモアの表現を試みているが、突飛な発言を行なうなど、ユーザにユーモアとして受容されにくい場合もある。著者らはこれを、ユーザに納得できる形で表現することでユーモアとして受容される可能性があるとしている。そこで我々は、対話型エージェントとのコミュニケーションにユーザが親しみを感じられるシステムの実現に向け、対話の中に自然なユーモアを取り入れることを研究課題とする。

4. 提案手法

認知科学研究者の大半が「不適合の認知」がユーモア生起に不可欠と主張している [7]。中でも漫才・落語・4 コマ漫画のようなユーモアは「不適合-解決モデル」で説明できる [8][9][10][11][12]。そこで我々は、エージェントによるユーモアのある対話実現に向け、認知科学領域で支持されている不適合-解決モデルを参考にする。ユーザが入力した単語とエージェントが出力した単語の概念距離を離すことで不適合を作り、その不適合を音が近いという視聴覚的類似性によって解決することで、ユーモアを生み出す。また、[13] で定義されている下記のボケの作り方も参考にする。

- ある音から連想する、意味の違う言葉を全て思い浮かべる。
- その中からできるだけ意味に差のある 2 つの言葉を選び出す。
- 選び出した言葉をタイミングに合わせて使う。

上記は人がボケを作成する際の作り方であるため、エージェントがボケを作成するにあたっていくつか変更を加える。聞き間違いとして聞き返すというシチュエーションに限定することでタイミングを合わせる必要をなくす。さらに一般的によく使われる単語を選ぶことで、ユーザが理解できない単語を出力することを防ぐ。

以上の定義に基づき、我々はエージェントによるボケの作り方を次のように定義する。

- 入力単語と出力候補単語の編集距離を算出し、この値が小さいほど高い s_e (Edit distance score) を与え

る*1。この s_e が高いほど音が近いとする。 s_e は下記のように計算される。

$$s_e = \frac{1}{1 + d_e} \quad (1)$$

d_e は入力単語と出力候補単語の編集距離である。

- 入力単語と出力候補単語の概念距離を算出し、この値が大きいほど高い s_s (Semantic score) を与える。この s_s が高いほど意味が近いとする。 s_s は下記のように計算される。

$$s_s = d_s \quad (2)$$

d_s は入力単語と出力候補単語の概念距離である。

- コーパス内での単語の出現数の対数を取り、出現頻度の高い単語ほど高い s_f (Frequency score) を与える。この s_f が高いほどユーザが単語の意味を理解しやすいとする。 s_f は下記のように計算される。

$$s_f = \log f \quad (3)$$

f は出力候補単語のコーパス内での出現回数である。このとき単語の出現頻度はべき分布に従うため、ごく一部の単語の出現頻度が極端に大きい。これらの単語が最終的な総合 Score に与える影響が大きくなりすぎないように、出現数の対数をとったものを s_f とする。また、同様の理由から s_f の最大値に制限を設ける。

- 上記の s_e , s_s , s_f の合計値を最終的な s (Score) とし、最も s の高い単語を出力単語とする。 s は下記のように計算される。

$$s = w_e s_e + w_s s_s + w_f s_f \quad (4)$$

w_e , w_s , w_f は重み係数である。

5. 実装

5.1 事前準備

Wikipedia 記事全文を形態素解析し、不要ページ、不要品詞を除去して分かち書きしたものをコーパスとし、出力候補とする単語の標準形の読み方リストと言語モデルを作成する。ここでの不要品詞とは IPA 品詞体系において、記号、助詞、助動詞、接続詞、副詞、連体詞、非自立、代名詞、接尾、数、サ変・スルと分類されるものを指す。読み方リスト、言語モデルの作成には MeCab[14], word2vec[15] を用いる。

5.2 各 Score の計算

5.2.1 s_e : 編集距離 Score

ユーザが入力した単語 A について、標準形の読み方 (カ

*1 編集距離とは 2 つの文字列がどの程度異なっているかを示す距離であり、1 文字の挿入・削除・置換によって一方の文字列をもう一方の文字列に変形するのに必要な手順の最小回数として定義される。

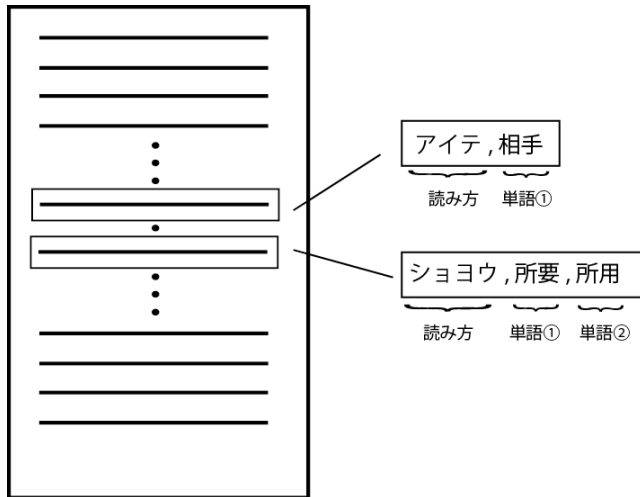


図 1 読み方リスト

表 1 s_e が高い例・低い例

元の単語	s_e が高い単語	s_e が低い単語
情報	乗法	若年者
花火	花見	廃棄物

表 2 s_s が高い例・低い例

元の単語	s_s が高い単語	s_s が低い単語
こんにちは	硬化	こんばんは
食事	小惑星	食費

表 3 s_f が高い例・低い例

s_f が高い単語	s_f が低い単語
駅前	詠嘆
目前	耄碌

タカナ) を MeCab を用いて取得し、コーパスを元に作成された単語の読み方リスト内に登録された単語のうち、最初の 1 文字が一致している読み方を持つ単語を出力候補単語とする。出力候補単語と単語 A の読み方の編集距離を計算し、その距離が近いほど高い Score とする。

5.2.2 s_s : 概念距離 Score

前述の出力候補単語それぞれに対し、単語 A との概念距離を計算し、その距離が近いほど高い Score とする。概念距離の計算には word2vec を用いる。

5.2.3 s_f : 出現頻度 Score

前述の出力候補単語それぞれに対し、単語ごとの Wikipedia コーパス内での出現数が多いほど高い Score とする。

5.3 出力単語の決定

以上 3 つの Score を算出し、それぞれ正規化処理、重み付けを行ってから合算したものを、単語ごとの最終的な Score とする。本稿では重み係数は全て 1.0 とする。算出された Score のうち、最も高い Score を保持する単語を出力単語 B とする。また、人名や地名といった固有名詞は、

表 4 返答がオウム返しになっている例

入力単語	出力単語
こんにちは	コンニチハ
りんご	リンゴ

ごく一部の有名なもの以外はユーザに理解されない可能性が高いと考え、出力候補単語から除く。また、単語 B が、単語 A を平仮名、あるいはカタカナ表記にしたものだった場合、エージェントの返答はオウム返しになってしまう。このような出力ではボケた返答になっていると言えないため、単語 B の次に最終的な Score が高い単語を新たな出力単語 B とする。単語 B を、ユーザが入力した単語 A と聞き間違えるという形で出力する。

6. 実験

6.1 実験目的

提案手法によるユーモアをユーザは面白いと感じるかを確かめる。また、提案手法は出力単語を選出するにあたって、編集距離・概念距離・出現頻度という 3 つの要素を参考にしている。これら 3 つの要素が面白さなどどのように関連しているのかも確かめる。

6.2 実験手順

本実験の被験者は 20 代の学生 (男性 6 名) である。順序効果を相殺するため、被験者は 4 パターンのシステムをランダムな順番で使用。このとき、システムの仕様を聞いたことによる先入観をなくすため、被験者には現在どの仕様のプログラムが使われているのかは説明しない。使用する 4 パターンのシステムは、ユーザの入力に対し下記のような返答をする。

- パターン 1
出現頻度 Score が一定以上の単語をランダムに選択し返す。本研究におけるベースライン方式とする。
- パターン 2
出現頻度 Score が一定以上の単語の中で、編集距離 Score が最も高い単語を返す。編集距離を重視しているシステムである。
- パターン 3
出現頻度 Score が一定以上の単語の中で、概念距離 Score が最も高い単語を返す。概念距離を重視しているシステムである。
- パターン 4
編集距離 Score, 概念距離 Score, 出現頻度 Score を総合的に加味した最終的な Score が最も高い単語を返す。本研究における提案手法を全て反映させたシステムである。

次に被験者にアンケート用紙を渡す。渡されたアンケート用紙に記載されている 20 個の単語を被験者自身が入力

表 5 各パターンの動作例

入力単語	パターン 1 出力例	パターン 2 出力例	パターン 3 出力例	パターン 4 出力例
寒中水泳 について教えて？	片耳？	寒中見舞い？	買い受ける？	環境経営？
ダイエット について教えて？	舵角？	ダイアップ？	第 44 回都市対抗野球大会？	大括弧？
体育館 について教えて？	ターンエー？	体育科？	たどる？	単位区間？

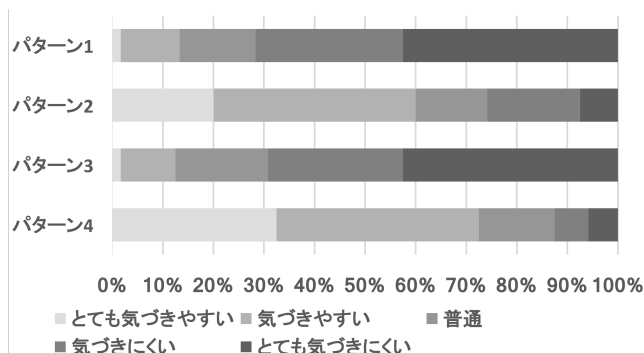


図 2 質問 A：被験者の回答 (N=6)

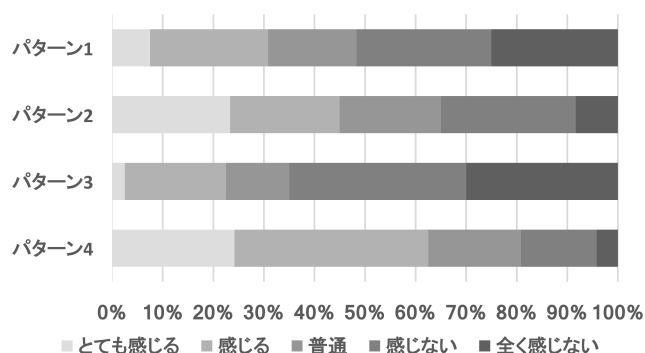


図 3 質問 B：被験者の回答 (N=6)

し、エージェントの出力を確認してもらう。入力“単語について教えて？”という形式にすることで、実際の対話に近づける。実験に用いる 20 の単語はあらかじめこちらで用意する。1 回の入出力ごとに被験者はアンケート用紙に記載されている 3 つの質問に 5 段階のリッカード尺度で答える。質問内容は次のとおりである。

- 質問 A. 出力された返答はボケているということに気づきましたか？
 - － 5. とても気づきやすい
 - － 4. 気づきやすい
 - － 3. 普通
 - － 2. 気づきにくい
 - － 1. とても気づきにくい
- 質問 B. 出力された返答はユーモアを感じましたか？
 - － 5. とても感じる
 - － 4. 感じる
 - － 3. 普通
 - － 2. 感じない
 - － 1. 全く感じない
- 質問 C. 出力された返答に意外性がありましたか？
 - － 5. とてもある
 - － 4. ある
 - － 3. 普通
 - － 2. ない
 - － 1. 全くない

6.3 結果と考察

パターン 2 とベースライン方式の結果に対し、Wilcoxon の符号順位和検定を行なうと、質問 A・質問 B において 5%水準で有意差を確認できた。比較結果より、ボケを作る際に編集距離 Score を参考にとすると、システムがボケてい

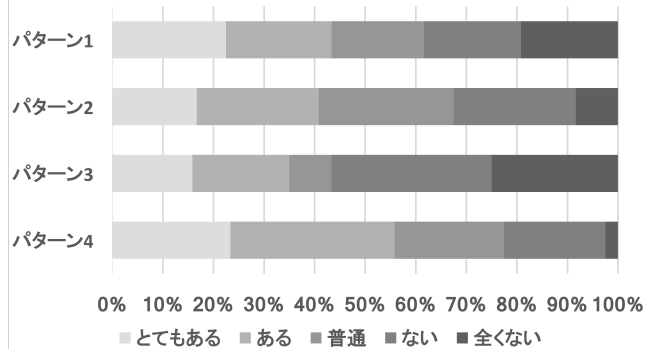


図 4 質問 C：被験者の回答 (N=6)

ることにユーザは気づきやすいということがわかった。これは、ランダムに選択した単語を出力単語としているベースライン方式に対し、パターン 2 では入力単語と出力単語に視聴覚的類似性が存在することが理由と考えられる。すなわち、入出力単語間での視聴覚的類似性が、不適合-解決モデルにおける解決部として機能したと考えられる。また、解決部単体ではあるが、ユーザはベースライン方式と比較してパターン 2 の方がよりユーモアを感じたとしている。このことから、システムがボケていることに気づきやすいと、ユーザはユーモアを感じやすいと考えられる。

パターン 3 とベースライン方式の結果に対し、Wilcoxon の符号順位和検定を行なうと、質問 B・質問 C において 5%水準で有意差を確認できた。しかし比較結果より、ボケを作る際に概念距離 Score を参考にするよりも、ランダムな単語を選択の方が意外性が高くなるということがわかった。これは、パターン 3 では入力単語と出力単語には概念距離が遠いという関連性があるのに対し、ベースライン方式の出力単語はランダムであるため、入力単語との関連性が全くないことが理由と考えられる。また、ユーザはパターン 3 と比較してベースライン方式の方がよりユーモ

アを感じたとしている。このことから、意外性とユーモアの関連性が窺える。

パターン4とベースライン方式、パターン4とパターン2、パターン4とパターン3の結果に対し、Wilcoxonの符号順位検定を行なうと、質問A・質問B・質問C、全ての質問において、どのパターンと比較した場合においても5%水準で有意差を確認できた。このことから編集距離 Score・概念距離 Score を同時に参考にしても、編集距離 Score を参考にして得られるボケていることへの気付きやすさは損なわれなと考えられる。また、出力単語をランダムに選定しているベースライン方式と比較して、パターン4は意外性が高いという結果になった。これはパターン4が編集距離 Score・概念距離 Score を同時に参考にし、不適合-解決モデルを満たす単語を出力したことが理由と考えられる。また、パターン3とベースライン方式の比較結果では、概念距離 Score を参考にしても出力単語の意外性が高くなるとは言えなかった。このことから、ユーザは感じた不適合を解決できたとき、より意外性を感じると考えられる。以上の比較結果より、どちらか一方の Score のみを参考に行っている場合よりも、編集距離 Score・概念距離 Score を同時に参考にした方が、ユーザはユーモアを感じたと言える。

またベースライン手法において、出現頻度 Score による出力候補単語の選定は、完全に機能しているとはいえなかった。被験者は、時折出力された単語に当惑する様子を見せるなど、その意味を正確に理解できていない場面があった。これは、設定した出現頻度 Score の制限の値が十分に高くなく、ユーザに理解できない単語を出力してしまったことを意味する。パターン2・パターン3においては、そのような様子はあまり見受けられなかった。これは入力単語の選定が人手によるものであったため、入力単語から計算で導き出される出力単語は、比較的ユーザに理解しやすい単語であったからだと考えられる。とはいえ全く見受けられなかった訳ではなく、自然なユーモアの表現において、ユーザに理解しやすい単語を出力することは大切だと言える。

以上より、本実験の結果から考察される事項は次のとおりである。

- ボケを作る際に編集距離 Score を参考にとすると、システムがボケていることにユーザは気付きやすい。
- システムがボケていることに気付きやすいと、ユーザはユーモアを感じやすい。
- ボケを作る際に概念距離 Score を参考にするよりも、ランダムな単語を選択の方が意外性が高くなる。
- 出力単語の意外性が高いと、ユーザはユーモアを感じやすい。
- 編集距離 Score・概念距離 Score を同時に参考にしても、編集距離 Score を参考にして得られるボケている

ことへの気付きやすさは損なわれない。

- 編集距離 Score・概念距離 Score を同時に参考にとすると、意外性が高くなる。
- ユーザは感じた不適合を解決できたとき、より出力単語に意外性を感じる。
- どちらか一方の Score のみを参考に行っている場合よりも、編集距離 Score・概念距離 Score を同時に参考にした方が、ユーザはユーモアを感じる。
- 自然なユーモアの表現において、ユーザに理解しやすい単語を出力することは大切である。

7. おわりに

本稿では普及した対話型エージェントとの無機質な対話に親しみを持たず、ユーザがエージェントを受け入れられないという問題の解消を目指し、ボケて返す対話型エージェントのプロトタイプシステムを提案した。ボケを作成する手法として、[13]のボケの作り方や、認知科学領域で支持されているモデルを参考にした。実験の結果、ユーザは提案手法によるボケを面白いと感じることがわかった。また、編集距離・概念距離・出現頻度という3つの要素を組み合わせることは、ユーモアを表現するにあたり有効であるとわかった。今後は提案システムの改善に向け、生成されるボケの面白さの向上を図りたい。この問題は最終的な Score 算出時の重み係数の値変更によって改善が期待される。また、単語のみならず文の入力に対してもボケて返すことができるエージェントの考案にも着手したい。文の入力に対するボケの作り方として、例えば文中の名詞を1つ選び、その単語についてボケを作るなどの手法が考えられる。あるいは、全ての名詞に対してボケなどの手法も考えられる。

本研究の期待される活用法として、ユーザとの良好なコミュニケーションがパフォーマンスや継続利用率の向上につながる場面での活用が期待される。例えば、NTT レゾナント社の教えて goo[16]にはユーザからの質問を AI が自動で応答する機能がある。このような質問掲示板で、AI がボケた返答をしてユーザに親しみを抱かせるといった活用法が考えられる。

参考文献

- [1] 井上宏:「笑い学」研究について, 笑い学研究, No. 9, pp. 3-15 (2002).
- [2] 伊勢崎隆司, 小林明美, 望月崇由, 山田智広: 笑い感情を誘起するロボットインタラクションの検討, 情報処理学会研究報告グループウェアとネットワークサービス (GN), Vol. 2017-GN-100, No. 7, pp. 1-5 (2017).
- [3] 真下遼, 梅谷智弘, 北村達也, 灘本明代: 文の感情を考慮した漫才ロボット台本自動生成手法の提案, DEIM Forum 2015 F4-4 (2015).
- [4] 中谷仁, 岡夏樹: ロボットの日常会話におけるユーモア生成の試み, 人工知能学会 2009 年全国大会論文集, 1J1-0s2-5

- (2009)
- [5] 吉田裕介, 萩原将文: 漫才形式の対話文自動生成システム, 日本感性工学会論文誌, Vol. 11, No. 2, pp. 265-272 (2012).
 - [6] 藤倉将平, 小川義人, 菊池英明: ユーモア発話の自動生成における単語間類似度導入によるユーモア受容性の向上, HAI シンポジウム 2014 (2014).
 - [7] Martin, R. A.: *The Psychology of Humor*, Elsevier Academic Press (2007).
 - [8] Coulson, S., & Williams, R. F.: Hemispheric Asymmetries and Joke Comprehension, *Neuropsychologia*, Vol.43, Issue 1, pp.128-141 (2005).
 - [9] Samson, A. C., Hempelmann, C. F., Huber, O., & Zysset, S.: Neural Substrates of Incongruity-Resolution and Nonsense Humor, *Neuropsychologia*, Vol.47, Issue 4, pp.1023-1033 (2009).
 - [10] Shultz, T. R.: The Role of Incongruity and Resolution in Children's Appreciation of Cartoon Humor, *Jnl. Experimental Child Psychology*, Vol.13, Issue 3, pp.456-477 (1972).
 - [11] Suls, J. M.: Cognitive Processes in Humor Appreciation, In *Handbook of Humor Research*, Vol.1: Basic issues, pp.39-57 (1983).
 - [12] 伊藤大幸: ユーモアの生起過程における論理的不適合及び構造的不適合の役割, *認知科学*, Vol. 17, No. 2, pp. 297-312 (2010).
 - [13] 織田正吉, 野村雅昭: シャレ・ダジャレ学事始 (ことはじめ)(第19回研究会), *笑い学研究*, No. 6, pp. 55-67 (1999).
 - [14] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/> (Last visited on 2017/4/1).
 - [15] Tomas Mikolov, Kai Chen, Greg Corrad, Jeff Dean: Efficient Estimation of Word Representations in Vector Space, In *Proceedings of Workshop at ICLR* (2013).
 - [16] 教えて goo, <https://oshiete.goo.ne.jp> (Last visited on 2017/4/1).