

文脈を考慮してボケるエージェントの実装と評価

長岡 大二¹ 中原 涼太¹ 呉 健朗² 鈴木 奨³ 宮田 章裕^{1,a)}

概要：現代社会において対話型エージェントは、我々の日常生活に浸透し始めており、今後も多くのユーザが利用すると考えられる。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は、このような対話型エージェントを受け入れられないという問題が懸念される。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してきた。しかし、従来手法では、ボケる際に文の文脈を考慮せずに置換単語を決定しているため、文全体で見た際にユーモア性が損なわれる可能性がある。そこで我々は、エージェントがボケる際に、文の文脈を考慮することで、ユーモア性を保てると考えた。この考えに基づき、文脈を考慮してボケる対話型エージェントを提案する。これは、文の概念を代表する単語を、音が近く、文の概念から意味が離れた単語に置換するというものである。本稿では、エージェントが文脈を考慮したボケを行うプロトタイプシステムを構築し、それをを用いた検証実験を行った。

Implementation and Evaluation of a Conversational Agent Replying with a Context-sensitive Joke

Daiji Nagaoka¹ Ryota Nakahara¹ Kenro Go² Sho Suzuki³ 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。しかしその一方で、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も一定数いると考えられる。そのような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。このような問題を解決するためには、エージェントとユーザの間に親和的な関係を築くことが必要であると言える。[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、ユーザは親しみを感ずることができるのではないか、という仮説をたて、それに基づき、[2], [3], [4]を提案してきた。しかし従来手法では、置換元である単語のみに着目して置換単語を決定し、置換する手法であったため、文脈を考慮できておらず、文全体で見たときユーモア性が損なわれる可能性がある。そのため、文の文脈を考慮して置換単語を決めることで、文全体のユーモア性を向上させられると考えられる。

本稿の貢献は次の通りである。

- ユーモアがある文脈を生成する対話型エージェントのシステムを提案したこと。
- 上記提案のプロトタイプシステムを構築し、ユーザ実験を行って有効性を検証したこと。

2. 関連研究

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、[5], [6], [7], [8]

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

² 日本大学大学院総合基礎科学研究科
Graduate School of Integrated Basic Sciences, Nihon University

³ 2018年日本大学文理学部卒業
Graduated from College of Humanities and Sciences, Nihon University in 2018

a) miyata.akihiro@acm.org

が挙げられる。[5]では、エージェン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いる。ユーザの笑い感情が誘起することがセンサデータから確認された会話文をデータセットとして、機械学習のアプローチに基づいてユーザの笑い感情を誘起するロボットインタラクションモデルの生成を行っている。[6]では、複合語と複合語を構成する各語の類音語を見出しとする辞書を用意し、複合語の前半部、もしくは後半部を類音語に置換することで駄洒落を生成している。[7]は、漫才形式の対話文の自動生成システムである。入力した文章から単語を選び、それを置換することでボケを生成している。またその後、ボケに対する否定・訂正するツッコミ文を生成することで、漫才形式を実現している。[8]では、Web ニュース記事本文の感情情報を抽出して感情を考慮し、おかしみの構造図に基づいた形式的なボケによる漫才台本自動生成手法を提案している。

ユーザと対話を行うエージェントとして [9], [10], [11], [12] が挙げられる。[9]は、ユーザの対話継続欲求を向上させるために対話システムによるユーモア発話の自動生成手法を提案している。ユーモア発話の自動生成には単語間類似度を用いている。[10], [11], [12]では、ユーザからの特定の入力に対して、あらかじめ設定された、ユーモアを含む応答文を返していると思われる。

3. 研究課題

日常生活で多く見られるようになった対話型エージェントは、今後も幅広い分野での活躍が期待される。しかしエージェントは、未だその対話の多くが無機質なものである。そのため、無機質な対話に親しみを持ってないユーザには、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 [5][6][7][8]。しかし、これらはエージェントが一方的に話すことでユーモアを表現しており、ユーザと対話は行っていない。[9]は対話中でのユーモアの表現を試みているが、エージェントが突飛な発言を行うことがあり、ユーザがエージェントの発話をユーモアとして受容しにくい場合がある。[9], [10], [11], [12]については、ユーモアを表現する場合、あらかじめユーザが入力する文と、それに対応する応答文が決められていると思われる。そのため、ユーザが入力できる文の自由度が低く、すぐに飽きられてしまう可能性がある。

これらの問題を解消するために、我々はエージェントの発言にユーモアを持たせることで、エージェントに対して親しみを持たせる研究を行ってきた [2][3][4]。しかし、従来の提案ではボケ単語を生成する際に、置換元である単語とのみ概念距離を離していたため、文との概念距離が離れていない場合がある。そのため、ユーモア性の獲得に必要

な不適合の生成ができておらず、ある単語を置換した後の、文全体としての面白さが損なわれる可能性がある。例えば、[2], [3], [4]では、「機械を操作する」という文章は、「機械を調査する」という文章に変換される。この場合、「操作」という単語が、音が近く、意味が遠く、聞き慣れた「調査」という単語に置換されている。しかし、文全体で見たときに意味が通じてしまっているため、不適合が生成できておらず、ユーモア性が損なわれている可能性がある。

以上のことから、我々はユーモアがある文脈を生成する対話型エージェントの実現を研究課題とする。

4. 提案方式

ボケを生成するために我々は、認知科学研究者の多くがユーモア生起に不可欠と主張している「不適合の認知」に着目している [13]。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 [14][15][16][17][18]。「不適合-解決モデル」とは、通常はまったく異種であり、関連がないと思われる思考や状況(不適合)が、納得できる方法で結び合わされる(解決される)ことでユーモアを生起させるとする理論である。3章の研究課題に対し、我々は、ある1つの文章を構成する単語群内では、共通する1つの概念が存在するのではないかという仮説をたてた。この仮説に基づき、文の概念を代表する単語を、音が近く、文の概念から意味が離れた単語に置換することを考えた。上記の検討に基づき、文全体を通してユーモア性が保てるように、文脈を考慮してボケの対話型エージェントを提案する。これにより、文全体を見た際でも不適合の生成をすることができるようになり、ユーモア性を維持できると考えられる。

提案方式は次の3つの手順からなる。まず、手順1として、ある1つの文を構成する単語群内で、共通する概念を検出する。次に、手順2として、検出された概念に最も意味が近い単語を、文中から選定する。そして、手順3として、選定された単語を、手順1で検出された概念から最も意味が離れた単語に置換する。上記3つの手順により生成された不適合の解決については、従来手法と同様に置換する単語を、置換元単語の音と近く、聞き慣れている単語にすることで実現する。

5. 実装

5.1 事前準備

Wikipedia 全文記事を形態素解析し、不要品詞などを除去して分かち書きしたものをコーパスとする。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、非自立語、代名詞、接尾、数、サ変・スルと分類されるものを指す。作成したコーパスから、MeCab[19]を用いて読み方辞書を、word2vec[20]を用いて言語モデルを作成する。読み方辞書とは、コーパス内に出

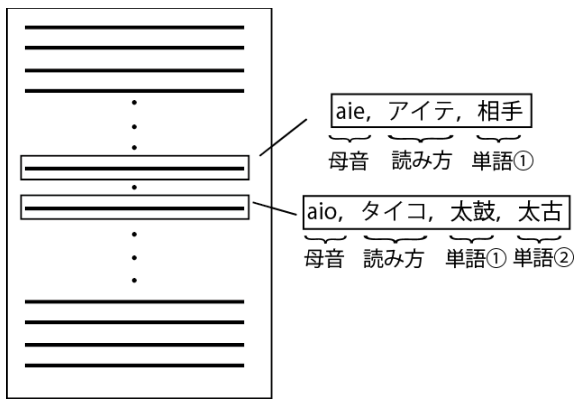


図 1 読み方辞書

現する単語の標準形，標準形の読み方（カタカナ），標準形の読み方をローマ字表記に変換したものから子音を削除し，母音のみにしたもののリストである（図 1）。

5.2 文脈を考慮した置換元単語の決定

文脈を考慮するためには，文の概念を代表する単語を置換する必要がある。4章で述べた，手順 1 と手順 2 は次のように行われる。入力文に対して，MeCab を利用して，文中の全名詞で構成される置換元候補群を作成する。その後，置換元候補群内のすべての単語の s_r (Replace score) を算出し，最も高い s_r を保持する単語を，置換元単語とする。このとき，名詞，動詞，形容詞は，文を構成する上で重要だと思われる。そのため，置換元候補群の各名詞について，それぞれ文中の，自身を除いた名詞，動詞，形容詞との概念距離を計算し，その距離が近いほど高い s_r を与える（図 2）。 s_r は下記のように計算される。

$$s_r(i) = \frac{1}{n-1} \sum_{k=1, k \neq i}^n \frac{1}{1+d_s(k)}$$

n は文中のすべての名詞，動詞，形容詞の数であり， i は s_r を算出する名詞の位置である。 $d_s(k)$ は k 番目の単語と s_r を算出する名詞との概念距離である。概念距離の計算には word2vec を利用して作成した言語モデルを用いる。

5.3 文脈を考慮した置換単語の決定

本節では，置換候補単語それぞれの各 Score (s_s, s_e, s_f) の計算方法について述べる。置換候補単語とは，下記の条件を満たすコーパス内の単語である。

- 頭文字の母音と末尾の文字が置換元単語と一致する。
- コーパス内での出現回数が 1000 回以上である。
- MeCab で人名・地域と判定されない。
- 日本語である。
- 置換元単語を，ひらがな・カタカナ表記にしたものではない。

5.3.1 s_s : 概念距離 Score

置換候補単語と，文中のすべての名詞，動詞，形容詞との

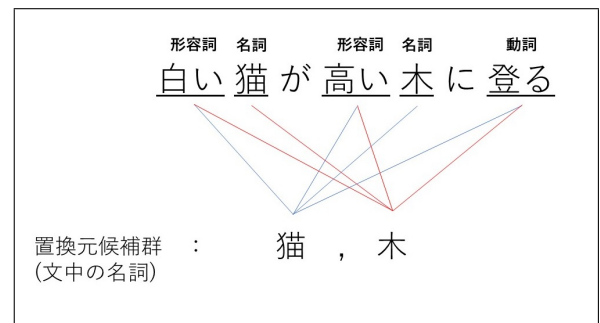


図 2 置換元単語の選定時の概念距離計算

概念距離を計算し，その距離が遠いほど高い s_s (Semantic score) を与える。 s_s は下記のように計算される。

$$s_s = \frac{1}{n} \sum_{k=1}^n d_s(k)$$

n は文中の名詞，動詞，形容詞の数の合計値であり， $d_s(k)$ は k 番目の単語と置換候補単語との概念距離である。

5.3.2 s_e : 編集距離 Score

置換候補単語と置換元単語の読み方の編集距離^{*1}を計算し，その距離が近いほど高い s_e (Edit distance score) を与える。 s_e は下記のように計算される。

$$s_e = \frac{1}{1+d_e}$$

d_e は置換候補単語と置換元単語の編集距離である。

5.3.3 s_f : 出現頻度 Score

Wikipedia コーパス内での置換候補単語の出現回数が多いほど高い s_f (Frequency score) を与える。 s_f は下記のように計算される。

$$s_f = \log f$$

f は置換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。

5.4 出力文の決定

算出した 3 つの Score (s_s, s_e, s_f) を，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，置換候補単語ごとの最終的な s_t (Total Score) とする。すなわち s_t は下記ようになる。

$$s_t = w_s s_s + w_e s_e + w_f s_f$$

w_s, w_e, w_f は重み係数である（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

^{*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。

表 1 被験者への質問一覧

Q1	出力された返答にユーモアを感じましたか？ 5: とても感じる ~ 1: 全く感じない
Q2	エージェントのボケに納得感は得られましたか？ 5: とても納得できた ~ 1: 全く納得できなかった
Q3	置換元の単語は文のトピックを代表する単語だと思いますか？ 5: とても思う ~ 1: 全く思わない
Q4	文の内容から置換後の単語の意味が離れていると感じましたか？ 5: とても感じる ~ 1: 全く感じない
Q5	エージェントに親しみを感じましたか？ 5: とても感じる ~ 1: 全く感じない

保持する置換候補単語を置換単語とし、文中の置換元単語と置換する。文中の置換元単語を置換した文を、出力文とし、出力する。

6. 検証実験

6.1 実験の目的

従来手法と比較して、文脈を考慮することでユーモア性が向上するかどうかを確認するため、下記の2点について検証する。

- 置換元の単語を、文の概念を代表する単語にすることで、よりユーモア性が向上するか。
- 置換後の単語を、文の概念から意味が離れた単語にすることで、よりユーモア性が向上するか。

6.2 実験条件

本実験の被験者は20代の学生12名(男10, 女2)である。実験は下記のような、Baselineと、置換元の単語・置換後の単語を選定する方法が異なる3つの提案方式(P1, P2, P3)を用いて行う。

- Baseline: 文の最初の単語を、その単語との s_s , s_e , s_f を合算した値が最も高くなる置換候補単語に置き換え、出力文とする。
- P1: 文の最初の単語を、その単語との s_e , s_f と、 s_i との s_s を合算した値が最も高くなる置換候補単語に置き換え、出力文とする。
- P2: 文の概念を代表する単語を、その単語との s_s , s_e , s_f を合算した値が最も高くなる置換候補単語に置き換え、出力文とする。
- P3: 文の概念を代表する単語を、その単語との s_e , s_f と、 s_i との s_s を合算した値が最も高くなる置換候補単語に置き換え、出力文とする。

6.3 実験手順

被験者には6.2節に示した4つのパターンをそれぞれ使用してもらい、アンケートに5段階のリッカート尺度で回答してもらい、被験者への質問と選択肢を表1に示す。被験者には、各パターンに対してそれぞれ、あらかじめ実

Agents Chatroom

何が知りたいの？

自動車って何？

まかせて！

自動車とは、原動機の動力によって【家田】を回転させ、軌条や架線を用いずに路上を走る車のことだよ

あ、「車輪」か

・・・

他に知りたいことはある？

あるよ

何が知りたいの？

図 3 実験に用いた Web アプリケーション

験者が用意した10個の単語を Web アプリケーション(図3)を用いて入力し、出力を確認してもらう。このとき使用した Web アプリケーションは、ユーザとエージェントがチャット形式で対話をするものである。また、用意した単語については、下記の条件を満たす単語からランダムに選ばれたものを使用する。

- コーパス内の出現回数が1000回以上である。
- その単語の Wikipedia 記事ページが存在し、入力単語から始まる概要文が書かれている。
- 概要文が、ユーザからの単語の意味を問う質問に対する回答として成立している文章になっている。

被験者が単語を入力する際には、実際のエージェントとの対話を行うシーンに近づけるために、「“単語”って何？」と入力してもらう。また、各パターンでは、その単語の Wikipedia 記事ページにある概要文の始めの一文をもとに、ユーモアを生成する。被験者には1回の入出力が終わるごとに Q1~Q4 に回答してもらう。Q5 については、各方式に対して10回の入出力を終えるたびに回答してもらう。このとき、順序効果を相殺するために、被験者ごとに使用するパターンの順番をランダムに入れ替える。さらに、被験者が変わるごとに、10個の単語をランダムに選び直し、被験者に入力してもらう各単語が実験結果に与える影響を少なくする。システムの仕様を聞いたことによって生じる先入観をなくすため、各パターンの出力文の生成方法は被験者には説明しない。

6.4 結果と考察

Q1 の回答結果を図4に示す。“出力された返答にユーモアを感じたか”という質問に対し、“とても感じた”または“感じた”と回答した被験者は、Baseline では36%、P1では32%、P2では37%、P3では29%であった。Q2 の回答結果を図5に示す。“エージェントのボケに納得感は得られましたか”という質問に対し、“とても得られた”または“得られた”と回答した被験者は、Baseline では38%、P1では33%、P2では34%、P3では32%であった。上記2点の回答結果から、必ずしも文脈を考慮することがボ

ケの納得感やユーモア性の向上につながるとは限らないことが示された。しかし、どのパターンもユーモアの感じやすさは一定の水準にとどまっており、改善の方法次第では今後向上する余地があると考えられる。Q3の回答結果を図6に示す。“置換元の単語は文のトピックを代表する単語だと思いますか”という質問に対し、“とても思う”または“思う”と回答した被験者は、Baselineでは62%、P1では59%、P2では52%、P3では51%であった。Q4の回答結果を図7に示す。“文の内容から置換後の単語の意味が離れていると感じましたか”という質問に対し、“とても感じた”または“感じた”と回答した被験者は、Baselineでは79%、P1では75%、P2では70%、P3では65%であった。Q5の回答結果を図8に示す。“エージェントに親しみを感じましたか”という質問に対し、“とても感じた”または“感じた”と回答した被験者は、Baselineでは41%、P1では41%、P2では59%、P3では41%であった。このことから、文脈を考慮しても親しみやすさに影響はないと考えられる。

Q1からQ5の質問に対して、それぞれのパターンを比較したが、有意差が見られなかった。今回の実験において、我々が設定した実験条件が適切ではなかったため、Baselineと提案方式で出力結果が同じになってしまう場合が多く見られた。そのため、被験者にとって各パターンの違いが分からず、実験結果に差が表れなかったと考えられる。出力結果が同じになってしまった原因として、次の2点が考えられる。1つ目は、Baselineと提案方式における置換元単語の位置に、差がない場合が多かったことである。これは、今回の実験でユーモアを生成する文をWikipedia記事ページの概要文としており、概要文の多くは先頭の単語を説明する文章になっているため、先頭の単語が文のトピックになることが多かったからだと考えられる。そのため、文の先頭にある単語が、その文の概念を表す単語となり、提案方式において置換元単語を選定する際に、文の先頭の単語が選ばれてしまったと考えられる。2つ目は、Baselineと提案方式で、出力される置換後単語が同じになってしまう場合が多かったことである。提案方式では文脈を考慮するために、文中の名詞、動詞、形容詞の限られた品詞のみを利用していた。そのため、提案方式の置換後単語を選定する処理において、十分に文脈を考慮出来ておらず、Baselineと同じ置換後単語が選ばれてしまったと考えられる。この問題は、利用する品詞の組み合わせや、文章の係り受け関係に着目することで、解決できると考えられる。

7. おわりに

本稿は、[2], [3], [4]のシステムに見られる、文全体でのユーモア性の喪失という問題の解消を狙ったものである。我々は、ボケる際に文脈を考慮することで、ユーモア性を保てるを考え、文中の名詞の中から、文の内容と関連性の

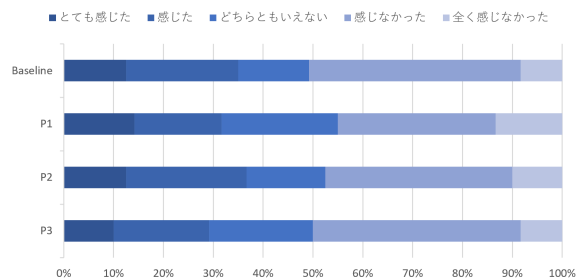


図4 Q1の回答 (N=12, 1人あたり10回答)

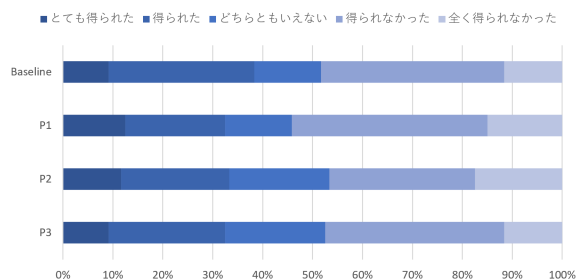


図5 Q2の回答 (N=12, 1人あたり10回答)

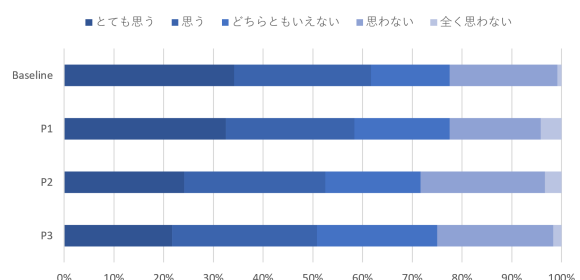


図6 Q3の回答 (N=12, 1人あたり10回答)

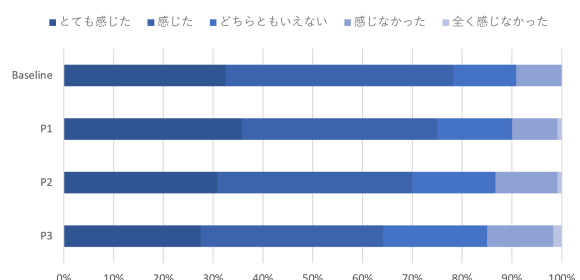


図7 Q4の回答 (N=12, 1人あたり10回答)

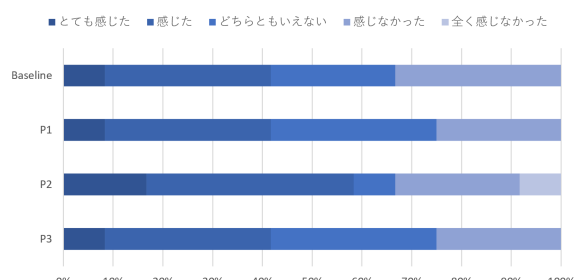


図8 Q5の回答 (N=12)

高い単語を選び、文の内容と意味が離れた単語に置き換えるシステムを提案した。またそのプロトタイプシステムを構築し、従来手法と比較してユーモア性が向上するかどうかの実験を行った。実験の結果、現時点では、提案方式によるユーモア性の向上は見られなかった。今後の課題としては、適切な条件で実験を行うことと、文脈を考慮する際のアプローチの仕方を変えることなどが挙げられる。

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] 鈴木奨, 呉健朗, 瀧田航平, 堀越和, 中辻真, 宮田章裕: ボケて返す対話型エージェントの基礎検討, 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-102, No.3, pp.1-6 (2017).
- [3] 呉健朗, 鈴木奨, 瀧田航平, 中原涼太, 中辻真, 宮田章裕: ボケて返す対話型エージェントの実装, 日本バーチャルリアリティ学会 VR 学研報, Vol.22, No.CS-4, pp.15-18 (2017).
- [4] 呉健朗, 鈴木奨, 瀧田航平, 堀越和, 中辻真, 宮田章裕: 対話型エージェントにおけるボケる返答機能の実装, インタラクシオン 2018 インタラクティブセッション (2018).
- [5] 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オンの検討, 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).
- [6] ビンステッドキム, 滝澤修: 日本語駄洒落なぞなぞ生成システム”BOKE”, 人工知能学会誌, Vol.13, No.6, pp920-927 (1998).
- [7] 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).
- [8] 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 Forum2015 F4-4 (2015).
- [9] 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).
- [10] Apple Inc.: Siri, <https://www.apple.com/ios/siri/> (Last visited on 2018/05/09).
- [11] SoftBank.: Papper, <https://www.softbank.jp/robot/> (Last visited on 2018/05/09).
- [12] Google Inc.: Google Assistant, <https://assistant.google.com/> (Last visited on 2018/05/09).
- [13] Martin: The Psychology of Humor, Elsevier Academic Press (2007).
- [14] S. Coulson and R. F. Williams.: Hemispheric Asymmetries and Joke Comprehension, *Neuropsychologia*, Vol.43, Issue1, pp.128-141 (2005).
- [15] A. C. Samson, O. Hempelmann, C. F. Huber, and S. Zysset.: Neural Substrates of Incongruity-Resolution and Nonsense Humor, *Neuropsychologia*, Vol.47, Issue 4, pp.1023-1033 (2009).
- [16] Shultz: 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).
- [17] Suls: Cognitive Processes in Humor Appreciation, In *Handbook of Humor Research*, Vol.1: Basic issues, pp.39-57 (1983).
- [18] 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).
- [19] MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://taku910.github.io/mecab/> (2018-04-24).
- [20] Tomas Mikolov, Kai Chen, Greg Corrad, Jefrey Dean: Efficient Estimation of Word Representations in Vector Space, In *Proceedings of Workshop at ICLR* (2013).