

笑い感情を誘起するロボットインタラクションの検討

伊勢崎 隆司^{1,a)} 小林 明美^{1,b)} 望月 崇由^{1,c)} 山田 智広^{1,d)}

概要：“笑う”活動は精神的、身体的にも有益であることが近年の研究により明らかとなってきた。笑い感情を誘起するインタラクション技術への取り組みは次第に盛んになってきているが、臨床上にユーザ実験を行い、得られた知見を集約するアプローチが多く、人手がかかる問題がある。本研究では、ユーザの状態に基いて笑い感情を誘起するインタラクションを自動的に更新するアプローチを採用。本稿では、笑い感情が誘起される入力文章と出力文章のデータセットからモデルを構築し、回答を生成する技術について述べ、ユーザ評価の結果を元に考察する。

Investigation of robot interaction to induce laugh emotion

TAKASHI ISEZAKI^{1,a)} AKEMI KOBAYASHI^{1,b)} TAKAYOSHI MOCHIZUKI^{1,c)} TOMOHIRO YAMADA^{1,d)}

1. はじめに

“笑う”という活動はリラックス状態の際に発生する活動であるため、精神的な効果が非常に大きいことがわかっている [1]。さらに、笑う際には表情筋や腹筋の活動が活性化される。これら筋肉の活動に伴い、筋力低下の防止や血流促進に伴う脳の活性化など、精神面だけでなく身体面への影響も大きく、日常的に笑う活動を行うことが重要であることがわかってきている。これまで人は人とのインタラクションを通じて様々なシーンにおいて笑う活動を行ってきた。今後は日常生活への普及が期待されるロボットとのインタラクションの機会が増えることが想定される。それに従い、ロボットとのインタラクションを通じた笑いの活動も増えることが想定される。

NTTでは、ロボットを用いたインタラクションを実現する完全クラウド対応型デバイス連携制御技術 R-env:連舞TMの研究開発を進めている。上記技術を使うことで、ユーザは各社から発売される様々なロボットをインターネットにつなぐことで、思いどおりのインタラクションをプロ

グラミングすることができる。ロボットを用いたインタラクションのプログラミングは容易になってきている一方、ユーザの笑いを引き起こすような高次のインタラクションをプログラミングするには、プログラミングに関する知見だけでなく、心理的な知見やユーザの嗜好などの情報を多面的に捉えて実装する高次のスキルが必要となる。

本研究では、日常生活への今後の普及が期待されるロボットを通じてユーザの笑い感情を誘起するインタラクションの実現を目指す。本取り組みで創出される技術は、ロボットのインタラクションをプログラミングする際に、ユーザの感情を変化させることが可能な高次の技術として利用可能である。ロボットのインタラクションにおいては、人とロボットがインタラクションを行う形態とロボット同士のインタラクションを人が視聴する形態がある。本研究では技術検討を進めるにあたり、複雑性を低めるために人が介在しない系として、ロボット同士のインタラクションを人が視聴する形態について考える。また笑い感情を誘起する視聴コンテンツとして、笑い感情の誘起が比較的容易に行えると想定される大喜利に着目して検討を深める。

2. 従来研究

ユーザの笑い感情を誘起するインタラクション技術については、これまでに様々な検討がなされてきた。ロボットやエージェントが発話する文章の生成についての技術の取

¹ 日本電信電話株式会社 NTT サービスエボリューション研究所
NTT Service Evolution Laboratories, NTT Corporation

a) izezaki.takashi@lab.ntt.co.jp

b) kobayashi.akemi@lab.ntt.co.jp

c) mochizuki.takayoshi@lab.ntt.co.jp

d) yamada.tomohiro@lab.ntt.co.jp

り組みが盛んに行われている。人の笑いを誘起する一つの語法として「なぞかけ」があるが、前田らは“なぞかけ”の生成を単語の意味的な関連性と音韻的類似性に基いて行っている [2]。2者間のインタラクションを通じて笑い感情を誘起する語法としては“漫才”がある。

灘本は、Web上のニュースの記事を用いて漫才台本を自動的に生成する手法について取り組んでいる [3]。Web上の記事を用いることで、最新の情報を用いることが可能であり、時が経つことで、異なる漫才台本が生成可能である。吉田らは、漫才形式の対話文を自動生成する技術に取り組んでいる [4]。一連の文章からなるストーリーに対して、対象単語に対して隣接する単語との共起が最大となる単語を Google N-gram [5] を用いて抽出し、置換することで聴衆に特定の違和感を生じさせ、笑い感情の誘起を実現している。

発話文章の他に、ロボットを用いた人の笑い誘起に関する検討の取り組みもある。林らは、ユーザとロボットの様々なインタラクション形態の中から、ロボット同士の漫才をユーザが視聴する形態を社会的受動メディアと定義し、社会的受動メディアとしてのロボット同士のインタラクションの有効性について検討を行っている [6]。被験者実験の結果、臨場感や漫才全体の評価については人の漫才映像と比較してロボット同士の漫才の方が有意に高いという結果を得ている。白井らは、人とロボットの漫才をテーマとしてワークショップを開催し、伊藤らの不適合理論 [7] の観点からロボットのプログラミングや人との協調動作についての検討を行っている [8]。

従来研究では臨床的にユーザ実験を通じて得られた知見を元にユーザの笑い感情を誘起するインタラクションのモデルを定義するアプローチが多い。ユーザ実験を通じて知見を集約してインタラクションモデルとして構築するアプローチでは、ユーザの行動観察やロボットの動作プログラミングから笑い感情を誘起する要因を分析するプロトコルであるため、専門的なスキルを保有する実験者と様々な実験データを必要とする。結果として多くの人手がかかるため、集約できる知見の量に限界がある。また、ユーザの笑い感情を誘起するインタラクションの定義するモデルの粒度についても検討が必要である。年齢や性別、嗜好といったユーザの様々な属性に応じて臨機応変に発話内容や動作を選択する必要があることが想定される。したがって、従来研究のように発話や動作の内容やタイミングについてだけでなく、ユーザの属性に応じてルールを定義する場合は、定義すべきルールの数が膨大になり、現実性が低い可能性がある。

3. 提案手法

本研究は、センサデータを用いてユーザの笑い感情が誘起することが確認された会話文をデータセットとして、機械学習的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目指す。機械学習的アプローチを採ることで定義すべきルールの数に依存せずにインタラクションモデルを生成できる。また、ユーザの属性をモデルへの入力として組み込むことで、詳細なユーザ適応が可能となり、高精度に笑い感情を誘起することが可能であると考えられる。機械学習的アプローチの重要なポイントとしては、大規模な学習データを用意することや、学習データを自動的に収集できる仕組みを設計することがある。すなわち、本提案手法の研究課題としては、下記の二点がある。

- 会話文をデータセットとした笑い感情を誘起するインタラクションモデルの生成
- センサデータに基づくラベリング済み学習データの収集

本稿では、大喜利をデータセットとした笑い感情を誘起するインタラクションモデルの生成について検討する。また、インタラクションモデルの出力としてはロボットの発話と動作の2種類があるが、本稿では簡略化のため発話のみに着目する。

3.1 笑い感情を誘起するインタラクションモデルの生成

入力文に対してユーザの笑い感情を誘起する出力文を生成する方法について述べる。本提案手法は学習フェーズと出力フェーズの2つがある。学習フェーズにおいては、入力文と出力文が組になっている入力文リスト *InputSentenceList* と出力文リスト *OutputSentenceList* のデータセットを用いる。

InputSentenceList と *OutputSentenceList* は N 個の文を含んでいる。

$$\mathit{InputSentenceList} = [S_{in_1}, S_{in_2}, \dots, S_{in_N}] \quad (1)$$

$$\mathit{OutputSentenceList} = [S_{out_1}, S_{out_2}, \dots, S_{out_N}] \quad (2)$$

S_{in_i} と S_{out_i} は文ベクトルであり、それぞれ *InputSentenceList* と *OutputSentenceList* に含まれる。文章は複数単語の連なりによって構成される。したがって、 S_{in_i} と S_{out_i} はそれぞれ複数の単語によって構成されている。したがって、各単語の意味や単語間の関係性をコンピュータ上で計算できるように、あらかじめ単語空間を構築する。各単語を構築した単語空間に射影し、単語ベクトル w を得る。文章ベクトルは複数の単語ベクトルを保持するため、 n 個の単語を保持する文章ベクトル S を下式のように記述する。

$$S = [w_0, w_1, \dots, w_n] \quad (3)$$

規定単語数を定義し、各文章ベクトル S の保持する単語数が規定単語数に満たない場合は 0 パディングを通じて各文

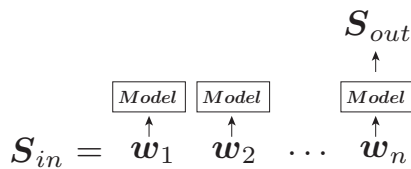


図 1 LSTM の入出力

Fig. 1 Input and output for LSTM learning



図 2 実験風景

Fig. 2 Experimental scene

章ベクトルの単語数を統一する。単語数が統一された文章ベクトル *InputSentenceList* と *OutputSentenceList* を用いて、下式が成立するようなモデルを機械学習にて実装する。

$$Model(InputSentenceList) = OutputSentenceList \quad (4)$$

機械学習方法の種類としては SVM や RandomForest, 深層 Neural Network 等様々なものがあるが, 文章などの一定の長さを持つ時系列データに対して有効である LSTM(Long Short-Term Memory) を用いることとする。図 1 に示すように, LSTM のモデルに対して入力文章 S_{in} を 1 単語ずつ入力し, 最後の単語を入力した際に出力単語群が連結した出力文章 S_{out} を出力する。

出力フェーズにおいては, 入力文章 S に対して, 各単語を単語空間に射影した単語ベクトルをリストとして保持し, 規定単語数を満たすように 0 パディングを行う。学習フェーズにおいて生成したモデルに対して S を 1 単語ずつ入力し, 末尾の単語を入力した際の出力単語群を出力文章として出力する。

4. 実験

本提案手法で創出した出力文章がユーザの笑い感情を誘起できているかを検証するために実験を行った。

4.1 実験システム

図 2 に示すように, 2 台の Sota (ヴイストーン社製, 280(H) × 140(W) × 160(D)mm)[9] を用いて大喜利の環境を模擬した。Sota は全身 8 自由度 (胴体 1 軸, 腕 2 軸 × 2, 首 3

軸) をもつプログラム可能なロボットであり, 姿勢, ジェスチャ, 音声による表現が可能である。左側の Sota が出題者, 右側の Sota が回答者という構成となっている。それぞれのロボットはネットワークで接続し, NTT が研究開発を行っている完全クラウド対応型デバイス連携制御技術 R-env:連舞TM を用いて発話と動作の実装を行った。

本提案手法は文章を構成する単語をベクトルで表現するためにあらかじめ単語空間を保持しておく必要がある。本実験では, 一般的な日本語の意味空間を構築することを目的として, 単語の意味や単語間の Wikipedia の本文データ (2.48GB) に対して, オープンソース形態素解析エンジン MeCab を用いて単語の分かち書きを行い, オープンソース Word2vec を用いて単語空間を構築した。Word2vec は 2 層から成るニューラルネットワークで, コーパスを入力として単語の特徴量ベクトルを出力する。

提示する大喜利の内容については, 任意の入力文に対して高精度に単語を出力することができなかつたため, あらかじめ笑いが誘起できそうな出力が得られる入力文章セットを用意してプログラムした。なお, 任意の入力文に対して笑いを誘起する単語が出力できていない理由としては, 提案手法の学習データが充分でないことが考えられる。なお, 本提案手法は各単語の出力までは実現できているが, 文章を構成する機能までは未実装である。したがって, 単語出力を提案手法により行い, 文章構築は筆者らが行った。

本提案手法が笑い感情を誘起できることを評価するための一つのアプローチとして, 人が作成した出力文章との比較を行った。人が作成する出力文章については, 筆者らが同じ入力文章に対して笑い感情を誘起されることが想定される出力文章を作成した。入力文章としては下記の 5 つを用いた。

- Q1 こんなガメラ嫌だ。どんなガメラ?
- Q2 こんなポケモン嫌だ。どんなポケモン?
- Q3 こんな大喜利嫌だ。どんな大喜利?
- Q4 こんな飴玉嫌だ。どんな飴玉?
- Q5 こんなティッシュ嫌だ。どんなティッシュ?

上記に対して, 提案手法を用いて構築した出力文章が下記となる。

- A1 一部, ビニール
- A2 セーブしかできない
- A3 内容が暗い
- A4 泥土
- A5 つながっている

筆者らが作成した出力文章が下記となる。

- A1 声がかわいい
- A2 メタモンしか出ない
- A3 お題がバイナリデータ
- A4 唾液味
- A5 湿っている

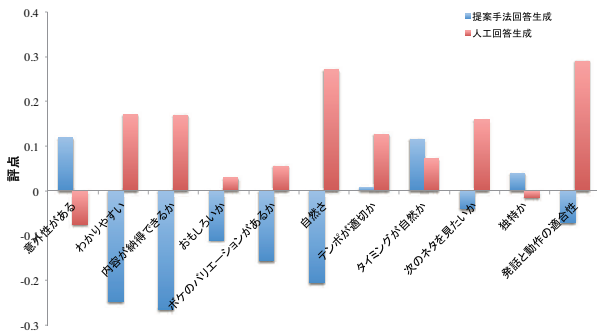


図 3 実験結果

Fig. 3 Experimental result

4.2 実験の手順

本実験では、6名の被験者（男性5人、女性1人）を用いた。提案手法による回答と人工回答の2パターンについて順番に視聴し、各出力文章を視聴した際の印象をアンケートにて取得した。2パターンの順序効果を考慮し、3人は提案手法の回答を視聴した後に人工回答を視聴し、残りの3人は逆の順番で視聴した。アンケートについては、下記項目を7段階（-3～3）にて取得した。

- 意外性がある
- わかりやすい
- 内容が納得できるか
- おもしろいか
- ボケのバリエーションがあるか
- 自然さ
- テンポが適切か
- タイミングが自然か
- 次のネタを見たいか
- 独特か
- 発話と動作の適合性

提案手法による回答5つと人工回答5つのそれぞれに対してアンケートを答えたため、被験者一人に対して10個のアンケートデータを得た。

5. 結果と考察

各被験者から得られたアンケートについて、各項目に対して全被験者の回答を平均した。実験の結果を図3に示す。本提案手法が人工回答に対して高い評点となった項目は、「意外性がある」、「タイミングが自然か」、「独特か」であった。一方で、人工回答の方が高い評点となった項目は「わかりやすい」、「内容が納得できるか」、「おもしろいか」、「自然さ」等が挙げられた。意外性や独特性は従来研究でも評価として用いられている指標であるため、笑いを誘起するためには重要なポイントであると考えられる。しかしながら、本提案手法から得られた出力文章は意外性や独特性が高く、結果としてわかりやすさや納得性、そして面白さに対して否定的な面を持ってしまったと考えられる。すなわち、意

外性や独特性を一定に保ちつつ納得性の高い文章を出力することが求められると考えられる。このような結果が得られた原因の一つとして、用いた学習データが意外性・独特性が高すぎる事が考えられる。今後の検討として、学習で用いたデータに対して本実験で用いたアンケート項目の観点で評価し、バランスよく学習させることでより納得性、面白さを向上できる可能性がある。

6. おわりに

本稿では、ユーザの笑いを誘起するロボットインタラクション技術の確立を目指し、センサデータを用いてユーザの笑い感情が誘起することが確認された会話文をデータセットとして、機械学習的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た。提案手法で生成された出力文章と人工的に作成した文章に対する印象評価実験を行った結果、提案手法は意外性や独特性が高いという結果となった一方、納得性や面白さについては低いという結果が得られた。この結果となった原因の一つとして学習データの属性の偏りが考えられる。すなわち、意外性や独特性の高いデータが多く、納得性や面白さが高いデータが少なかったことが考えられる。今後は、様々な観点で学習データを評価し、バランスよく学習させることで、より高精度に笑いを誘起するインタラクションモデルの生成を検討していく。

参考文献

- [1] 松阪崇久：ヒトはなぜ笑うのか？：行動学の視点から、笑い学研究，No. 21, pp. 5-18 (オンライン)，入手先 (<http://ci.nii.ac.jp/naid/110009840886/>) (2014).
- [2] 前田実香，鬼沢武久：単語の関連性とおもしろさを取り入れたなぞかけ生成，感性工学研究論文集，Vol. 5, No. 3, pp. 17-22 (2005).
- [3] 灘本明代：漫才ロボット：Webを用いた漫才台本自動生成，システム制御情報学会研究発表講演会講演論文集，Vol. 60, p. 4p (オンライン)，入手先 (<http://ci.nii.ac.jp/naid/40020836203/>) (2016).
- [4] 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).
- [5] GoogleJapan: 大規模日本語 n-gram データの公開，Google (オンライン)，入手先 (<https://japan.googleblog.com/2007/11/n-gram.html>) (参照 2016-12-18).
- [6] 林宏太郎，神田崇行，宮下敬宏，石黒 浩，萩田紀博：ロボット漫才：社会的受動メディアとしての二体のロボットの利用，日本ロボット学会誌 = Journal of Robotics Society of Japan，Vol. 25, No. 3, pp. 381-389 (オンライン)，DOI: 10.7210/jrsj.25.381 (2007).
- [7] 伊藤大幸：ユーモア経験に至る認知的・情動的過程に関する検討：不適合理論における2つのモデルの統合へ向け，認知科学，Vol. 14, No. 1, pp. 118-132 (オンライン)，DOI: 10.11225/jcss.14.118 (2007).
- [8] 日比野太郎，原 理奈，トカチヨフサワ，白井宏美：「人とロボットの漫才」におけるユーモア生起不適合理論の観点から，慶應義塾大学 (オンライン)，入手先

(<http://shirai.sfc.keio.ac.jp/ORF2016/index.html>) (参照
2016-12-18).

- [9] ヴイストーン株式会社：普及型ロボットプラットフォーム
「Sota:Social Talker」, ヴイストーン株式会社 (オンライン),
入手先 (<https://www.vstone.co.jp/products/sota/>) (参照
2016-12-19).