

令和 6 年 6 月 17 日現在

機関番号：32692

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K12191

研究課題名（和文）雑談対話システムへの個性の付与とそのコミュニケーションへの影響について

研究課題名（英文）Development of a chat dialog system with personality and its impact on communication

研究代表者

岩下 志乃（Iwashita, Shino）

東京工科大学・コンピュータサイエンス学部・准教授

研究者番号：00360503

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：個性を持った雑談対話システムの構築を目的とした。人と既存のチャットボットが雑談対話（チャット）をする実験を行い、話者の性格特性と対話行為に関するJAISTタグを付与した72名分の対話コーパスを作成して公開した。話者の性格特性と発言特徴の関係を分析した結果、話者本人が自覚している性格よりもユーザなどの他者から見た性格の方がシステムの印象に直結することを確認した。作成した対話コーパスを学習データとして事前学習済み言語モデルをファインチューニングすることで、雑談対話の応答生成を行うモデルを構築し、社会的スキルと特性シャイネスに関して、性格印象の特性を反映できた。

研究成果の学術的意義や社会的意義

本研究では、個性を持つ雑談対話システムを構築することと、そのために必要な個性の特徴を紐づけた対話コーパスを構築することを目的とした。話者の性格特性と対話行為に関するJAISTタグを付与した72名分の対話コーパスを作成して公開することで、対話システムにおける性格の付与を目指す研究に貢献した。また、チャットベースの対話において、自覚している性格と表出された性格が異なり、表出された性格の方が重要であることを示したことで、対話システムに個性を付与する際の着目点を示すことができた。さらに、性格特性の一部を対話システムに付与することができ、ユーザに対して相性の良い雑談対話システムの実現に寄与できた。

研究成果の概要（英文）：The purpose of this study was to construct a chat dialog system with personality. We conducted an experiment in which a person and an existing chatbot engaged in a chat dialogue (chat), and created and published a corpus of dialogues for 72 people to which the personality tags of the speaker and the JAIST tags related to the dialogue act were assigned. As a result of analyzing the relationship between the speaker's personality and the characteristics of his/her speech, we confirmed that the personality of the speaker as seen by others, such as the user, is directly related to the impression of the system, rather than the speaker's personality as perceived by the speaker himself/herself.

By fine-tuning a pre-trained language model using the dialogue corpus as training data, we constructed a model for generating responses to chat dialogues, and were able to reflect the characteristics of personality impressions in terms of social skills and shyness.

研究分野：自然言語処理，感性情報処理

キーワード：雑談対話システム 個性 性格特性

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

近年、対話システムの需要が高まっている。IBM Watson を導入したコールセンターや Apple Siri によるスマートフォンの Q&A 機能など、特にユーザの目的を達成するためのタスク指向型の対話システムが研究・開発されてきている。一方で、目的が明確でない雑談対話システム(非タスク指向型対話)は、人間に癒しや娯楽を与え、親近感や愛着、信頼といった人間関係の形成にも寄与するなど、タスク指向型対話とは異なる効果を持つ。雑談対話システムでは、コミュニケーションそのものが目的であるため、ユーザの個性を理解することや、対話システム自体の個性をどう表出するかが重要になる。対話システムが個性を持つことにより、人間側の対応に変化が現れる可能性が高い。

雑談対話システムは、キーワード照合ルールを用いる方法や、対話コーパスとの照合を行う方法、深層学習を用いる方法がある。現在、実装されている雑談対話システムは、前者 2 つの方法を用いたものが主であるが、ルールを人手で用意するコストや、コーパスの質が悪いことによる不適切な応答文が出力されるといった問題がある。深層学習を用いる方法では、学習に適したコーパスが無いという問題があり、実用には至っていない。

2. 研究の目的

本研究の目的は、以下の 4 つである。

- (1) 対話における「個性」を分類できるようにすること
- (2) 深層学習に適した対話コーパスを構築すること
- (3) 個性を持つ雑談対話システムを構築すること
- (4) 個性を持つ雑談対話システムに対する印象や返答方法が個人によってどの程度異なるのかを検証すること

これまでに人間の個性を分類するための研究は多く行われてきており、人間側の個性に応じて対応を変化させるような雑談対話システムが構築されてきている。本研究では、雑談対話システムに個性パラメータを導入することにより、個性的な雑談対話システムを構築することと、人間側の個性との関係により対話がどのように変化するのが明らかにすることを目的としている。この目的を達成することにより、ユーザに対して相性の良い雑談対話システムを用意して会話することで、システムに対する好感度が上がったり、スムーズに対話できるようになったりすることが期待できる。このようにスムーズな対話が可能になれば、タスク指向の対話においてもユーザの意図を楽に推論できるようになり、タスクを効率よく遂行できるといった貢献ができる可能性がある。また、深層学習に適した雑談対話コーパスを構築することで、今後の雑談対話研究に貢献できると考える。

3. 研究の方法

本研究は以下の 3 段階で実施する。

- (1) 被験者が性格診断アンケートへの回答とチャットボットとの対話を行う実験を実施し、収集した性格特性データと対話ログと紐付けることで、性格情報をもつ雑談対話コーパスを作成する。また、雑談対話システムの発話に対する印象評価と評価者の個人特性との関連についても調べる。次に、他者から見た性格の印象(性格印象)に着目し、収集した対話ログを第三者が見て話者の性格を推測する実験を行い、性格印象データを作成する。このデータを用いて、対話における話者の性格印象と発言特徴の関係を分析する。
- (2) 1)の分析結果をふまえて、性格印象データと紐付けた対話コーパスを学習データとして事前学習済み言語モデル T5 をファインチューニングすることで、雑談対話の応答生成を行うモデルを構築し、性格印象の再現性について検証を行う。併せて、ChatGPT を用いた性格印象の再現性についても検証を行う。
- (3) ユーザに合わせた対話を行うために、ユーザ発話中の単語カテゴリ・感情語数・単語総数を使用した 3 つの手法でユーザの性格特性を推測する手法について検証を行う。

4. 研究成果

- (1) 雑談対話コーパスの作成と分析

被験者が性格診断アンケートへの回答とチャットボットとの対話を行う実験を実施した。被験者はチャットボットと以下の3つのテーマについて10分間ずつ、計30分間のチャットを行った。

- 新しいクラスで初めて知り合った隣の席の学生と初めて会話する感じで
- 好きな について(の内容は自由、途中で変えてもよい)
- 小学校の時の思い出

収集された対話の全発話に対し、人手でJAIST タグ付き自由対話コーパス[1]において付与されている9種類の対話行為「自己開示」、「質問(YesNo)」、「質問(What)」、「応答(YesNo)」、「応答(平叙)」、「あいづち」、「フィラー」、「確認」、「要求」を付与した。また、チャットボットの発話には破綻の有無を表すタグも付与した。

チャット終了後、被験者の個人特性を取得するためのアンケート調査を行った。性格特性に関する調査項目は、Big Fiveの短縮版であるTIPI-J[2]を10項目、コミュニケーションスキルENDCORE[3]を6項目、社会的スキルKiSS-18[4]を18項目、特性シャイネス尺度[5]を16項目である。これらのアンケート結果から被験者ごとに以下の項目を含む性格特性データを作成した。

- TIPI-Jから得られる項目：外向性、協調性、勤勉性、神経症傾向、開放性
- ENDCOREから得られる項目：自己統制、表現力、解読力、自己主張、他者受容、関係調整
- KiSS-18から得られる項目：社会的スキル
- 特性シャイネス尺度から得られる項目：特性シャイネス

表1にタグ付きコーパスのデータ例を、表2に性格特性のデータ例を示す。表1において<ID>は被験者ID、Sはチャットボットを表す。本研究では今回収集したデータのうち被験者63人分、計6,135ペアの発話からなる、

表1 タグ付きコーパスのデータ例

話者	発話	破綻	対話行為
<ID>	アンパンマンは好きですか？		質問(YN)
S	ご想像にお任せします。		応答(平叙)
<ID>	僕は好きじゃないです		自己開示
S	意外なことにコッペパンマンはまだいないんだよね。		確認
<ID>	面白いですね		応答(平叙)
S	たくさんキャラクターが出てくる作品として、ギネスに認定されたくらいなんだ。		自己開示
<ID>	好きな教科は何ですか？		質問(WH)
S	最近では教科書もタブレット化されているよね。	x	確認

表2 性格特性のデータ例

ID	外向性	協調性	勤勉性	...	社会的スキル	特性シャイネス
001	5.50	6.50	3.00	...	3.83	2.69
002	2.50	5.50	3.00	...	3.33	4.06

自分で自覚している性格と他者から見た性格は同じとは限らない。本研究ではある人のふるまいを他者が見た時に感じる性格の印象を性格印象と呼ぶ。個性をもつ対話システムへの応用を考えた場合、話者本人が自覚している性格(性格特性)よりもユーザなどの他者から見た性格(性格印象)の方がシステムの印象に直結するため重要と考えられる。そこで、雑談対話コーパスにおけるチャットボットとの対話のやり取りを第三者が見て話者である被験者の性格を推測する実験を行い、コーパスに含まれる63人の性格印象データを構築した。

話者の性格印象と発言特徴の関係を分析した。発言特徴としてはLinguistic Inquiry and Word Count(LIWC)[6]の日本語版J-LIWC2015[7]における69カテゴリの表現の出現頻度を用いる。性格印象4種類と発言特徴79種類の各組み合わせに対してピアソンの積率相関係数を求めた。

これらの結果から、性格印象ごとの発言特徴の傾向をまとめると以下のようになる。

- 協調性が高いように見える人は、応答(平叙)が多く、質問(WH)と要求が少なく、文字数が多い。また、interrog(「何」、「どこ」、「どんな」など)とipron(「何」、「それ」、「その」など)とpronoun(「何」、「私」、「僕」など)とsocial(「好き」、「ゲーム」、「人」など)が少ない。
- 勤勉性が高いように見える人は、応答(平叙)が多く、あいづちが少なく、文字数が多い。また、casepart(「の」、「は」、「が」など)が多く、assent(「そう」、「うん」、「そうですね」など)とinterrogとipronとnetspeak(「(笑)」、「マジ」、「(汗)」など)とnonflu(「あっ」、「やった」、「あれ」など)とpronounが少ない。
- 社会的スキルが高いように見える人は、応答(平叙)が多く、応答(YN)と要求が少なく、

文字数が多い。また、affiliation(「好き」、「ゲーム」、「君」など)とinterrogとipronとpronounとsocialが少ない。

- 特性シャイネスが高いように見える人は、質問(WH)が多く、文字数が少ない。また、affiliationとinterrogとipronとnumber(「何」、「ゼロ」、「1」など)とpronounが多く、differ(「けど」、「でも」、「なら」など)が少ない。

4種類の性格印象をそれぞれ目的変数とし、本研究で対象とした79個の発言特徴の中からステップワイズ法により説明変数を7個選択して重回帰分析を行った。一連の分析結果から、自己診断に基づく性格特性と他者から見た性格印象の間に相関は無く、性格特性とは異なり性格印象は協調性、勤勉性の双方について以下に示すような発話特徴と明確な相関が見られることが分かった。

1. 発話行為の質問(Question)およびLIWCのinterrogと負の相関がある。
2. 発話行為の要求(Command)と負の相関がある。
3. LIWCのpronounと負の相関がある。
4. LIWCのangerと負の相関がある。

性格特性・コミュニケーションスキルと破綻発話への不快感の関係を明らかにするために、TIPI-JとENDCOREの各得点を予測変数、標準化済みの不快感得点を目的変数として重回帰分析を行った。その結果、TIPI-Jの「神経症傾向」とENDCOREの「自己統制」が不快感に影響を与えていることが分かった。

(2) 雑談対話の応答生成を行うモデル構築と性格印象の再現

これらの分析結果から、雑談対話コーパスと性格印象データを用いて、事前学習済み言語生成モデルT5[8]をファインチューニングすることで、特定の性格をもつように見える雑談対話の応答生成を行うモデルを構築した。ベースとなる事前学習済みモデルは、Hugging Faceで配布されているsonois/t5-base-japaneseを使用し、Hugging Face Transformersライブラリを用いて実装を行った。

まず、一般的な対話のやり取りのパターンを学習するためにTwitter(現X)から収集した大量のツイート・リプライ対を用いてベースとなるモデルの第1段階のファインチューニングを行う。次に、雑談対話コーパスにおけるチャットボットの発話とそれに続く被験者の発話の組を学習データとして第2段階のファインチューニングを行い、4種の応答生成モデルを構築した。特定の性格印象の高低を指定して生成した応答文が指定どおりの性格印象をもつか検証する実験を行った。学習に使用した被験者63人分の対話データに含まれていない部分からチャットボットの発言を100個抽出し、1つの文につき4種類の性格印象の高低を変更して合計800個の応答文を生成した。

得られた800個の文をそれぞれ入力文と組にした形で3人に提示し、性格印象を4段階で評価してもらった。そして文ごとに3人の評価の平均を求めて、その文の評価値とする。種類ごとに各100文の評価値の平均を求めたものを図1に示す。平均の差の検定(対応のあるt検定)を行った結果、社会的スキルに関しては5%水準、特性シャイネスに関しては1%水準で有意な差が見られた。社会的スキルが高いと指定して生成した応答文の方が社会的スキルが低いと指定して生成した応答文よりも社会的スキルの評価値の平均が高いという意図どおりの結果が得られた。一方、協調性と勤勉性に関して同様の傾向が見られたが、統計的に有意な差は得られなかった。

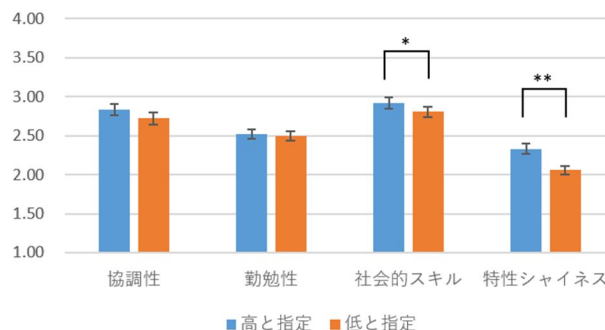


図1 アンケートによる生成文の評価

(3) ユーザの性格特性を推測する手法の提案

ユーザ発話中の単語のカテゴリ、単語総数、感情語数の3要素を用いた性格推定手法について検討した。カテゴリとしてはJapanese Linguistic Inquiry and Word Count(JLWC)[9]に登録されているカテゴリのうち、性格推定に使えるカテゴリを表3のように分類した。

表3 性格特性によるJIWCカテゴリの分類

	正のスコア	負のスコア
開放性	一人称, 知覚, 欲求	リスク
外向性	二人称, 社会的関係, 同意	ネットスラング
神経症傾向	負の感情, 宗教, 死	数量詞, 洞察
調和性	正の感情, 同意	否定, 疑問, 非流暢
誠実性	数量詞, 時間, 仕事, お金	砕けた口調, フィラー, ネットスラング, 同意

感情語による性格推定では,文中に使用されている感情語の数を使用してスコアを算出する。この推定手法では,ML-Ask[10]を使用して入力文を解析し,取得できた感情語の数を数える。これは,情動性や感受性を測ることが出来るため,開放性と神経症傾向の推定に使用する。文章の長さによる性格推定では,文全体に含まれる単語数を使用してスコアを算出する。これは,ユーザの積極性を測ることが出来るため,開放性と外向性の性格推定に使用する。上記の特徴を反映した対話システムにおいて,システムの性格推定の精度と性格表現を評価するための実験を実施した。表4に実験協力者の客観的性格と,チャットボットの性格変化の過程を示す。性格推定システムでは,実験協力者ごとに異なる性格変化を示した。また,対話中にチャットボットの性格が変更された場合,どの実験協力者も性格の変化を感じ取ることが出来ており,性格ごとの特徴に違いが見られた。対話履歴から著者が判断した被験者性格特性とシステムの性格特性は72%程が概ね一致していた。

表4 チャットボットの性格変化

実験協力者の性格印象	チャットボットの性格変化
A 神経症傾向(低), 調和性(高)	神経症傾向(低) 調和性(高) 誠実性(低)
B 外向性(高), 調和性(低)	調和性(高) 開放性(低)
C 外向性(高)	外向性(高)
D 開放性(高), 誠実性(低)	開放性(高), 神経症傾向(低)
E 開放性(低), 調和性(低)	神経症傾向(高) 誠実性(高) 開放性(低)
F 誠実性(高), 神経症傾向(低)	開放性(高) 誠実性(高), 神経症傾向(低)
G 開放性(高)	開放性(高), 外向性(高)
H 開放性(高), 外向性(高)	開放性(高) 外向性(高)

<引用文献>

- [1] 福岡知隆, 白井清昭: 対話行為に固有の特徴を考慮した自由対話システムにおける対話行為推定, 自然言語処理, 24(4), 523-547, 2017.
- [2] 小塩真司, 阿部晋吾, カトローニ ピノ: 日本語版 Ten Item Personality Inventory (TIPI-J)作成の試み, パーソナリティ研究, 21(1), 40-52, 2012.
- [3] 藤本学, 大坊郁夫: コミュニケーション・スキルに関する諸因子の階層構造への統合の試み, パーソナリティ研究, 15(3), pp.347-361, 2007.
- [4] 菊池章夫: 思いやりを科学する: 向社会的行動の心理とスキル, 川島書店, 1988.
- [5] 相川充: 特性シャイネス尺度の作成および信頼性と妥当性の検討に関する研究, 心理学研究, 62(3), 149-155, 1991.
- [6] Chung C. K., and Pennebaker J. W.: Linguistic Inquiry and Word Count (LIWC): Pronounced "Luke, " ... and other useful facts, in McCarthy P. M., and Boonthum-Denecke C. eds. Applied Natural Language Processing: Identification, Investigation and Resolution, Information Science Reference, 206-229, 2012.
- [7] Igarashi T., Okuda S., and Sasahara K.: Development of the Japanese version of the linguistic inquiry and word count dictionary 2015, Frontiers in Psychology, 13, 2022.
- [8] Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., and Liu P. J.: Exploring the limits of transfer learning with a unified text-to-text transformer, Journal of Machine Learning Research, 21(140), 1-67, 2020.
- [9] 奈良先端科学技術大学院大学, 日本語感情表現辞書 JIWC -Dictionary, <https://sociocom.naist.jp/jiwc-dictionary/>
- [10] Michal Ptaszynski, Pawel Dybala, Wenhan Shi, Rafal Rzepka and Kenji Araki, "A System for Affect Analysis of Utterances in Japanese Supported with Web Mining", Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, 21(2), 30-49, 2009.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計17件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 杉本徹, 仲村詩穂, 岩下志乃, 伊藤紀子, 林篤司
2. 発表標題 雑談対話における発言特徴と性格印象の関係分析と対話システムへの応用に向けた検討
3. 学会等名 第39回ファジィシステムシンポジウム
4. 発表年 2023年

1. 発表者名 ト秋予, 伊藤 紀子
2. 発表標題 雑談対話システムの発話に対する印象評価と評価者の個人特性との関連
3. 学会等名 第31回日本機能言語学会秋期大会
4. 発表年 2023年

1. 発表者名 田村 克, 櫻 リベカ, 大竹 正彦, 林 篤司, 岩下 志乃
2. 発表標題 ユーザの性格特性に基づき自身の性格を変化させる雑談対話システム
3. 学会等名 第19回感性工学会春季大会
4. 発表年 2024年

1. 発表者名 松村 崇太郎, 櫻 リベカ, 大竹 正彦, 林 篤司, 岩下 志乃
2. 発表標題 ユーザ感情に共感したモーションで応答する雑談対話エージェント
3. 学会等名 第19回感性工学会春季大会
4. 発表年 2024年

1. 発表者名 山本 茉莉, ト 秋予, 伊藤 紀子
2. 発表標題 チャットボットとの雑談対話に対する 印象評価と評価者の性格特性 ・コミュニケーションスキルとの関連
3. 学会等名 言語処理学会 第29回年次大会
4. 発表年 2023年

1. 発表者名 伊藤紀子, 岩下志乃, 杉本徹, 林篤司, ト秋予
2. 発表標題 ユーザの特性情報付きチャットボットとの雑談対話コーパスの概要
3. 学会等名 言語資源ワークショップ
4. 発表年 2022年

1. 発表者名 杉本徹, 岩下志乃, 伊藤紀子, 林篤司
2. 発表標題 性格情報を持つ対話コーパスの構築と対話システムへの応用に向けた検討
3. 学会等名 第38回ファジィシステムシンポジウム
4. 発表年 2022年

1. 発表者名 仲村詩穂, 杉本徹
2. 発表標題 性格情報を持つ対話コーパスに基づいた話者の性格特性と発言特徴の関係の分析
3. 学会等名 第38回ファジィシステムシンポジウム
4. 発表年 2022年

1. 発表者名 竹内慶介, 杉本徹
2. 発表標題 研究室紹介システムにおけるユーザの興味度と知識量を考慮した話題誘導
3. 学会等名 情報処理学会第84回全国大会
4. 発表年 2022年

1. 発表者名 関萌水, 林 篤司, 岩下 志乃
2. 発表標題 雑談対話における話題遷移の検出
3. 学会等名 第37回ファジィシステムシンポジウム
4. 発表年 2021年

1. 発表者名 安藤 大智, 林 篤司, 岩下 志乃
2. 発表標題 エージェントへの意識モデルの導入と食欲を例にした個性獲得シミュレーション
3. 学会等名 第37回ファジィシステムシンポジウム
4. 発表年 2021年

1. 発表者名 Toru Sugimoto, Yusuke Temmanyu
2. 発表標題 Response generation reflecting the personality of a chat dialogue system
3. 学会等名 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS&ISIS2020) (国際学会)
4. 発表年 2020年

1. 発表者名 鈴木勸太、杉本徹
2. 発表標題 Encoder-Decoderモデルを用いた文章表現を豊かにする執筆支援システム
3. 学会等名 言語処理学会第27回年次大会
4. 発表年 2021年

1. 発表者名 Daichi Ando and Shino Iwashita
2. 発表標題 A Study on Reinforcement-Learning Agents with Personality Through the Implementation of Character Parameters
3. 学会等名 Joint 11th International Conference on Soft Computing and Intelligent Systems and 21st International Symposium on Advanced Intelligent Systems (SCIS&ISIS2020) (国際学会)
4. 発表年 2020年

1. 発表者名 井 凌泓, 伊藤 紀子
2. 発表標題 雑談対話システムとの対話にみられる対話破綻とユーザのコミュニケーション方略 ユーザ特性による違い
3. 学会等名 第28回日本機能言語学会秋期大会
4. 発表年 2020年

1. 発表者名 杉本徹, 和田史織, 天満谷侑亮
2. 発表標題 ユーザの性格を考慮した対話応答生成への取り組み
3. 学会等名 第35回ファジィシステムシンポジウム
4. 発表年 2019年

1. 発表者名 井 凌泓, 伊藤 紀子
2. 発表標題 雑談対話システムとの対話にみられる対話破綻とユーザのコミュニケーション方略
3. 学会等名 第27回日本機能言語学会秋期大会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>チャットボットとの対話コーパスと性格特性・社会的スキルのデータ公開について https://www2.teu.ac.jp/iws/chat_corpus.pdf</p>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	杉本 徹 (Sugimoto Toru) (30277280)	芝浦工業大学・工学部・教授 (32619)	
研究分担者	伊藤 紀子 (Ito Noriko) (00391863)	同志社大学・文化情報学部・准教授 (34310)	
研究分担者	林 篤司 (Hayashi Atsushi) (60777895)	国立研究開発法人農業・食品産業技術総合研究機構・農業口 ボテイクス研究センター・特別研究員 (82111)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------