

## 科学研究費助成事業 研究成果報告書

平成 30 年 6 月 4 日現在

機関番号：16101

研究種目：基盤研究(B) (一般)

研究期間：2015～2017

課題番号：15H02770

研究課題名(和文)多属性感性の共感性を用いた発話誘発による対話継続法

研究課題名(英文) Dialogue continuation by utterance induction using multi-attribute sensitivity empathy

研究代表者

泓田 正雄 (Fuketa, Masao)

徳島大学・大学院社会産業理工学研究部(理工学域)・教授

研究者番号：10304552

交付決定額(研究期間全体)：(直接経費) 11,600,000円

研究成果の概要(和文)：コンピュータによる質問と、人間の応答に対して多属性感性の抽出を行い、それらの共感的感性を判定することにより、想定外発話に対するコンピュータの返答を生成し、会話継続を行うことを目的とする。質問文からは、話題感性、時制、質問の種類を抽出し、応答文からは、話題感性、心情感性、時制、肯定・否定を抽出する。これらの組み合わせから返答文の作成を行った。実験の結果、返答の成功率が90.0%と高い精度を得られた。

研究成果の概要(英文)：The purpose of this research is to continue conversation by generating computer responses to unexpected utterance. For generating computer responses, multi-attribute sensitivities from questions of the computer and human's response are extracted, and empathic sensibilities are determined. Topic sensitivity, tense, and type of question are extracted from the question sentences, and topic sensitivity, emotion sensitivity, tense, affirmation / negation are extracted from the response sentence. The reply sentence was generated from these combinations. From experimental results, the accuracy of the response was 90.0%.

研究分野：自然言語処理

キーワード：感性情報学 対話処理

1. 研究開始当初の背景

和み・癒し効果のある対話（コミュニケーション）による「心のケア」支援の重要性が議論されている。この対話では、共感的な対話継続が重要であるが、国内外の質問応答研究は、少ない対話回数で回答を得る効率化を目的とし、対話継続には適用できない。また、高齢者や要介護者への対話支援は重要であるが、国内外の研究成果は少ない。これまで、軽・中度の認知症患者への対話実験により、リハビリテーション効果がある対話継続度の評価法で研究実績を上げてきたが、感性理解処理がないことで、共感的感性による意思疎通できないという課題が未解決である。

2. 研究の目的

コンピュータによる質問と、人間の応答に対して多属性感性の抽出を行い、それらの共感的感性を判定することにより、想定外発話に対するコンピュータの返答を生成し、会話継続を行うことを目的とする。

3. 研究の方法

本研究では、コンピュータによる質問と人間による応答に対してコンピュータの返答を生成する。

(1) 質問文の解析

質問文からは、話題感性、時制（過去、現在、未来）、質問の種類（「はい」か「いいえ」で答えられるか）を抽出する。話題感性の抽出は、形態素解析とYahoo キーフレーズ抽出を用い、文の主格、または目的格を話題とする。時制、質問の種類は、パターンマッチングの照合ルール（図1）を作成し、多属性照合を行うことで抽出した。また、質問文中の動詞、形容詞、形容動詞を述語として抽出する。

```
#pat:<過去パターン>:<過去>
{
<助動詞(た・だ) / 終止1「た」>
<助動詞(です) / 連用>
<助動詞(た・だ) / 「である」終止「た」>
>
<助動詞(ます) / 連用>
<助動詞(ない) / 連用1「なかっ」>
}
```

図1. 質問文の照合ルールの例

(2) 応答文の解析

応答文からは、話題感性、心情感性（良、悪、無）、時制（過去、現在、未来）、肯定・否定（肯定、否定、希望）を抽出する。話題感性と時制は、応答文と同じ手法で抽出する。心情感性と肯定・否定についても、パターンマッチングの照合ルールを作成し、多属性照合を行うことで抽出を行う。また、質問文中

の動詞、形容詞、形容動詞を述語として抽出する。肯定・否定と心情は、ユーザの意図を理解することで、相槌や次の質問など適切な返答を作成するために使用する。

(3) 返答文の作成

まず、質問文の種類と応答文の話題を比較し、質問文の種類が「はい」か「いいえ」で答えられないであり、応答文の話題が「無し」であった場合、話題が一致していないとする。次に、質問文と応答文から抽出された話題を、日本語 Wordnet を利用して比較する。2つの話題の関係を調べ、一方の話題を上位概念へとさかのぼるともう一方の話題にたどり着くような、単純な上位、下位の関係にある場合、話題が一致しているとし、そうでない場合は話題が一致していないとする。質問文と応答文のいずれかの話題がない場合や日本語 Wordnet に存在しない場合は、質問文と応答文から抽出した述語を話題と同様に日本語 Wordnet を利用して比較する。質問文と応答文のいずれかの述語がない場合は話題が一致しているとし、いずれかの述語が日本語 Wordnet に存在しない場合は話題が一致していないとする。

次に、質問文の時制と応答文の時制を比較し、時制が一致しているかを判定する。応答文の時制が現在である場合を除いて、質問文と応答文の時制が異なる場合は時制が一致していないとする。

表1. 特徴の組み合わせと返答の対応

時制	肯否	心情	返答(時制は質問, 応答によって変更する)
一致	肯定	良	それはいいですね。どうしてですか?
		無	(相槌作成) どうでしたか?
		悪	そうなんです。ところで、(別の質問)
	否定	良	そうなんです。では何ですか?
		無	そうなんです。では何ですか?
		悪	そうだったんですか。では何でしたか?
	希望	良	そうなんです。では何ですか?
		無	そうだったんですか。では何ですか?
		悪	そうだったんですか。ところで、(別の質問)

これまでに抽出した特徴から返答を作成する。質問文と応答文で話題が一致しているかどうか、時制が一致しているかどうか、応答文の肯定・否定と心情が何かといった特徴の全ての組み合わせについて、テンプレートを準備し、そのテンプレートを用いて返答を作成する。用意したテンプレートは、応答に対する相槌と「話題を広げる返答」、「正しい答えを求める返答」、「違う話題へ移る返答」のいずれかの組み合わせになっている。特徴の組み合わせと返答の対応の表（一部）を表1に示す。話題や時制が一致していなくても応答文の心情が「良」であれば話題を広げる場合があるが、これはユーザがその話題について話したがっているとしてユーザに話を合わせるためである。また、特徴の組み合わせのうち質問文と応答文で話題と時制が一致し、応答文が肯定文で心情が無だった場合、

応答文の内容を確認する相槌を作成する。質問文の話題の直後にある格助詞「が」「を」「に」、副助詞「は」を助詞として抽出して「<話題><助詞><述語>」という文を作成し、文末を応答文の時制に合わせて整形したものを相槌とする。質問文と応答文がともに「話題無し」であった場合は「<述語>」という文を作成する。

#### 4. 研究成果

##### (1) 特徴抽出の実験結果

提案手法の時制、肯定・否定、心情といった特徴抽出の精度評価を目的とし、応答文からの特徴抽出の正解率から評価をおこなった。実験データとして、700個の応答文から作成した多属性照合のルール、提案手法のルール作成に使用していない応答文500個を使用し、時制、肯定・否定、心情の判定をおこない、正解率を求めた。肯定・否定については応答データに記載された正解と比較をおこない、時制、心情については人手で判定をおこなった。

時制の正解率は97.8%、肯定・否定の正解率は92.4%、心情の正解率は98.4%といずれも高い正解率を得られた。誤判定をした理由として、多属性照合のルール不足が挙げられる。例として、「思い出したくないし全く懐かしくありません」という応答文が心情「良」として判定されてしまった。これは「懐かしい」が含まれる場合に心情「良」と判定するルールを設定したためである。そのため、この実験結果より新しくルールを追加することでより判定の精度を上げる。

##### (2) 返答文作成の実験結果

提案手法の有効性に対する評価を目的とし、質問に対して適切でない応答をした場合の返答の成功率から評価をおこなった。また、比較対象としてNTTドコモの雑談対話APIの対話と比較をおこなった。

実験データとして、50個の質問文と700個の応答文から作成した多属性照合のルール、提案手法のルール作成に使用していない質問文30個、用意した質問文に対してそれぞれ適切な応答文30個を使用し、応答文からランダムに選んだものを質問文への応答とすることで想定外発話を再現した。対話システムの評価実験において、返答の成功率を用いて評価をおこなった。返答の正誤の判定は人手でおこなった。

実験の結果対話数30回のうち、20回が質問に対して意味の通らない想定外発話となった。実験結果より、正しい返答をした数が27回、返答の成功率が90.0%という結果が得られた。比較対象のNTTドコモの雑談対話APIは任意の質問を用意できないため同条件ではないが、対話を続ける中でシステムからの質問があった際に、用意した応答文からランダムに選んだものを応答とすることで実験をおこなった。結果として成功率は13.3%であり、提案手法の方が良い結果を得られた。

正しい返答をした例を図2に示す

システム：青い山脈を読んだことがありますか？ ユーザ：食べませんでした。 システム：そうだったんですね。 ところで、青い山脈を読んだことがありますか？
システム：東京都は、過ごしやすい天気でしたか？ ユーザ：うどんが食べたいです。 システム：そうなんですか。 ところで、東京都は過ごしやすい天気でしたか？
システム：炊事の仕事で、手は荒れましたか？ ユーザ：よく食べます。大好きです。 システム：それはいいですね。 ところで、炊事の仕事で、手は荒れましたか？
システム：すいとんを食べましたか？ ユーザ：はい。 システム：すいとんを食べたんですね。どうでしたか？

図2. 正しい返答の例

返答の作成に失敗した原因は3つ挙げられる。1つ目の原因は、話題が概念辞書に載っていない場合である。例として「今、食べたいものはありますか？」という質問に対して「広島県の福山です。」と応答した場合、応答文からは「福山」が話題として抽出されるが、福山は概念辞書に載っておらず、応答文からは述語も存在しない。そのため、この場合は「はいかいいえで答えられる質問」に対して「はい」と答えたと判断されてしまい、システムは話題を広げる返答を作成した。これは実験結果から新たに概念辞書を更新することによって改善されると考えられる。2つ目の原因は、概念辞書と実際の概念の差異である。例として、「林業の仕事は、きつかったですか？」という質問に対して「よかったです。」と応答した場合、林業の仕事は良い仕事だったと捉えることができるが、「きつい」と「よい」は単純な上位、下位概念の関係にあたらないため、提案した手法では話題が一致していないと判断されてしまい、システムはもう一度同じ質問を繰り返した。提案した手法ではこのように意味的に近いはずの単語を判別することができないことがあるため、概念辞書から単語の意味類似度を算出するなど、新たな比較基準を設定することで改善できると考えられる。3つ目の原因は、返答作成のためのテンプレート不足である。例として、「東京オリンピックがありましたか、覚えていますか？」という質問に対して「知ってはいますが覚えていません。」と応答した場合、「そうなんですか。では何でしたか？」という返答を作成した。しかし、質問や応答によっては意味が通らない返答となるため、より汎用性の高いテンプレートを作成する、質問と応答で話題が一致して応答が否定文だった場合ユーザはその話題への興味がないと判断して他の話題へ移るといった改善案が考えられる。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表](計2件)

笠原 諒, 泓田 正雄, 森田 和宏, ユーザからの想定外発話に対応した対話システム FIT2017(第16回情報科学技術フォーラム), 2017

笠原 諒, 泓田 正雄, 森田 和宏, 想定外発話に対応した対話システムの提案, 平成28年度電気関係学会四国支部連合大会, 2016

## 6. 研究組織

(1)研究代表者

泓田 正雄 (Fuketa, Masao)

徳島大学・大学院社会産業理工学研究部  
(理工学域)・教授

研究者番号: 10304552

(2)研究分担者

森田 和宏 (Morita, Kazuhiro)

徳島大学・大学院社会産業理工学研究部  
(理工学域)・准教授

研究者番号: 20325252

任 福継 (Ren, Fuji)

徳島大学・大学院社会産業理工学研究部  
(理工学域)・教授

研究者番号: 20264947