

ユーザ感情を考慮した返答発話生成ルールの自動生成手法

田渕 悠真 目良 和也 黒澤 義明 竹澤 寿幸
 広島市立大学情報科学部
 tabuchi@ls.info.hiroshima-cu.ac.jp
 {mera, kurosawa, takezawa}@hiroshima-cu.ac.jp

1 はじめに

現在様々な分野で自然言語対話システムが普及しつつある。さらに近年は指示された作業を実施するタスク指向型対話だけでなく、雑談のような非タスク指向型の対話についても研究が進められている。しかし従来の自然言語対話システムには“字面でしか判断できない”という問題がある。例えばシステムからの「元気ですか?」という問いかけに対して、力なく「元気だよ...」と答えても、音声認識処理では「元気だよ」という文字列としてしか情報を得られないため、システムはユーザが元気なものとして対話を進めてしまう。しかし人間同士の対話の場合、たとえ「元気だよ」と答えていても、声の調子や表情などのノンバーバル情報からも相手の状態を判断し、「本当は元気じゃないのに強がっているな」というような判断をすることができる。

この問題を解決するため、我々はユーザの感情を推定し、推定した感情も考慮して返答発話を返すシステムを提案している[1]。このシステムは音響分析によりユーザ感情を推定し発話内容と推定感情の両方を考慮した対話を行うことができる。このシステムは言語情報と推定した話者感情から返答発話生成ルールデータベースを検索し、条件にマッチした返答発話生成ルールを用いて作成した発話を返す。したがって、このシステムで様々な日常会話を成立させるためには返答発話生成ルールの大幅な追加が必要不可欠である。

しかしながら、この対話システムで用いられる返答発話生成ルールは、どのような発言に対してこの返答発話を返すのかだけでなく、どのような話者感情の際にこの返答発話を返すのかまで定義しなければならない。しかし、話者感情の条件付けの作業にはかなりの作業コストがかかり、さらに作業者の主観が入り過ぎてしまう恐れがあるという問題がある。

そこで本研究では、大規模テキストコーパスから収集した返答発話候補の直前の発話（先行発話）のテキスト情報から推定した話者感情の傾向に基づいて、返答発話生成ルールの話者感情条件部を自動生成することを目指す。そのためにまず、同様の返答発話に対する先行発話を大量に収集し、ある返答発話がどのような話者感情のときによく用いられるのかを分析する。そして、収集した先行発話-返答発話対に対して、返答発話から推定される先行発話時

の話者感情を条件として追加する。

2 感情を考慮した自然言語対話システム

2.1 対話システムの処理の流れ

図 1 に本研究で用いる自然言語対話システム[1]の処理の流れを示す。まずマイクでユーザ発話を取得する。そして取得した音声から音響分析ツールで算出した音響的特徴を機械学習器に入力することによってユーザの感情を positive, negative, neutral の3種類から1つ選ぶ。同時に、音声認識エンジンを使用して発話音声を文字列に変換する。この文字列と推定した話者感情の両方を使って返答発話生成ルールデータベースを検索し、条件にマッチした返答発話生成ルールを用いて作成した発話を返す。

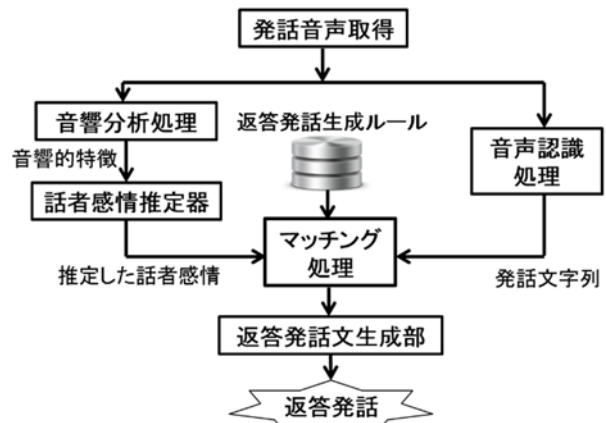


図 1 感情を考慮した自然言語対話システム

2.2 返答発話生成ルール

[1]で用いている返答ルールは、“先行発話”，“システムの直前発話”，“話者感情”を条件として持っており、全ての条件に適合すれば“返答発話”を返す。“先行発話”は直前のユーザの発話文字列を表す。“システムの直前発話”は先行発話の前にシステムが出力した発話を表す。“話者感情”は先行発話を音響分析して得られる感情の種類を表す。これにより、同じ字面のユーザ発話に対しても口調に合わせて異なるリアクションをとることができる。“返答発話”は条件に合致した場合にシステムが出力する返答発話の文字列を表す。

表 1 に返答発話生成ルールの例、図 2 に表 1 のル

ールを用いた対話例を示す。最初にユーザが悲しい声で「おはよう」と言った場合、同じ先行発話を条件とするルール 1 とルール 2 のうち、話者感情が negative であるルール 2 が適用され、「元気そうじゃないね。どうしたの?」と返答する。それに対してユーザが悲しい声で「大丈夫」と答えると、今度はルール 5 が適用され、「そうか、何か楽しい音楽を流そうか?」と返答する。もし最初の「おはよう」が positive な話者感情での発話だったなら、ルール 2 ではなくルール 1 が適用され、システムの返答および以下の対話が変わっていく。

表 1 返答発話生成ルール

No.	ルール選択条件			返答発話
	先行発話	システムの直前発話	話者感情	
1	おはよう	(条件無し)	positive or neutral	いい天気だね
2	おはよう	(条件無し)	negative	元気そうじゃないね、どうしたの?
3	試験に落ちたよ	元気そうじゃないね、どうしたの?	negative	今度もっと頑張ろう
4	大丈夫	元気そうじゃないね、どうしたの?	positive	良い日を過ごしてください
5	大丈夫	元気そうじゃないね、どうしたの?	negative	そうか、何か新しい音楽を渡そうか?
6	大丈夫	(条件無し)	neutral	何が大丈夫なの?

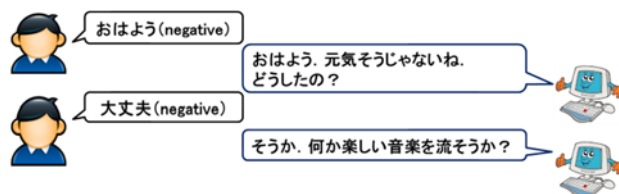


図 2 返答発話生成ルールを用いた対話例

3 返答発話生成ルールの自動生成手法

3.1 返答発話生成ルールの生成手順

表 1 のような返答生成ルールを作成する際、発話文字列だけでは推定できない話者感情の情報が必要となる。テキスト対話コーパスからの対話対の収集を想定した場合、話者感情の推定のためには顔文字やエモートアイコンが本文に付加されている必要がある。しかし、Line や SMS などの個人的なチャットログとは違い、一般に公開されているテキスト対話コーパスにこのような情報が付与されていることは多くない。とはいえ、プライバシーの観点から個人的なチャットログが大量に公開/提供されることは期待できない。しかし、感情を人手で推定すれば、作業者の主観が入ってしまうことや、大量の作業コストがかかってしまうという問題がある。

そこで本論文では「ある感情状態にある相手の先行発話に対する返答としてよく用いられている返答発話は、同じ感情状態で発せられる別の先行発話に対する返答としても有効である」という仮定に基

づいた感情条件付き返答発話生成ルールの自動生成手法を提案する。図 3 に提案手法の手順を示す。

手順 1: 大規模テキスト対話コーパスから、同一の表現とみなせる返答発話に対する先行発話をまとめて収集する。図 3 の例では、返答発話 r1 に対する先行発話として、U1, U2, U3, ..., Um が収集されている。

手順 2: テキスト表現から感情推定が可能な先行発話に対してのみ、先行発話の話者感情を推定する。感情推定手法として本研究では感情語データベースを用いた手法を用いる (3.2 節参照)。図 3 の例では、感情推定が可能な先行発話 (U2, U3, ..., Um) に対して、それぞれ positive, negative のいずれかの感情が推定されている。また U1 はテキストから感情推定ができなかったものとする。

手順 3: 先行発話の話者感情を集計し、最も多い感情タイプを返答発話に対する先行発話の一般的な話者感情とする。図 3 の例では、**返答発話 r1 は negative な先行発話に対してよく用いられると判断される。**

手順 4: 感情推定が出来なかった先行発話について、返答発話生成ルールを作成する。このときの先行発話の話者感情は手順 3 で推定した感情とする。図 3 の例では、【先行発話: U1, システムの直前発話: (条件無し), 話者感情: negative → 返答発話: r1】というルールが生成される。

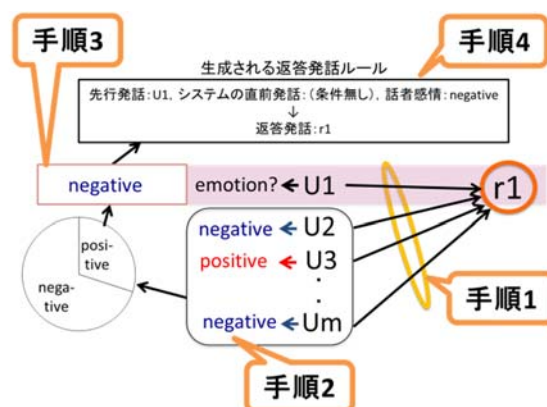


図 3 返答発話生成ルールの生成手順

3.2 発話文字列からの感情推定処理

発話文字列からの話者感情推定は、感情語データベースとのマッチングと発話文の構文構造を考慮して行う。推定する感情は positive, negative, 判定不能の 3 クラスである。

以下に感情推定の手順を示す。最初に推定する発話文字列に対して形態素解析を行い、感情語データベースにマッチする語句の有無を調べる。発話文字列中でデータベースにマッチした語句の感情クラスが全て同じならば、その感情クラスを話者感情とす

る。マッチングした語句の感情クラスが一致していない場合、[2]の手法に基づき主述語の感情クラスを優先する。発話文字列中に感情語データベースにマッチする語が全く無い場合は、感情推定出来ない発話とする。

感情語データベースは、現代形容詞用法辞典[3]と感情表現辞書[4]の感情極性情報だけでなく、三省堂こどもこくごじてん[5]に作業員1名が感情を付与したものを加えて構築した。本研究で用いた感情語データベースの収録語数は4,830である。

4 評価実験

提案手法を用いて生成した返答発話生成ルールの妥当性、特に話者感情条件の妥当性について評価実験を行う。

4.1 使用するテキスト対話コーパス

本研究では日本語自然会話書き起こしコーパス(旧名大会話コーパス)[6]を使用する。このコーパスには69,104発話が含まれている。コーパス中のある発話をユーザによる先行発話、その直後の発話をシステムの返答発話と考え、先行発話と返答発話の対話対として収集する。なお、長すぎる文に対しては感情推定精度が下がることと、対話システムの返答として不適切なもの(身内話や昔話など)が多かったため、本研究では先行発話、返答発話ともに20文字以下のもの23,315発話を対象とする。

4.2 返答発話生成ルールの作成

提案手法による返答発話生成ルール作成のため、実験に用いる23,315件すべての発話に対して話者感情推定を行った。結果を表2に示す。

表2 コーパスの感情推定結果

	positive	negative	判定不能
個数	5,606	1,561	16,148

表2より、コーパス中の全発話の推定感情の比率にそもそも偏りがあることが確認できる。式(1)の結果、positiveの比率は0.78となっている。

$$\text{positive 比率} = \text{positive} / (\text{positive} + \text{negative}) \quad (1)$$

よってコーパス中の全発話の推定感情数の比率を、先行発話の推定感情数に偏りがあるかの判定の際に考慮する。

評価対象となる返答発話生成ルールを作成するため、まずコーパスに頻出する語を含む返答発話を60件作成し、各返答発話に対する先行発話を収集して、先行発話の推定感情の比率を計算した。返答発話文字列をどこまで同じとみなすかの判定は、今回は人手で行った。そして感情推定された対話対と感情推定不能だった対話対がともに15個以上である

返答発話のうち、先行発話の話者感情のpositiveの比率が0.78以上のもの12件と0.78以下のもの12件に提案手法を適用し、273件の返答発話生成ルールを作成した。作成したルールの“システムの直前発話”条件は今回無しと設定している。評価実験では、作成されたルールのうち、話者感情がpositiveのルール11件と話者感情がnegativeのルール11件について評価を行った。今回評価実験の対象とする返答発話生成ルールを表3に示す。

表3 提案手法で作成された返答発話生成ルール

No.	ルール選択条件			返答発話
	先行発話	システムの直前発話	話者感情	
P1	白い花でしょ	(条件無し)	positive	きれいだよね
P2	うわー、きれいやねえ	(条件無し)	positive	きれいだよね
P3	ほいでね、このネコちゃんがいる	(条件無し)	positive	きれいだよね
P4	きれいですよねえ、坂多いけど	(条件無し)	positive	きれいだよね
P5	いやー、すげえかわいいじゃん	(条件無し)	positive	かわいいね
P6	そう。で、このワンちゃんもいるんだよね	(条件無し)	positive	かわいいね
P7	すごーい、ほれ。見て	(条件無し)	positive	かわいいね
P8	この赤いのもかわいいね	(条件無し)	positive	かわいいね
P9	どうでしょうか	(条件無し)	positive	かわいいね
P10	しゃれてるねー	(条件無し)	positive	かわいいね
P11	このお魚おいしい。ちっちゃい	(条件無し)	positive	おいしい
N1	夜はやだね	(条件無し)	negative	怖いね
N2	ちょっと手術はね	(条件無し)	negative	怖いね
N3	すごい怖い	(条件無し)	negative	怖いね
N4	心霊スポットとか	(条件無し)	negative	怖いね
N5	やーねー。白い粉ねー	(条件無し)	negative	怖いね
N6	何でいいじゃん	(条件無し)	negative	嫌です
N7	作って	(条件無し)	negative	嫌です
N8	ねえ、付き合いで行ってよ	(条件無し)	negative	嫌です
N9	乗れるからだいじょうぶ、トラシク	(条件無し)	negative	嫌です
N10	なんで?	(条件無し)	negative	嫌です
N11	ちょっと見て行ってもいいよね	(条件無し)	negative	大丈夫

4.3 評価実験の手順

作成した返答発話生成ルールの適切さを評価するための実験を以下の手順で行った。なお被験者は22~24歳の大学生および大学院生(男性10名、女性3名)である。

手順 1A: 先行発話を音声で被験者に提示する。音声は、音声創作ソフトウェアCeVIO Creative Studio S [7]を用いて作成した。聞き間違いが無いように、音声と同時に発話文字列も被験者に提示した。

手順 1B: 手順 1A と同時に被験者に話者感情を想起させるノンバーバル情報を提示する。positiveな話者感情では“笑顔の写真”と“元気そうな声”、negativeな話者感情では“怒った顔の写真”と“怒った声”としている。なおCeVIO Creative Studio S [7]による音声作成時にpositiveな感情音声は「元気」パラメータ100、negativeな感情音声は「怒り」パラメータ100と設定して音声を作成した。

手順 2: 先行発話の提示後, 返答発話をテキストで提示する.

手順 3: 手順 1B の話者感情を反転した先行発話と手順 2 の返答発話のセットを同様に被験者に提示する

手順 4: positive な先行発話に対する返答発話の適切さと negative な先行発話に対する返答発話の適切さについてそれぞれ 3 件法 (適切/どちらでもない/不適切) で回答してもらう.

4.4 実験結果

被験者の回答結果の分布に対してカイ 2 乗検定を行い, 有意差の有無について調査した. 適切な回答数と不適切の回答数の有意差の分析結果を表 5 に示す. 表内の表記は, “表 3 のルール番号 (提示した話者感情)” となっているため, P?(positive)あるいは N?(negative) という対話対はルールに則った話者感情が, P?(negative)あるいは N?(positive) という対話対ではルールと逆の話者感情が提示された場合の返答発話に対する評価となっている.

表 5 返答発話生成ルールに基づく返答発話の適切さ

適切 > 不適切		有意差 無し	適切 < 不適切	
p<0.01	p<0.05		p<0.05	p<0.01
P1(pos)	P4(pos)	P3(pos)	P2(neg)	P1(neg)
P2(pos)	P6(pos)	N1(pos)	P7(neg)	P3(neg)
P5(pos)	P9(pos)	N2(pos)	P10(neg)	P6(neg)
P7(pos)	P10(pos)	N3(pos)		
P8(pos)	N7(pos)	N4(pos)		
P9(pos)	N3(neg)	N5(pos)		
P11(pos)	N5(neg)	N6(pos)		
N8(pos)		N9(pos)		
N11(pos)		N10(pos)		
N1(neg)		P4(neg)		
N2(neg)		P5(neg)		
N4(neg)		P8(neg)		
		P9(neg)		
		P11(neg)		
		N6(neg)		
		N7(neg)		
		N8(neg)		
		N9(neg)		
		N10(neg)		
		N11(neg)		

表 5 の結果, positive な話者感情を条件とする返答発話生成ルールを positive な先行発話に適用したところ 11 件の返答発話全てが適切と評価された. また, negative な話者感情を条件とする返答発話生成ルールを negative な先行発話に適用したところ 11 件中 5 件の返答発話が適切と評価された. 逆に条件とは逆の話者感情を提示した場合, positive な話者感情のルールでは 6 件が不適切と判定されたが, negative な話者感情のルールでは条件部に設定し

た感情と逆の話者感情を提示しているにも関わらず適切な返答だと評価された事例が 2 件あった.

次に, 同じ返答発話生成ルールを positive な先行発話に適用した場合と negative な先行発話に適用した場合の, 適切であるという回答数の有意差について表 6 に示す.

表 6 話者感情による返答発話の適切さの回答の有意差

positive > negative		有意差 無し	positive < negative	
p<0.01	p<0.05		p<0.05	p<0.01
P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, N11		N6, N7, N8, N9, N10	N3	N1, N2, N4, N5

表 6 の結果, positive な話者感情を条件とする返答発話生成ルールでは 11 件全て positive な話者感情を提示した際の返答が有意に適切だと評価された. しかし, negative な話者感情を条件とする返答発話生成ルールでは, 5 件が negative な話者感情を提示した際の返答が有意に適切, 5 件が有意差無し, 1 件は逆に positive な話者感情を提示した場合の方が有意に適切だと判定された.

これらの結果から, 本提案手法は positive な話者感情を条件とする返答発話生成ルールの自動生成には有効であるといえる.

5 今後の課題

今後の課題としては, レーベンシュタイン距離や類似度を用いて同一表現とみなせる返答発話を収集する手法の開発に取り組む予定である.

参考文献

- [1] Takumi Takahashi, *et al.*, “Natural Language Dialog System Considering Speaker’s Emotion Calculated from Acoustic Features,” Proc. of IWSDS2016, 2016.
- [2] 菅原久嗣, Alena Neviarouskaya, 石塚満, 日本語テキストからの感情抽出, 2009 年度人工知能学会全国大会発表論文集, 2009.
- [3] 飛田良文, 浅田秀子, 現代形容詞用法辞典, 東京堂出版, 1991.
- [4] 中村明, 感情表現辞書, 東京堂出版, 1993.
- [5] 三省堂編修所, 三省堂こどもこくごじてん, 三省堂, 2002.
- [6] 日本語自然会話書き起こしコーパス, <https://nknet.ninjal.ac.jp/nuc/templates/nuc.html>, 2017 年 1 月 17 日アクセス
- [7] CeVIO Official Web, <http://cevio.jp/>, 2017 年 1 月 17 日アクセス