

共感・発話促進のための 多様な聞き手応答を生成する傾聴対話システム

石田 真也 井上 昂治 高梨 克也 河原 達也

京都大学 大学院情報学研究科 知能情報学専攻

1. はじめに

近年、人と対話を行うシステムが日常生活においても盛んに利用されるようになってきている。対話の種類として、タスク遂行型の対話だけでなく雑談対話の研究も行われている [1]。雑談対話の一つとして、傾聴対話 [2] がある。システムが傾聴対話を行うことで、ユーザの話したいという欲求の充足や高齢者の認知機能の維持などが期待される [3][4]。

傾聴対話を実現するためには、ユーザが発話しやすいように、システムのふるまいを設計する必要がある。従来研究では、相槌の韻律 [5] や形態 [6] が扱われている。また、相槌以外の聞き手応答として、我々は、ユーザの発話内容に応じて、慣用的表現の語彙的応答、一部を繰り返す繰り返し応答、詳細を問う掘り下げ質問の選択と生成 [7] に取り組んでいる。また、[8] では聞き手応答を評価応答を含む 6 種に分類している。本研究では、より自然な傾聴対話を行うために、上記の掘り下げ質問、繰り返し応答、語彙的応答、評価応答という聞き手応答に加えて、システムが自分自身に関する話をする自分語りの生成を行う。上記の各応答を提案手法で生成したのち、ユーザ発話の音声認識、焦点解析の結果や、生成した応答などの情報を特徴量として、適切な応答の種類を選択する。

2. 傾聴対話システムの全体像

提案するシステムの概要を図 1 に示す。はじめに、ユーザ発話に対して音声認識と、韻律情報の抽出を行う。次に、音声認識結果に対して形態素解析・係り受け解析を、これらに基づいて焦点解析も行う。形態素解析には Juman を、係り受け解析には KNP を、焦点解析には吉野らの手法 [9] を用いた。上記の解析結果を用いて、傾聴対話のためのシステム応答を生成する。ここでは、各応答モジュールが独立に生成した応答の中から、応答選択器によってふさわしい応答の種類を選択し生成する。

以下、各応答生成モジュールについて述べる。語彙的応答は「そうですか」のみをそのまま出力する。繰り返し応答は以下の例のように、焦点単語と文末表現を用いて繰り返す。

ユーザ「先週友人と沖縄に行ってきました」
システム「沖縄ですか」←繰り返し応答

2.1 掘り下げ質問

掘り下げ質問は先行するユーザ発話の詳細を問うものである。本研究では、焦点解析によって得られた焦点（存在しない場合は先行発話の述語）と疑問詞のペアを基に掘り下げ質問を生成する。焦点とは、ユーザ発話内の単



図 1: 応答生成システムの全体像

表 1: ラベルに対応する評価応答

	ポジティブ	ネガティブ
客観的	素敵ですね	大変ですね
主観的	いいですね	残念ですね

語のうち、その後の対話でトピックとなりうるものを指す。候補となる疑問詞は「いつ」「どこ」「誰」「どんな」もしくはその派生（「いつの」「どこの」など）である。候補の疑問詞の中から焦点（もしくは述語）と相性の良いものを絞り込むために、単語 N-gram 確率を用いる。この値が最も高い疑問詞と焦点（もしくは述語）のペアを用いて掘り下げ質問を生成する。例を示す。

ユーザ「昨日友達と海に行ってきました」
システム「どこの海ですか？」←掘り下げ質問

ユーザ「昨日兄と夕食を食べました」
システム「どこで食べましたか？」←掘り下げ質問

2.2 評価応答

評価応答は先行するユーザ発話に対してシステムが評価を示す応答である。本研究では小林らの日本語評価極性辞書 [10] を用いて評価応答を生成する。当該辞書内には、用言・名詞の両者の各単語について、「ポジティブかネガティブか」と「客観的か主観的か」の 2 つのラベルが付与してある。提案手法では、ユーザ発話を文節毎に区切り、各文節にポジティブ/ネガティブ、主観/客観のラベルを当該辞書を参照しながら付与し、各発話内の多数決で発話文全体のラベルを決定する。そして、そのラベルに応じてシステムの発話を図 1 のように決定する。

2.3 自分語り

自分語りは先行するユーザ発話に関連した事象に対して、聞き手自身の考えや情報を提示する応答である。この自分語りは対話をより自然にする目的に加えて、直接掘り下げ質問を生成できないようなユーザ発話に対しての緩衝となり、その直後に掘り下げ質問が続くことの自然さを実現する効果が期待される。自分語りを行うため

Attentive Listening System Generating Various Responses for Empathizing and Encouraging User Utterances: Masanari Ishida, Koji Inoue, Katsuya Takanashi, Tatsuya Kawahara (Kyoto University)

には、システムに背景知識を持たせる必要があるが、本研究では単純化して、先行発話の焦点単語と述語をテンプレートに当てはめるスロット形式を採用する。テンプレートを以下に示す。

- 私も (焦点) が好きです
- 私も (焦点) です
- 私は (焦点)(述語) したことはありません
- 私も (焦点)(述語) してみたいです

また、これらの応答の中から言語モデル的に正しいものを選択するために、単語 N-gram 確率を用いて、その値が最大になるものを応答の候補とした。

2.4 応答選択器

応答生成モジュールによって生成された5つの応答の内から一つを選択する。言語的・文脈的な観点から、傾聴の聞き手応答としてふさわしい応答を予測する。学習に用いる素性は以下である。

- ユーザ発話文の末尾の表現 (文末表現・終助詞)
- ユーザ発話全体に含まれる内容語の数
- 焦点解析結果
- 応答生成における N-gram 確率
- ユーザ発話の末尾 100ms の韻律情報
 - ・パワーの最大値, 平均
 - ・F0の最大値, 最小値, 平均, レンジ
- ユーザ発話後からのポーズ長
- 直前のシステム応答の種類

3. 対話データ

応答選択器の学習や評価に用いるデータセットを作成するために、対話データへのアノテーションを行った。使用した対話データは、語り手役の被験者の語りをアンドロイド ERICA が傾聴する 1対1の WOZ 対話である。ERICA の音声発話と非言語行動の制御は別室の操作者が行っている。今回は1セッション10分程度の対話データを9セッション使用する。

4. 評価実験

応答選択器の性能を評価した。先述の対話データにおいて実際に ERICA が発話した応答の種類を予測を行った。ただし、ERICA の発話のうちで、提案した5種の聞き手応答のどれにも該当しないものには「その他」ラベルを付与している。ERICA の各発話が生じたタイミングにおいて、2.4節で挙げた特徴量を入力とし、応答ラベルを出力とする。学習モデルはロジスティック回帰を用いた。また、データセット内の ERICA の463発話を5分割し、交差検定によって精度を計算した。今回は ERICA が実際に出力した応答ラベルと、システムが予測したもののうちで ERICA の応答とは異なるが文脈的に妥当だと考えられるものを正解としている。

各応答ラベルの適合率、再現率、F値、チャンスレートを表2に示す。評価応答のF値が比較的強くチャンスレートの値を下回った。その要因として、学習の際の素性の選択が不十分であることが考えられる。評価応答は語り手が述べる自身の体験や感情に対して聞き手が評価を示す応答であるので、発話内容における行為の主体が語り手自身であるかを判断するための情報が必要となる。また、自分語りの精度も低かった。自分語りも評価応答

表2: 応答選択器の評価実験の結果

	適合率	再現率	F 値	チャンスレート
語彙的応答	133 / 245	106 / 164	0.59	0.35
評価応答	11 / 29	8 / 89	0.16	0.19
繰り返し応答	67 / 142	50 / 101	0.48	0.22
自分語り	5 / 13	4 / 47	0.14	0.10
掘り下げ質問	18 / 34	12 / 62	0.28	0.13
合計	180 / 463	180 / 463	0.44	0.23

と同様に、相手に共感を示す目的でなされるため、精度が低い要因は、基本的に評価応答と類似していると考えられる。一方で、そのほかの種類に応答はチャンスレートよりも高い精度を示している点や、全ての応答の種類合計におけるF値がチャンスレートの値を上回っていることも評価できる。語彙的応答の精度が高いことは、システムが高い頑健性を持っていることを意味するため、評価できる。今回用いたデータセットでは、応答の種類が「その他」に該当するものが多数存在し、その種類も多岐(挨拶や確認応答など)にわたっていたため、適切に学習に生かせていないことが考えられる。さらなる精度向上のために、「その他」の応答の扱い方に目を向けていきたい。

5. おわりに

本稿では、自然な傾聴対話システムに必要と考えられる5種類の応答生成手法について述べた。従来の一般的な応答に加え、聞き手自身に関する自分語りを生成した。また、複数の応答候補から統計的手法によりふさわしい応答を選択する手法を提案した。傾聴対話コーパスを用いた実験の結果、評価応答を除きチャンスレートよりも高い精度で聞き手応答が選択できることを確認した。今後は、データセットの各発話に対して許容される応答の種類を追加で付与するアノテーションを行うことや、本稿で述べた5種類の応答が出やすい設定のもと収録した、新たな対話データを評価実験で使用することで、応答選択器の性能を向上させたい。その後、被験者実験により傾聴対話システム全体の主観評価および定量評価も行う予定である。

謝辞 本研究は、JST ERATO 石黒共生ヒューマンロボットインタラクションプロジェクトの支援を受けて実施された。

参考文献

- [1] 河原達也. 音声対話システムの進化と淘汰: 歴史と最近の技術動向. 人工知能学会誌, Vol. 28, pp. 45–51, 2013.
- [2] 下岡和也ほか. 音声対話ロボットのための傾聴システムの開発. 言語・音声理解と対話処理研究会, Vol. 58, pp. 61–66, 2010.
- [3] R. Butler. The life review: An interpretation of reminiscence in the aged. *Psychiatry*, Vol. 26, pp. 65–76, 1963.
- [4] 大武美保子ほか. 回想法から見た共想法の考察と連携の可能性. 人工知能学会全国大会, 6 2010.
- [5] 上里美樹ほか. 傾聴対話における相槌の韻律的特徴の同調傾向の分析. 言語・音声理解と対話処理研究会, Vol. 70, pp. 7–13, 2014.
- [6] 山口貴史ほか. 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成. 人工知能学会研究会資料, Vol. 76, pp. 49–54, 02 2016.
- [7] 石田真也ほか. 傾聴対話システムのための発話を促す聞き手応答の生成. 人工知能学会研究会資料, Vol. 77, No. 01, pp. 1–6, 08 2016.
- [8] 伝康晴. 対話への情報付与. 小磯花絵 (編), 日本語コーパス3: 話し言葉コーパス設計と構築, pp. 101–130. 朝倉書店, 東京, 2015.
- [9] K. Yoshino, et al. Conversational system for information navigation based on pomdp with user focus tracking. *Computer Speech Language*, Vol. 34, pp. 275–291, 2015.
- [10] 小林のぞみほか. 意見抽出のための評価表現の収集. 自然言語処理, pp. 1483–1486, 2005.