

6Q-01

# 傾聴対話システムのための多様な聞き手応答の生成

石田 真也<sup>†</sup> 井上 昂治<sup>‡</sup> 中村 静<sup>‡</sup> 高梨 克也<sup>‡</sup> 河原 達也<sup>‡</sup>

<sup>†</sup> 京都大学 工学部情報学科      <sup>‡</sup> 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により、話し相手としてユーザの欲求を満たす効果が期待されている。

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、システムのふるまいとして相槌に焦点が当てられてきた。システムが相槌をうつ場面において、相槌の韻律をユーザ発話へ同調させたり [3]、相槌の形態をユーザ発話の情報に応じて決定すること [4] などが取り組まれてきた。本研究では、相槌以外に、より多様な聞き手応答の生成を実現する。具体的には、ユーザの発話内容に応じて、慣用的表現で同意を示す「語彙的応答」、先行発話の一部を繰り返す「繰り返し応答」、先行発話の詳細を問う「掘り下げ質問」を行う。[5] では相槌を「応答系感動詞」、「感情表出系感動詞」、「語彙的応答」、「評価応答」、「くり返し」、「共同補完」の6つに分類している。提案手法の2つはこのうちの「語彙的応答」と「くり返し」に対応していて、さらに「掘り下げ質問」を生成することを目指す。ユーザ発話、それに対する焦点・応答文の例を図1で示す。ユーザ発話に対して焦点解析を行い、生成される各応答に対してその適切さを判別することで、候補から最も適した応答を選択する。

## 2. システム概要

提案するシステムの概要を図2に示す。入出力ともテキスト形式である。はじめに、ユーザ発話文に対して、形態素解析・述語項構造解析を行う。形態素解析にはJumanを、述語項構造解析にはKNPを使用する。焦点解析モジュールは、ユーザ発話の中で焦点となる内容語を抽出した上で、焦点の有無に応じて下記のように処理を分岐する。焦点がある場合(図2の判別器1)には、焦点に関する掘り下げ質問と繰り返し応答のどちらが適切かを判定する。焦点がない場合(図2の判別器2)には、述語に関する掘り下げ質問と語彙的応答のどちらが適切かを判定する。選択された種類の応答生成モジュールで応答文を生成し出力する。各モジュールの詳細について、次節以降で説明する。

### 2.1 焦点解析

入力ユーザ発話を形態素解析・述語項構造解析したものである。焦点解析として先行研究 [6] の手法を用いる。この手法ではユーザ発話に対して、「各節の出現位置」、「各格要素の出現位置」、「各単語の品詞」、「品詞とその出現位置のペア」、「各節が担っている格」、「文に存在している全ての格」、「各格とその出現位置のペア」、な

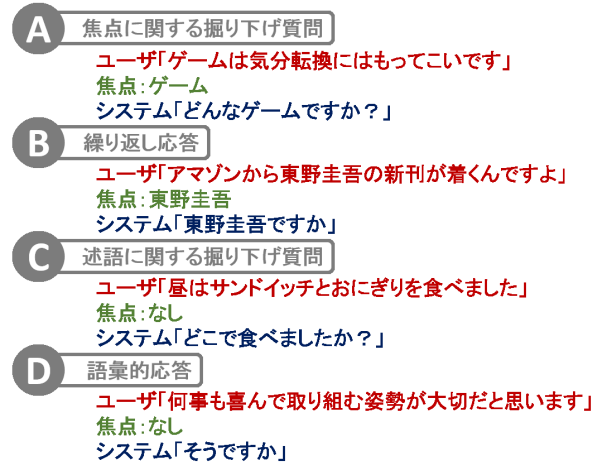


図1: システム応答例

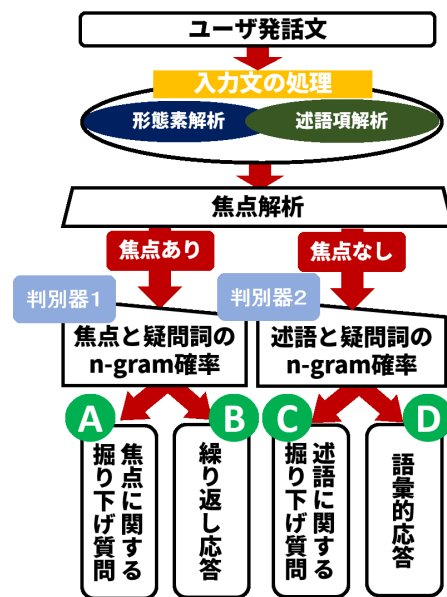


図2: システム概要

どの特徴量を抽出し、条件付確率場(以下CRF)で学習を行い、焦点を解析する。[6]の対話システムから抽出した計918個ユーザ発話にアノテーションを付加したものを学習データとして検証を行ったところ、ユーザ発話に焦点があるかどうかの判定の精度は99.9%で、焦点となるフレーズの抽出精度は78.5%と報告されている。

### 2.2 判別器

判別器は図2に示す2種類である。

#### 2.2.1 判別器1

入力をユーザ発話内の焦点となる単語として、その焦点に関する掘り下げ質問の生成に適した疑問詞(もしくは疑問詞を選択する。ここでは、その焦点と疑問詞のそれぞれとのn-gram確率に基づいて適切さを判定する。n-

Various Response Generation for Attentive Listening System: Masanari Ishida, Koji Inoue, Shizuka Nakamura, Katsuya Takahashi, Tatsuya Kawahara (Kyoto University)

gram は, Julius ディクテーションキット<sup>§</sup> に付属されているもので, 現代日本語書き言葉均衡コーパス (BCCWJ) から学習されている. 疑問詞が1つの形態素からなるときは bi-gram 確率, 2つの形態素からなるときは tri-gram 確率を求める. 疑問詞を次のように分類する.

- ・どんな/どの/なんの/なにの/何の
- ・いつの/何時の ・どこの/何処の
- ・だれの/誰の

各カテゴリ毎に確率の総和を計算し, その値が閾値以上となるカテゴリが存在する場合, その中で最大のものの代表的な疑問詞 (各カテゴリの先頭に記述されている疑問詞) と焦点となる単語のペアを「焦点に関する掘り下げ質問モジュール」へと出力する. 閾値以上となる疑問詞が存在しない場合は焦点となる単語のみを「繰り返し応答モジュール」へと出力する.

### 2.2.2 判別器 2

入力をユーザ発話内の述語として, その述語に関する掘り下げ質問の生成に適した疑問詞 (もしくは疑問詞なし) を出力する. 以下のカテゴリに分類された各疑問詞と述語との n-gram 確率に基づいて判定する.

- ・何を/なにを
- ・誰を/だれを
- ・何に/なにに
- ・誰に/だれに
- ・何から/なにから
- ・誰から/だれから
- ・何まで/なにまで
- ・誰へ/だれへ
- ・どこで/何処で
- ・誰と/だれと
- ・どこへ/何処へ
- ・いつ
- ・どこから/何処から
- ・いつまで/何時まで
- ・どこまで/何処まで
- ・いつから/何時から
- ・どこに/何処に

前節と同様にカテゴリ内の確率の総和を求め, 閾値を超えたカテゴリがある場合, 最大のカテゴリの代表的な疑問詞と述語のペアを「述語に関する掘り下げ質問モジュール」へと出力する. どのカテゴリも閾値を超えない場合, 「語彙的応答モジュール」へ進む.

### 2.3 応答生成モジュール

応答生成モジュールは図2に示されているように4種類存在する. 図1に各々の例を示す. 語彙的応答モジュールは常に「そうですか」という応答を出力する. 繰り返し応答モジュールは, 入力となる焦点の単語を用いて「○○(焦点)ですか」という応答を出力する. 焦点に関する掘り下げ質問は, 入力となる疑問詞と焦点の単語を用いて「△△(疑問詞)○○(焦点)ですか?」という応答を出力する. 述語に関する掘り下げ質問は, 入力となる疑問詞と述語を用いて「△△(疑問詞)○○(述語)か?」という応答を出力する.

### 3. 評価実験

本稿では, 雑談対話コーパス [7] から抽出した 100 のユーザ発話を評価実験に用いた. 焦点解析は正しく行えるという前提で, 評価を行う. 59 個の焦点, 41 個の述語を手動で抽出し, それぞれの疑問詞のラベルを正解デー

焦点有無	応答種類	再現率	適合率
焦点あり	A 掘り下げ質問	9/21	9/12
	B 繰り返し応答	36/38	36/47
焦点なし	C 掘り下げ質問	2/14	2/5
	D 語彙的応答	27/36	27/27
合計		74%	74%

表 1: 実験結果

タとして手動で選択した. 判別器の素性となる特徴量は, 入力となる焦点の単語と各疑問詞との n-gram 確率のみである. 閾値に関しては実験的によいものを定めた. 判別器 1 においてはサンプル中で掘り下げ質問が正解とした焦点と疑問詞のペアの n-gram 確率が登録されていない場合が多かったため, 閾値は 0 に近い値として, 掘り下げ質問が可能なのは無条件に生成することにした. 判別器 2 においてはサンプル中では誤った掘り下げ質問として生成されることが多かったため, 掘り下げ質問の再現率を損なわない程度に閾値を高くした. これらの判別の問題は焦点/述語という入力データに疑問詞の種類 (もしくは疑問詞なし) のラベリングの問題に帰着できる. 評価結果を表 1 に示す. A の再現率が低いのはそもそも n-gram 確率に登録されていない単語が多数存在したことで焦点として選ばれるべき特徴的な単語は逆に n-gram 確率に登録されていないことが原因であると考えられる. また, C の適合率が低いのは, 文の格構造を利用していないため, 既に述べられている情報に関して掘り下げ質問をすることが原因であると考えられる.

### 4. おわりに

本稿では, 傾聴対話システムを実現するために, ユーザ発話内の焦点となる単語の有無と, 焦点となる単語/述語と疑問詞との n-gram 確率をもとに適切な応答文を生成する手法を述べた. 本稿では, 焦点解析は正しく行えるという仮定を置いた. 今後は, 2.1 節で述べたの従来法 [6] に上記 n-gram 確率の特徴を加味することによる焦点解析の自動化に取り組み, 本システムの性能向上を目指す.

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, No.1, pp.45-51, 2013.
- [2] 下岡和也, 徳久良子, 吉村貴克, 星野博之, 渡部生聖: “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] 上里美樹, 吉野幸一郎, 高梨克也, 河原達也: “傾聴対話における相植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.
- [4] 山口貴史, 井上昂治, 吉野幸一郎, 高梨克也, Nigel Ward, 河原達也: “多様な相植をうつ傾聴対話システムのための相植形態の予測,” 人工知能学会研究会資料, SLUD-B502-1, 2015.
- [5] 伝康晴: “講座日本語コーパス 3: 話し言葉コーパス設計と構築,” 小磯花絵編著, 朝倉書店, pp.101-130.
- [6] 吉野幸一郎, 河原達也: “Conversational system for information navigation based on POMDP with user focus tracking,” Computer Speech & Language, Vol.34, pp.275-291, 2015.
- [7] 東中竜一郎, 船越孝太郎: “Project Next NLP 対話タスクにおける雑談対話データの収集と対話破綻アノテーション,” 人工知能学会 言語・音声理解と対話処理研究会 第 72 回, pp.45-50, 2014.

<sup>§</sup><https://github.com/julius-speech/dictation-kit>