

質問タイプの分類に基づく用例なし質問に対する応答生成

磯西 爽太¹ 田中 滉己² 井上 昂治² 高梨 克也² 河原 達也²¹ 京都大学 工学部情報学科² 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

音声対話システムはスマートフォンアプリやスマートスピーカなど、様々な形で実用化されている。また、人間らしい社会的な対話を指向したシステムでは、システム自身の趣味や嗜好といったプロフィール情報が設定されている。例えば、我々は自律型アンドロイド ERICA [1] によるお見合い対話の実現を目指している。このような対話を対象にシステムを実装する際には、プロフィールの一貫性を保つために想定されるユーザからの質問とその応答の用例を事前に用意しておくことが多い。しかし、用例にない質問がなされた場合、「お答えできません」や「どうでしょう」などの画一的な応答を返すしかなくなってしまう。この場合、対話の自然さを損ない、ユーザの対話に対する意欲が低下してしまう可能性がある。関連研究として、ユーザ発話の理解に失敗したときに聞き返すエラーハンドリング [2] や対話の破綻検出 [3] などが挙げられるが、用例がない質問に対する対処方法として直接適用することはできない。また、どのような発話に対しても何かしらの応答を生成できるという点では、ニューラルネットワークによる応答生成 [4] は有用と考えられるが、システムの発話の一貫性を保つのが難しいという問題がある。本研究で提案する方法は、用例なし質問に対して、システムが質問を理解しているかのように見せかけた応答を生成する。具体的には、質問を「個人情報」、「好み」、「習慣」などのタイプに分類し、そのタイプに応じた応答フレームを用いて応答を生成する。はじめに、質問タイプの分類を行うロジスティック回帰モデルを提案する。次に、質問文とそれに対応する上記の応答フレームのデータを用いて、質問文から応答フレームを直接生成するニューラルネットのモデルも提案する。

2. 質問タイプ

プロフィール情報をやりとりする対話における質問のタイプを分類する。ここでは、「個人情報」「習慣」「好み」「経験」「将来に対する希望」「現在までに対する主観」「システムに関する知識」「一般の知識」の8種類を用いる。それぞれの分類基準を表1に示す。「個人情報」と「システムに関する知識」についてはどちらもシステムに関する事実を表すが、「個人情報」はプロフィール情報をやりとりする対話で頻繁になされるものであり、どのようなシステムの設定であっても対応すべき質問に相当する。質問の例を以下に示す。

- 個人情報：「お名前は何ですか」「趣味は何ですか」
- 習慣：「休みの日は何をしますか」
- 好み：「旅行は好きですか」
- 経験：「パンケーキは食べたことがありますか」
- 将来に対する希望：「クリスマスはどのように過ごしますか」

Response generation for out-of-database questions based on classification of question types: Sota Isonishi, Koki Tanaka, Koji Inoue, Katsuya Takanashi, and Tatsuya Kawahara (Kyoto Univ.)

表1: 質問タイプの分類基準

質問タイプ	基準
個人情報	システムに関する、行為ではない現在の事実を聞いているもので、初対面の対話でよく聞かれるもの。
習慣	現在または普段の行為について聞いているもの。
好み	好きなものについて聞いているもの、趣味自体を聞く質問は個人情報に含む。
経験	過去の行為について聞いているもの。
将来に対する希望	未来の未確定の希望や予定について聞いているもの。
現在までに対する主観	現在あるいは過去のことに対してのシステムの考えについて聞いているもの。
システムに関する知識	システムに関する、行為ではない現在の事実について聞いているもので、個人情報に分類されないもの。
一般の知識	システムとは関係のない一般的事実について知っているか聞いているもの。

- 現在までに対する主観：「仕事は忙しいですか」
- システムに関する知識：「ペットは飼っていますか」
- 一般の知識：「ウユニ塩湖を知っていますか」

3. 質問タイプの分類と応答生成

入力される質問に対して、そのタイプを分類する。分類結果に基づき、対応する応答フレームを用いて応答生成する。

3.1 質問タイプの分類

入力特徴量は、学習データの質問文に含まれる機能語、代名詞、副詞的名詞、形容動詞の語幹となる名詞の Bag-of-Words である。ただし、学習データの1%以上のサンプルに出現するもののみを使用する。また、過去形の動詞の有無も特徴量に加える。出力は質問タイプである。ただし、質問が命題質問 (YES / NO 質問) か集合質問かにより後述の応答フレームを切り替える必要があるため、この分類も同時に扱う。したがって、8種の質問タイプと、命題質問か集合質問かの組合せである16種の多クラス分類を行う。ただし、後述の対話コーパスに少数しか含まれないものはまとめて「その他」とする。また、「個人情報」の質問は用例として用意されるべきであると考えられるため、ここでは対象外とする。したがって、実際には10種の多クラス分類を扱う。上記のモデルをロジスティック回帰を用いて学習する。

3.2 応答フレームによる応答生成

質問タイプの分類結果に基づき、対応する応答フレームを用いて応答文を生成する。質問タイプ毎に応答フレームを使い分けることで、システムが質問自体は理解しているようにみせかける。応答フレームには、ユーザの質問文に含まれる焦点語や動詞などがスロットとして埋め込まれている。応答フレーム自体は「反応」と「質問」の2パートで構成する。「反応」ではユーザの質問に対して否定的な回答を簡潔に返すことを方針とする。これにより、答えるのが難しい話題がそれ以上掘り下げ

表 2: 応答フレーム

質問タイプ		応答フレーム
命題	習慣	うーん、[動詞] ないです。[ユーザ名] さんは普段[焦点語]は[動詞] ますか。
	経験	うーんないですね。[ユーザ名] さんはどうですか。
	好み	うーんあまりです。[ユーザ名] さんは[焦点語]は好きですか。
	システムに関する知識	うーん、どうでしょう。[ユーザ名] さんはどうですか。
集合	一般の知識	うーん、知らないです。
	習慣	うーん、色々です。[ユーザ名] さんは[疑問詞(+名詞)][動詞] ていますか。
	経験	うーん、特に何も[動詞] ませんでした。[ユーザ名] さんはどうですか。
	好み	うーん、特にないです。[ユーザ名] さんは何が好きですか。
	将来に対する希望	うーん、特に考えていません。[ユーザ名] さんは何かありますか。
	システムに関する知識	うーん、どうでしょう。[ユーザ名] さんはどうですか。
その他	うーん、どうでしょう。[ユーザ名] さんはどうですか。	

表 3: 質問タイプの分類結果

質問タイプ	適合率	再現率	F 値	
命題	習慣	36/45	36/40	0.847
	経験	3/ 6	3/10	0.375
	好み	6/ 9	6/15	0.500
	システムに関する知識	3/ 7	3/18	0.240
集合	一般の知識	12/18	12/15	0.727
	習慣	36/46	36/40	0.837
	経験	25/34	25/30	0.781
	好み	33/48	33/40	0.750
	将来に対する希望	4/ 4	4/14	0.444
	システムに関する知識	4/ 6	4/18	0.333
その他	1/22	1/ 5	0.074	

られないようにする。「質問」では逆に質問を投げ返す。これにより対話の主導権を得ることができ、場合によっては用例がある話題へ転換することが可能となる。各質問タイプに応じた応答フレームを表 2 に示す。ただし、実際の分類問題で使用される質問タイプ 10 種類のみについてである。例えば、「映画はよく見ますか」という質問文が与えられた場合には、その質問タイプとして「習慣 (命題質問)」と判定し、応答文は「うーん見ないです [ユーザ名] さんは普段映画は見ますか」となる。

4. Seq-to-seq モデルによる応答フレーム生成

前節の方法では、質問タイプの分類に失敗した場合には、不自然な応答フレームを選択してしまう。そこで、ニューラルネットワークの一種である Seq-to-seq モデルを用いて、質問文の単語列を入力とし、応答フレームの単語系列を直接出力するようにする。前節で述べた質問タイプの分類の正解データをもとにして、入力と出力の単語系列を用意し、これを学習データとして用いる。ただし、入出力の単語は分散表現を用いる。

5. 実験

使用した対話データは、遠隔操作されたアンドロイド ERICA と被験者とのお見合い練習対話である。対話数は 31 である。対話中でなされた質問に対して質問タイプのアノテーションを人手で行った。実験では、話題の最初になされたもので「個人情報」以外のタイプに分類された 245 件を用いた。

はじめに、質問タイプの分類について、5 分割の交差検証により評価した。多クラスロジスティック回帰の出力値がいずれも 0.3 未満の場合は「その他」クラスに分類した。結果を表 3 に示す。出現数が多い質問タイプは

高い精度で分類できていることがわかる。「システムに関する知識」については分類精度が低くなった。このタイプに分類される質問文には、定型的な表現がなく、表現に幅があるためと考えられる。

次に、Seq-to-seq モデルによって生成される応答フレームについて評価した。31 対話のうち、21 対話 (170 発話) を学習用、5 対話 (37 発話) を検証用、別の 5 対話 (38 発話) をテスト用として使用した。テストデータ 38 発話のうち、22 発話の出力系列が正解系列と全文一致した。これは、出力系列である応答フレームのパターンが限られているため、今回のデータ量でも Seq-to-seq モデルの学習が行えたと考えられる。多クラスロジスティック回帰による質問タイプの分類に基づいた応答フレームの選択と、Seq-to-seq モデルとの出力を比較すると、いくつかの応答で Seq-to-seq モデルによる改善がみられた。例えば、「やっぱり研究が忙しいと家事がおざなりになったりしてしまうこともありますか。」という習慣の命題質問に対して、多クラスロジスティック回帰では「好み (集合質問)」と誤ってしまい「うーん、特にないです。[ユーザ名] さんは何が好きですか。」と出力され、やや不自然な応答になっている。これに対して Seq-to-seq では「うーん、どうでしょう。[ユーザ名] さんはどうですか。」と出力された。

6. おわりに

本稿では、用例にない質問に対する応答を、質問タイプの分類に基づいて応答フレームを選択するシステムを提案した。さらに、質問文の単語系列から応答フレームの系列を直接生成する Seq-to-seq モデルも提案した。実験では、質問タイプの分類で誤った場合のいくつかの質問文において、Seq-to-seq モデルでは出力する応答フレームが改善されることを確認した。今後の課題として、生成した応答フレーム内の焦点語や動詞などのスロットの内容を、質問文から抽出することが挙げられる。さらに、最終的な応答文の自然さなどについての印象評定実験も行う予定である。

本研究は JST ERATO JPMJER1401 の支援を受けた。

参考文献

- [1] Tatsuya Kawahara. Spoken dialogue system for a human-like conversational robot ERICA. *IWSDS*, 2018.
- [2] Gabriel Skantze. Exploring human error handling strategies: Implications for spoken dialogue systems. *ISCA Tutorial and Research Workshop on Error Handling in Spoken Dialogue Systems*, 2003.
- [3] 東中竜一郎 et al. 対話破綻検出チャレンジ 2. 人工知能学会研究会資料, *SIG-SLUD*, 2016.
- [4] Oriol Vinyals et al. A neural conversational model. *arXiv preprint arXiv:1506.05869*, 2015.