The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

テンプレートを用いた疑問文生成による対話応答文DBの拡充

Augmentation of Dialogue Database by Generating Interrogative Sentences Using Templates

長内洋太 *1 Yota Osanai　尾形朋哉 *1 Tomoya Ogata　小町守 *1 Mamoru Komachi　下川原（佐藤）英理 *1 Eri Sato-Shimokawara　和田一義 *1 Kazuyoshi Wada
山口亨 *1 Toru Yamaguchi　高谷智哉 *2 Tomoya Takatani

*1首都大学東京 Tokyo Metropolitan Univercity　*2トヨタ自動車株式会社 TOYOTA MOTOR CORPORATION

In recent years, there are many dialogue systems selecting an appropriate response from a dialogue database. However, when the number of sentences in the database is small, it may not contain an appropriate sentence in an open domain setting. Therefore, it is desirable to prepare a large number of candidate sentences and to fully expand the dialogue database beforehand. In this research, we generate templates from handwritten responses from actual dialogue logs to augment interrogative sentences for a dialogue database.

1. はじめに

近年，コンピュータが自然言語を用いて人と会話を行うための対話システムの研究が行われている．対話システムには，大別して，経路案内や Web 検索などの特定の目的のために発話，応答を行う task-oriented なもの [Williams 07] と，特定の目的がない，主に雑談を行う non-task-oriented なもの [Wallace 09] の 2 種類がある．

non-task-oriented な対話システムにおいて，事前に応答の候補となる文を生成，または収集し，データベースに大量に蓄えておくことで，ユーザ発話に適切な応答を文選択によって出力するという手法が古くから研究されている [Higashinaka 06][柴田 09]．この手法は，応答文を出力する際，フィルタリングやスコアリングを用いることによって，非文を出力しづらいという強みがある．しかし，データベース内にユーザ発話に対する応答文としてふさわしい文が存在しない場合，適切でない文を選択し，応答文として出力してしまうという問題が存在する．そのため，応答文のデータベースは様々なユーザのあらゆる発話に対応できるよう，応答文となり得る様々な応答文を大量に保持していることが望ましい．

また，比較的小規模な日本語対話データや，Twitter などのスラングやノイズの多いデータセットを用いて生成された応答文は，それらを除外するためのフィルタリングに非常にコストがかかるという問題点がある．

そこで，本研究では，実際の対話ログから応答となる疑問文を手書きで作成し，テンプレート生成の観点から分類する．その後，日本語形態素解析器 JUMAN++の辞書情報を外部リソースとして用いて，分類ごとにルールベースの手法を用いて作成したテンプレートを適用することで，文選択の選択肢となりうる候補文を簡易かつ大量に生成する手法を提案する．

本研究の貢献は，以下の通りである．

1. ユーザ発話に対して応答疑問文を人手で作成し，それらを応答可能性の観点から分類した
2. ユーザの 1 発話から，対話を続けるための日本語の対話応答候補文のテンプレートを作成した

連絡先: 長内 洋太，首都大学東京，osanai-youta@ed.tmu.ac.jp

2. 関連研究

テンプレートから日本語応答文を生成する研究として，加藤らの研究 [加藤 15] や山内らの研究 [山内 14] がある．

加藤ら [加藤 15] は，話題となる単語とともにユーザ発話内の感情を表現する語を抜き出し，話題語のカテゴリと感情の一致を考慮してテンプレートに挿入した．話題を一致させることは対話のロボット応答における違和感を低減することが示されたが，固有表現抽出の精度が悪いことや，感情語の活用や品詞を考慮していないことなどに問題があり，生成文の流暢性が低いという問題がある．本研究では，彼らのいう‘話題語’に着目し，どのような品詞が話題語に該当し，どのようなカテゴリであれば応答の生成にテンプレートを用いることができるかの分類を行い，自然な応答文が生成されるテンプレートを作成した．また，ユーザ発話への応答疑問文を事前に人手で作成し，単語を話題語として抽出すべきかどうかで分類した．

また，山内ら [山内 14] は，日本語文について，ある特定の分野に関する用語の情報が記述された用語データベースを用意し，ユーザとの会話の話題をその特定分野に誘導することで，テンプレートを用いた応答文生成を行なう．ただし，特定の分野の単語に対してのみ情報を所持しているため，他の分野の単語に対しては適切な応答文生成ができないという問題がある．そこで，本研究では，日本語形態素解析器 JUMAN++の辞書情報を用いて，ある分野の単語のみでなく，分野を限定せずに，日本語の応答疑問文を生成した．

3. テンプレートの生成

テンプレートによって応答疑問文を作成できるユーザ発話群の発見のため，対話データ中のユーザ発話の分類を行う．その後，応答疑問文を作成できるものに対して応答文を人手で作成し，作成した応答疑問文に対して，テンプレートを作成できる応答疑問文群発見のための類型化を行う．

応答疑問文の作成において，文脈を考慮する場合，応答疑問文の候補は多岐に渡る．したがって本研究では，文脈を考慮せず，ユーザの 1 発話のみを見るという制約のもとで応答疑問文を作成した．また，ユーザ発話のどの内容を受けた応答疑問文であるかを明確化するために，ユーザの使用した単語（見出

表 1: 各分類におけるユーザ発話の件数と例

分類	件数	ユーザの発話の例	作成した応答文
単語抽出が必要	733	**本**を読んだりラジオを聴いたりしてます。 特別な**習い事**はしてないです。	最近はどんな**本**を読まれましたか？ 何か**習い事**をしてみたいと思ったことはありますか？
単語抽出が不要	72	おはよう，今起きました。 ただいま帰りました。	おはようございます。今日は何をされるご予定ですか？ お帰りなさい、お疲れ様でした。
User-negative	125	ダメです、全く泳げません。 全く興味ありません、できません。	- -
User-question	54	あなたのお名前は？ はい、毎日してます。一緒にしませんか？	- -
単語抽出不可	150	そうですね。 うーん…。	- -
書き起こし文なし	66	-	-

し語）を一語以上使用することも制約として加えた．

以上の制約下で作成した応答疑問文を用いて，応答疑問文同士で類似した箇所を探し，テンプレートを作成する．

3.1 データセット

実際に用いる対話ログに関しては，先行研究 [下川原 16] で収集，編集したデータを扱う．このデータは，主にロボットが高齢者のユーザに対して質問を行い，ユーザがそれに返答をする雑談形式の対話ログである．ユーザの発話ログには，ユーザ発話を聞き取って書き起こされたものを扱う．本研究では，2014 年 6 月に収集した 6 ユーザの発話から各 200 文，計 1,200 文を抜き出し，それに対して分類を行い，疑問文を作成可能なユーザ発話に対して応答文を人手で作成した．

3.2 データセットの分類と応答疑問文作成

対話ログにおけるユーザ発話文の分析を行う．応答疑問文を作成するかどうかで場合分けし，それぞれに対しさらに分類を行う．応答疑問文の作成が可能である場合は，ユーザ発話から単語を抽出して応答疑問文を作成したか否かによって 2 つの分類に分かれる．疑問文の作成が不可，もしくは応答を作成すべきでないとき，これらは，‘ 単語の抽出が可能であるか不可であるかに関わらず，ユーザ発話が否定的な応答をしている ’，‘ 単語の抽出の可否に関わらず，ユーザ発話が疑問文である ’，‘ 抽出するべき単語がユーザ発話に含まれない ’，‘ ユーザ発話データが存在しない ’ の 4 つの分類に分かれる．

単語抽出が必要　ユーザ発話内に，ユーザとシステム間で今行われている会話の話題が読み取れる単語，主に自動詞が存在し，それを抜き出して用いることで応答疑問文が作成できる場合が該当する．

単語抽出が不要　主に ‘ おはよう ’ や ‘ ただいま ’，‘ ありがとう ’ などの挨拶や「言語行為の定型表現・運用上固着した語彙 [加藤 06]」，それにより応答が疑問文である必要がない場合などが，抽出の必要がない場合に該当する．

User-negative　ユーザが ‘ 嫌い ’，‘ 全くない ’ 等の語を含む否定的な発話をした場合が該当する．この場合，ユーザは 1 つ前でロボットが提示した話題に対して，会話の続行を望んでいないと推測できる．したがって，単語抽出によって疑問文の作成が可能な場合であっても，他の話題についての発話を行なった方が適切であると判断したため，疑問文を作成していない．

User-question　ユーザの発話が疑問文である場合が該当する．この発話に対する応答の作成には，ユーザやロボットのパーソナリティ情報や外部のリソース等から情報を取得する必要があり，ユーザ発話に含まれる単語をテンプレートに挿入して疑問文を作成することは適切ではない．

表 2: 各名詞の細分類内訳と例

細分類	該当する名詞の例	件数
普通名詞	服, メール, テニス	353
地名	日本, 東京, パリ	79
サ変名詞	会話, 経験, 予定	79
時相名詞	梅雨, この後, 最近	57
数詞	100, 何, 一	10
副詞的名詞	前, ところ	3
組織名	ユニクロ	1

単語抽出不可　感嘆詞のみであったり，聞き取りがうまくいかずに名詞等が途中で途切れている場合などの，主にユーザ発話に自動詞が存在していない場合が該当する．

書き起こし文なし　ユーザが実際に発話していない場合に加え，ノイズやユーザの声量の問題で，ユーザ発話が正しい文として書き起こされていない場合が含まれる．

以上の分類に基づき，‘ 単語抽出が必要 ’ と ‘ 単語抽出不要 ’ の応答疑問文を作成可能なユーザ発話に対して，疑問文を人手で作成した．各分類におけるユーザ発話の件数と，作成した応答を表 1 に示す．ユーザ発話における単語抽出を行う場合，ユーザ発話に含まれる単語を用いつつ応答を作成した．

3.3 抽出した単語の分類

テンプレートに挿入される頻度の高い品詞の特定のため，ユーザ発話の分類 ‘ 単語抽出が必要 ’ に対して作成した各応答文 733 件について，応答作成で抽出された単語の品詞ごとに分類を行う．ユーザ発話に含まれる 1 単語を抽出して作成した応答文について，日本語形態素解析器 JUMAN++（version 1.01）を用いて，抽出した単語の品詞分類を行った．

抽出した品詞を分類したところ，名詞 470 件（例：犬，ご飯，今日），動詞 153 件（例：起きる，食べる，行く），形容詞・形容動詞 110 件（例：静か，長い，若い）となり，抽出される単語の品詞は名詞が最も多かった．そのため，本研究では名詞に対するテンプレートを考えることにした．

名詞を用いて作成した応答文に対して，JUMAN++で解析した結果における，名詞の細分類，JUMAN++辞書における ‘カテゴリ’，‘ドメイン’ の両分類を用いることで，各名詞の属性付けを行う．ここで言及した JUMAN++のカテゴリとは，語の上位下位関係に基づくものであり，‘人’，‘動物’，‘植物’ 等 22 種存在する．ドメインは単語をオントロジー横断的に意味情報で区分したものであり，‘文化・芸術’，‘スポーツ’，‘健康・医学 ’ などの 12 種が存在する [黒橋 16]．

表 3: カテゴリ・ドメインの組の内訳と例 - 上位 10 件

カテゴリ名	ドメイン名	該当する名詞の例	件数
時間	-	昨日, この後	60
抽象物	-	天気, 流行	59
抽象物	家庭・暮らし	メール, 買い物	28
人工物-その他	家庭・暮らし	ゴミ, クーラー	23
人工物-食べ物	料理・食事	食べ物, 缶詰	23
人	家庭・暮らし	主人, 女の子	21
抽象物	文化・芸術	文化, 演奏	18
抽象物	スポーツ	テニス, スポーツ	15
人工物-その他	-	話題, ボールペン	13
動物	-	虫, ペット	13

作成した応答文のうち，名詞を抜き出して作成した応答文の名詞について品詞細分類を付与した内訳を表 2 に，カテゴリとドメインの組の出現頻度上位 10 件と該当する名詞の例を表 3 に示す．この中で，JUMAN++の解析結果において，名詞の解析候補が複数存在する名詞は，複数の分類のテンプレートに挿入することが考えられるため，それぞれの解析候補を別の名詞として考えて出現頻度を計算した．また，解析結果が一つであっても，複数のカテゴリやドメインを持つ単語に対しては，意味の詳細な絞り込みのため，その単語の持つ複数のカテゴリ，または複数のドメインの組み合わせを独立した分類として考える．また，カテゴリとドメインのどちらか，もしくはその両方を持たない語も存在した．

3.4 テンプレートの作成

ユーザ発話から，名詞を一つ抽出して挿入することで応答文となるテンプレートを作成する．

今回は，抽出した各名詞と述語の共起などは考えず，述語を各テンプレートに固定する形で作成した．このうち，作成できたテンプレートの該当する分類を表 5 に示す．表 5 において，`<word>`で示されるのが実際にユーザの発話から抜き出した単語である．

また，表 2 の‘サ変名詞’に該当する名詞は，それをする主体が必要であるという点で述語と扱いが類似しており，今回は挿入する名詞から除外している．

3.5 考察

内訳の上位に位置する，カテゴリが‘時間’や‘抽象物’である場合，ドメインが‘家庭・暮らし’である場合などの非常に広範に渡る分類では，カテゴリとドメインに加えて，名詞の細分類を用いても抽出した名詞の情報を絞りきれず，その分類に該当する名詞が適用できるテンプレートの考案が難しい場合がある．そのため，さらに細かい名詞の区分が必要である．

カテゴリ‘人’の分類では，‘妻’，‘嫁’などの名詞がよく見られた．これらの単語は，ユーザが発話者である場合と，システム発話の場合において名称が変わる名詞である．今回の手法ではこれらの単語への対応は難しいが，人称名詞の辞書を用いて活用することで，テンプレート作成が可能になると考える．

また，カテゴリ‘抽象物’に該当する名詞に見られるように，適切な述語が，抽出した名詞ごとに様々である分類において，述語をテンプレートに固定する今回の手法は適していない．このような場合，抽出した各名詞と共起しやすい述語を同時に挿入するテンプレートを作成することで，適切な出力文が得られる可能性がある．

表 3 において，‘家庭・暮らし’のドメインが頻出であることは，データセットが高齢者とロボットの日常的な対話ログであることに起因すると考えられる．

4. テンプレートを用いた応答文生成実験

3.4 節で考案したテンプレートに対して，ユーザ発話文から抽出した単語を挿入して応答文を生成し，自然な応答文が出力されるかどうかを検証する．今回は，表 4 に示す 3 つの分類のテンプレートに対して，対話ログデータ内の全単語を挿入して応答文を生成し，カテゴリのみ・ドメインのみ・両方を用いた場合の 3 つの場合において検証を行った．全体として，3 つの分類に属する名詞 118 語から，応答候補文 557 文が生成された．テンプレートに単語を挿入した出力例を表 5 に示す．

表 5 に示した例のうち，カテゴリとドメインの両方を用いて単語を挿入した場合，出力は非文ではなく，簡潔かつ応答文の範疇を逸脱していないといえる．また，カテゴリのみもしくはドメインのみを用いた場合の自然な文の出力割合が，該当するカテゴリ，ドメインによって大きく異なることがわかる．例として示した分類のうち，‘抽象物’というカテゴリのみを用いた場合は，この区分が非常に広範であるため，作成したテンプレートに適する単語を抽出できていないことが問題である．ドメイン‘料理・食事’については，ドメイン側が広範な区分であることと，‘冷蔵’や‘煮（る）’のような事態性名詞が普通名詞として挿入されることが問題となっていた．

しかし，広範な分類の名詞を挿入した場合など，依然として文として意味が通じないものが生成された．今回の設定では，全体の 8.98 %の出力が適切でない応答文であり，挿入した単語によって，その単語を挿入した応答文全てが不自然である場合と，ある特定のテンプレートに挿入した場合のみ不自然な文が出力された場合が存在した．適切でない出力の事例としては，以下の 3 つが見られた．

挿入する単語の粒度の問題　表 5 の‘スポーツの他に何かスポーツはされるのですか?’等の出力に見られる，該当の分類に属する単語同士の意味的な粒度が揃っていない場合である．この原因として挙げられるのは，カテゴリまたはドメインの意味的なばらつきが原因で，そのテンプレートに挿入するのにふさわしくない単語が挿入されてしまうことである．この問題に対しては，異なる辞書を用いることによって，単語の分類をさらに細かく行うことが必要がある．この問題については，テンプレート自体に含まれる単語や，それに類する単語に，フィルタをかけてテンプレートに挿入しない等の対策が必要である．

形態素解析の誤りによる問題　JUMAN++の解析結果が実際の単語の意味と異なる解析を行なった場合である．例としては，表 5 の細分類‘地名’における，‘将棋’のような単語である．この問題は，テンプレート作成の段階での対処が難しいため，異なる出力を行なった場合に，順次人手で JUMAN++辞書を修正，拡張する等の対応が望まれる．

外部知識や環境情報の問題　ロボットの応答が発話者にとって当然の事柄であり，会話の進行を妨げてしまう場合も存在する．例としては，日本人に‘日本にはよく行かれるのですか?’等の応答を出力する場合である．ただし，この場合を適切でないと判断するのは難しい．なぜなら，この分類の不自然な出力文は，ユーザやロボットの立ち位置や周囲の環境などの状況によって，自然な文になりうるからである．この場合，応答候補文の生成としては問題ないが，実際に候補文の中から応答文を選択する際に，ユーザの出身がどこであるかなどのユーザの登録情報や，ユーザとロボットが今どこで会話しているかなどの対話環境といった，追加情報の取得が必要になる．

表 4: 各分類における出力の統計

分類			カテゴリ・ドメイン			カテゴリのみ			ドメインのみ		
細分類	カテゴリ	ドメイン	単語数	出力文数	正文割合	単語数	出力文数	正文割合	単語数	出力文数	正文割合
地名	-	-	79	395	89.9	-	-	-	-	-	-
普通名詞	抽象物	スポーツ	15	90	98.9	296	1,939	9.3	16	96	92.7
	人工物-食べ物	料理・食事	24	72	87.5	38	114	61.4	61	183	38.8

表 5: 作成したテンプレートと対話ログから抽出した単語を用いた出力例.
（文頭に‘*’のあるものは適切でない出力例，‘?’のあるものは疑問の残る，または議論の必要性のある出力例）

分類			作成テンプレート	出力
細分類	カテゴリ	ドメイン		（**ゴシック**が挿入単語）
地名	-	-	<word>には行ったことはありますか？ <word>にはよく行かれるのですか？ <word>と聞くと何を思い浮かべますか？ ご自宅から<word>へはどう行かれるのですか？ <word>へ行ってみたいと思いますか？ <word>でよく行くお店はありますか？	*将棋に行ったことはありますか？ ?**日本**にはよく行かれるのですか？ **横須賀**と聞くと何を思い浮かべますか？ ご自宅から**新宿**へはどう行かれるのですか？ **パリ**へ行ってみたいと思いますか？ **新宿**でよく行くお店はありますか？
普通名詞	抽象物	スポーツ	<word>はよくされるのですか？ <word>の他に何かスポーツはされるのですか？ <word>の観戦に行かれたことはありますか？ <word>はいつ始められたのですか？ <word>を始めたきっかけは何だったのですか？ <word>をしていて楽しい時はどんな時ですか？	**テニス**はよくされるのですか？ ***スポーツ**の他に何かスポーツはされるのですか？ **サッカー**の観戦に行かれたことはありますか？ **ジョギング**はいつ始められたのですか？ **体操**を始めたきっかけは何だったのですか？ **野球**をしていて楽しい時はどんな時ですか？
	人工物-食べ物	料理・食事	<word>はよく食べるのですか？ <word>はお好きですか？ ご自分で<word>を作られたことはありますか？	**刺身**はよく食べるのですか？ **ご飯**はよく食べるのですか？ ?**食べ物**はお好きですか？ **ヨーグルト**はお好きですか？ ご自身で**漬物**を作られたことはありますか？ ご自身で **弁当**を作られたことはありますか？

5. おわりに

本研究では，ユーザの発話に対して人手で応答文を作成し，作成した応答文の単語を，日本語形態素解析器 JUMAN++の辞書におけるカテゴリやドメイン等の情報を活用することでテンプレートを作成した．ユーザの発話から文脈となる単語を抽出し，JUMAN++の辞書上で該当する各分類のテンプレートに挿入することで，その単語を使いつつ対話を続行するための文が得られることがわかった．

提案手法は，生成したテンプレートと単語の条件が合致すれば一定数の文が得られる一方，テンプレートを考案することが難しい単語や品詞においては，他の手法を用いた応答文の生成が必要になる．また，カテゴリとドメインの分類ごとにテンプレート数に偏りがあるため，一部のカテゴリやドメインにおいては十分な応答文が得られない問題や，カテゴリまたはドメインが非常に広範であるために，単語情報を絞り込めない場合，適切なテンプレートの作成が難しいという問題が存在する．

また，今回の手法では名詞のみに焦点を絞ったため，ユーザ発話が名詞を含まない場合を考慮できていない点が課題として残る．ユーザ発話が名詞を含まない場合には，今までの対話ログから会話の文脈を考慮しつつ，ユーザ発話に含まれる述語を抽出し，テンプレートに挿入することが考えられる．

今後は，提案手法に加えて，ユーザ発話から得られる情報のみならず，ロボットとユーザの位置関係や名前等のパーソナリティを利用したテンプレートによる応答生成と，意味的な観点からの客観的な評価を行う予定である．加えて，一つの名詞を抽出してテンプレートに挿入する手法において，挿入する名詞と関連性の高い述語を同時に挿入できるテンプレートを作成することで，さらに豊富な応答文を得る可能性についても探っていきたい．

参考文献

[Higashinaka 06] Higashinaka, R., Prasad, R., and Walker, M. A.: Learning to Generate Naturalistic Utterances Using Reviews in Spoken Dialogue Systems, *Proceedings of ACL*, pp. 265–272 (2006)

[Wallace 09] Wallace, R. S.: *The Anatomy of ALICE* (2009)

[Williams 07] Williams, J. D. and Young, S.: Partially Observable Markov Decision Processes for Spoken Dialog Systems, *Computer Speech & Language*, Vol. 21, No. 2, pp. 393–422 (2007)

[下川原 16] 下川原 （佐藤）英理, 篠田 遥子, 李 海妍, 高谷 智哉, 和田 一義, 山口 亨：高齢者と音声対話ロボットの雑談履歴の解析, *JRSJ*, Vol. 34, No. 5, pp. 309–315 (2016)

[加藤 06] 加藤 重広：日本語文法入門ハンドブック, 研究社 (2006)

[加藤 15] 加藤 和樹, 柴田 千尋, 田胡 和哉：文のテンプレートの学習および感情を考慮した会話文の生成, 情報処理学会第 77 回全国大会講演論文集, pp. 175–176 (2015)

[黒橋 16] 黒橋 禎夫：日本語形態素解析システム JUMAN++ version1.0 マニュアル (2016), http://lotus.kuee.kyoto-u.ac.jp/nl-resource/jumanpp/jumanpp-manual-1.0p2.pdf

[山内 14] 山内 祐輝, Neubig, G., Sakti, S., 戸田 智基, 中村 哲：対話システムにおける用語間の関係性を用いた話題誘導応答文生成, 人工知能学会論文誌, Vol. 29, No. 1, pp. 80–89 (2014)

[柴田 09] 柴田 雅博, 冨浦 洋一, 西口 友美：雑談自由対話を実現するための WWW 上の文書からの妥当な候補文選択手法, 人工知能学会論文誌, Vol. 24, No. 6, pp. 507–519 (2009)