

意味属性パターンを用いた マイクロブログ中の発言に対する自動対話行為付与

目黒 豊美^{1,a)} 東中 竜一郎² 杉山 弘晃¹ 南 泰浩¹

概要: Twitter等のマイクロブログには様々な情報があり、これらを分類する研究が多くされている。もし、マイクロブログに対話行為を付与することができれば、ユーザの間でのやり取りを分析したり、対話システムへの発話生成などに応用できる可能性がある。本稿では、マイクロブログ中の発言に教師あり学習の手法を用いて、対話行為を付与する手法を提案する。マイクロブログのデータは、対話データに比べ、多種多様な話題や語彙を含み、崩れた日本語が多い。そこで、我々は多種多様な語彙や崩れた日本語を補うために、シソーラスを用いて抽象化した単語 N-gram と文字 N-gram を特徴量として用いることを提案する。評価実験の結果、Bag-of-Ngrams を用いるベースライン手法に比べて、精度が高いことが分かった。

キーワード: 対話行為, 対話システム, マイクロブログ

Dialogue act tagging for microblog utterances using semantic category patterns

Abstract: In this paper, we propose dialogue act tagging for utterances in microblogs. The dialogue act estimator is built by using support vector machines (SVMs). To cope with the variety of words and expressions in microblogs, the feature vector uses N-grams of characters and words. In addition, the feature vector of word N-grams are abstracted into semantic categories by using a thesaurus. In our experiment, the proposed model outperformed naive baselines based on word N-grams.

Keywords: dialogue act, dialogue system, microblog

1. はじめに

最近、Twitterに代表されるマイクロブログサービスが様々なユーザによって盛んに利用されている [1]。マイクロブログは、通常のブログと比較して、短い文が投稿される。また、特定のユーザに向けた投稿ができ、まるで対話をしているようにユーザ間でコミュニケーションを取ることができる。マイクロブログの利用者はスマートフォンの浸透と共に、爆発的に増加している [2]。

本稿はマイクロブログデータに対話行為を自動的に付与することを目的とする。対話行為とは、質問や共感といった、ユーザの発話意図のことである。自然言語で表現され

たユーザの発言に対話行為を付与できれば、様々なアプリケーションに応用できる。例えば、東日本大震災の際には、Twitter上でリアルタイムに情報交換をユーザ同士で行うことで、被害情報や安全情報の共有がなされた。宮部ら [3] は、災害のような緊急事態における Twitter の利用傾向を分析することで、今後の緊急時に適切な情報を提供するシステム開発を行う上での知見を抽出している。この研究では、同じ発言を広めるリツイートや、特定の人へ返信するリプライの中に出現する単語をキーワードとし、分析を行っている。しかし、リツイートやリプライだけでなく、例えばあるキーワードを含み、かつ「質問」という対話行為である発言 (例: 仙台の〇〇町の被害状況はどうですか? 等) とそのリプライを抽出することができれば、誰がどのような情報を欲しいのか、それにどう答えるべきであるかということが分析できるだろう。そのような分析ができれば、災害時により適切な情報を提供するシステムを

¹ NTTコミュニケーション科学基礎研究所, NTT Communication Science Laboratories

² NTTメディアインテリジェンス研究所, NTT Media Intelligence Laboratories

^{a)} meguro.toyomi@lab.ntt.co.jp

構築することができる可能性がある。

また、ユーザの発言に対話行為を付与することで、非タスク指向型対話システム分野の研究にも役立てることができる。非タスク指向型対話システム研究は、タスク指向型対話に対比し、対話の話題、内容が多岐に渡るため、開発が難しい。また、人手でルールを書き尽くすのはコストが高い。そこで、Twitter やブログなど、ウェブから自動的に集められるデータを用いて、非タスク指向型対話システムを構築する手法が研究され始めている [4], [5]。Ritter ら [6] は、統計翻訳の手法を応用し、大量の発話対を用い、ユーザ発話をシステム発話に翻訳するという枠組みで生成を行っている。ここで発話対は、Twitter から自動的に大量に収集している。稲葉ら [7] は Twitter の発言中から対話の話題に沿った発言を抽出し、システムの発話とする手法を提案している。これらの研究では、システムの発話として選択された発言の発話意図となる対話行為を考慮していないが、自然な流れで対話を行うためには、システムの対話行為が重要となるだろう。もし、対話行為がマイクロブログデータに付与できれば、より適した発話を選択できることになる。

上記に述べた用途以外にも、質問と直後の意見を表す発話のみを抽出することで対話を要約したり、マイクロブログのマイニング用途として、特定の単語や商品名が入った質問だけを抽出し、どのような質問が増加傾向であるかを分析するなどの評判分析などに活かすという応用も考えられる。このように、マイクロブログの各発言に自動的に対話行為を付与することができれば、様々な応用が可能である。

本稿では、既に様々な先行研究で作成されている対話行為付与済みの対話データを用い、マイクロブログデータへ対話行為を付与する手法を検討する。ここで既存の対話データを用いる理由は、すでに対話行為を付与された対話データが多く存在しており、このようなデータは容易に入手することが可能だからである [8]。このような対話データをいことで、新たに Twitter へ人手で対話行為を付与する必要がなく、低コストで対話行為付与装置を構築できると考えられる。しかし、対話データとマイクロブログは一般に言語的特徴が異なることは否めない。例えば、既存の対話データは、初対面同士が対話する機会が多い。また、対話者は「対話に参加していない第三者(主に研究者)」が観測している状態を理解した上で対話を行っている。そのため、マイナーな話題は避け、丁寧な言葉で話している場合が多く、話題や語彙の豊富さは、マイクロブログに比べて小さい場合が多い。これに対し、マイクロブログは主に知り合い同士が多岐に渡る話題で自由に対話を行っており、崩れた日本語が用いられる場合が多い。そのため、対話データをもとにして Twitter に対話行為を付与しようとすると、既存の対話データでは補いきれない話題や語彙の

豊富さと、崩れた日本語を原因とした形態素解析誤りが問題となる。本研究では、これらの問題を解決するために、ソーラスを用いた単語 N-gram と文字 N-gram を特徴量として用いることを提案する。なお、本稿ではマイクロブログとして Twitter を用いる。

以降、2 節では先行研究を紹介する。3 節ではマイクロブログの語彙の豊富さと形態素解析誤りを補う特徴量を提案する。4 節では、提案手法とベースラインの手法を比較し、その結果と考察を述べる。最後に 5 節で、まとめと今後の課題について述べる。

2. 先行研究

対話文に対話行為を付与する従来手法は、大きく分けて、対話行為付与済みデータを用いた教師あり学習に基づく手法と、教師データを用いないクラスタリングに基づく手法にわけることができる。

対話行為付与済みデータを用いた教師あり学習に基づく手法としては、Stolcke らが、Dialog Act Markup in Several Layers (DAMSL) のタグセットを、決定木の手法を用いてタグ付する手法を提案している [9]。彼らは、音声対話をタグ付の対象としており、音声認識の結果や、韻律など音声特有の特徴量を用いている。磯村ら [10] は、独自に収集したテキスト対話の各発話に、DAMSL に基づくタグセットを自動的に付与している。特徴量は単語 unigram, bigram を用い、対話行為付与モデルを Conditional random field (CRF) で学習している。CRF は系列モデリングの手法の一つで、対話の流れを入力とし、出力として対話行為を付与することができるものである。比較実験によって、Support Vector Machine (SVM), Naive Bayes と比較し、CRF が学習手法として優れていたと報告している。

クラスタリングを用いた研究事例には、大竹の手法がある [11]。この研究は、K-means を用いた手法で、対話行為数となる K の値を事前に設定し、各クラスにコーパス中の発話を自動的に振り分ける。そして、各クラスを各対話行為を表すクラスとして考え、発話に対話行為として、クラスのラベルを割り当てる垣鏑ら [12] は Y-クラスタリングと呼ばれるアルゴリズムを用いて同様の分類を試みている。これらの研究では、データから自動的に対話行為のクラスを生成している。これらのクラスタリングによる手法は教師データが必要でないという利点があるものの、得られた分類の妥当性が必ずしも人間の直感と合わないところもある。そのため、本稿では、人手で予め設定した対話行為セットによって自動タグ付する手法を取る。

本研究と従来研究との違いは、我々の対話行為を付与する対象が、対話文ではなく、マイクロブログであるという点である。現在までに、既存の対話データから学習したモデルで、マイクロブログに教師あり学習の手法を用いて対話行為を付与する技術は見られない。

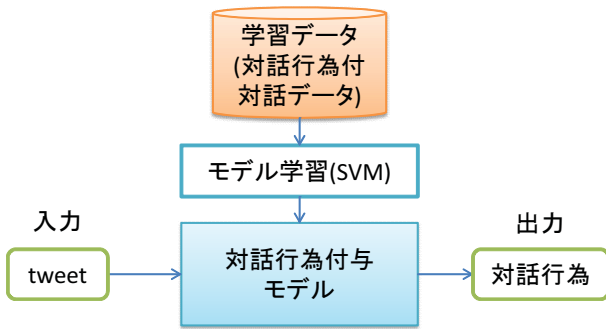


図 1 提案方法の学習と対話行為付与の流れ。
Fig. 1 The flow of the proposed method.

3. 提案手法

提案手法は、図 1 の流れで対話行為付与を行う。まず、対話行為付の対話データを学習データとして、後述の特徴量 (シソーラスを用いた単語 N-gram と文字 N-gram) を抽出する。学習機には SVM を用いる。磯村らは CRF が良いと報告しているが、多くの特徴量を取り入れやすいこともあり、本稿では SVM を用いる。

1 節で述べたように、Twitter へ対話行為を付与するモデルを既存の対話データから学習するため、対話データから取得できる素性で Twitter にも有効に機能する素性を設定する必要がある。ここで、Twitter データの豊富な語彙への対応と、崩れた日本語への対応の二点が問題となる。本稿では、それぞれシソーラスと文字 N-gram を用いることで対応する。

3.1 シソーラスを用いた特徴量

既存の対話データを利用するにあたり、Twitter の話題や語彙が豊富さが問題になる。限られた状況で収録された対話データは話題や語彙の豊富さでは、Twitter に比べて劣る。そのため、対話データを学習データとして利用しようとすると、語彙のサイズが小さくなり、対話データから抽出される特徴量だけでは Twitter には十分でない可能性がある。

語彙サイズの問題を回避するための方法として、シソーラスを用いて、単語の意味属性を単語の代わりに用い、抽象化して、特徴量とすることが考えられる。そこで我々は、シソーラスとして日本語語彙大系 [13] の意味属性を用い、抽象化された単語 N-gram を用いて特徴量を作成することを提案する。似たカテゴリの単語を、同じ特徴量として扱うことができれば、スパース性を軽減することができる。日本語語彙大系では、一般名詞と固有名詞に、2715 種類の意味属性が付与されている。例えば、「うどん」と「ラーメン」は同じ「麺類」に分類される。形態素の情報を用いてユーザの発話から特徴量を生成する際に、意味属性に抽象化することで、「うどん」と「ラーメン」を同じ

表 1 形態素解析の結果例。意味属性の列の括弧内の数値は、意味属性のクラス番号。

Table 1 An example of morphological analysis.

表記	品詞	標準表記	意味属性	抽象化
中学校	名詞	中学校	学校 (405)	学校
に	格助詞:連用	に	-	に
行	動詞語幹:IKU	行く	出発 (2132)	行く
き	動詞活用語尾	き	-	き
ます	動詞接尾辞:終止	ます	-	ます

表 2 1 から作成した特徴量の例。単語 N-gram は 1~2gram。文字 N-gram は trigram を使用した場合とする。

Table 2 An example of morphological analysis.

単語 unigram	{s}, 学校, に, 行く, き, ます, {/s}
単語 bigram	{s}-学校, 学校-に, に-行く, 行く-き, き-ます, ます-{/s}
文字 trigram	{s}-中-学, 中-学-校, 学-校-に, 校-に-行, に-行-き, 行-き-ま, き-ま-す, ま-す-{/s}

特徴量として扱うことができるようになる。

具体的な特徴量作成であるが、tweet 中の単語について、日本語語彙大系に意味属性が関連づけられている単語は、すべて意味属性に置き換える。ただし、動詞は数も少なく、また、対話行為を決定する重要な要素と考えられるため、意味属性による抽象化をしない。なお、動詞は終止形を用いる。意味属性が付与されない助詞などはそのまま用いる。形態素解析には、NTT の JTAG [14] を使用する。

「中学校に行きます」という文章を形態素解析すると、表 1 のように解析される。「抽象化」の列には、抽象化した後の特徴量が書かれており、「中学校」は「学校」を表す意味属性に抽象化されている。また、動詞は終止形を用いるため、「行」は「行く」として扱われる。この抽象化後の単語を用いて、単語 N-gram を作る。例えば、単語 bigram であれば、文頭と文末に、開始記号 {s} と終了記号 {/s} を挿入し、「{s}-学校」「学校-に」「に-行く」等とする。表 2 の単語 unigram, bigram が、抽象化後の単語から得られる特徴量の例である。

単語 N-gram の N の数の決定は、対話行為を決定するにあたり重要とされる、文末表現を網羅できるように通例より長めに設定する。ここで、文末表現とは、最後の内容語から文末までの語を指す。

3.2 文字 N-gram を用いた特徴量

既存の対話データは一般に不特定話者がテキスト対話を行ったものであるため、話し方は丁寧で、崩れた日本語は少ない。一方、Twitter は、知り合い同士がフランクに話しており、表現、話題などが多岐に渡る。崩れた日本語も多用されており、形態素解析の誤りが起こりやすい。例えば、実際の tweet である「みせたらーじゃないの」は、標準

表記として「店-た-ろ-う-じゃない-の」と解析され、「見せ」という単語が「店」と誤って解析されている。この例で単語 N-gram を抽出しようとする、「店-た-ろ」といった元の文章には含まれない特徴量が抽出されることになる。このような誤りは、マイクロブログを扱う上でよく起こるが、この問題を回避するには、形態素解析誤りに強い素性を抽出する必要がある。本稿では、この対処に文字 N-gram を用いることを提案する。文字 N-gram とは、対話行為を付与する文を文字ごとに分割して作成したものであり、形態素解析を行った結果を用いず、直接元の文章から抽出するため、解析誤りにロバストだと考えられる。また、対話行為を決定するために重要な文字列は機能表現が多く、名詞などの内容語に比べて数が限られるため、対話データだけでも、十分に学習できる可能性がある。文字 N-gram の N の値の決定に際しては、形態素解析器で用いる辞書内に含まれる内容語以外のエントリが十分カバーできるような N の値とする。

4. 評価実験

提案方法を検証するために比較実験を行った。SVM の学習には、ツールキットとして LIBLINEAR^{*1} を用いた。

4.1 実験の概要

対話行為付与済みの対話データのみを持っている場合の、Twitter への対話行為付与の精度を検証する。そのために、対話データを学習データに用いて、先行研究の手法と提案手法の比較を行う。

我々は対話データを学習データとする手法を提案しているが、この手法では、対話データに含まれる特徴のみしか扱われない。しかし、対話行為付与の対象である tweet には、対話データには含まれない tweet 独自の表現が含まれ、それが対話行為を決定する可能性もある。そこで、本手法の限界を知る意味で Twitter データを学習データとして学習を行い、対話データのみからでも十分に学習できているかを検証する。

4.2 実験データ作成手順

本実験に用いる対話データとして、先行研究 [15] で我々が収集した聞き役対話データ (以下、対話データ) を用いる。Twitter のデータとしては、Twitter の各文に対話行為を付与したデータを独自に作成して用いる。それぞれのデータの詳細を以下に示す。

対話データ 本データは、我々が「聞き役対話」と呼んでいる、二人の実験参加者がテキストチャットを行った対話で、一人がユーザの話を積極的に聞く聞き役、一人が任意の話題について話を行う話し役を行った対話

表 3 対話データと Twitter の対話数、発話数、アノテーションの一致度。一致度は、対話行為のコーエンのカッパ係数。

Table 3 Statistics of collected dialogues and dialogue-act annotation. Inter-annotator agreement means agreement of dialogue-act annotation using Cohen's κ .

	対話データ	Twitter
対話数	1260	1008
発話数	57262	2300 (その他を除く)
一致度 (κ)	0.53	0.54

である。完全に面識のない者同士が、別室に用意された PC で、テキストのみで対話を行ったものである。

Twitter データ Twitter をクロールし、大量に収集されたものの中から、in-reply-to で繋がっている発話列を抽出して Twitter データとする。これは、誰かが何かしらの返答をした対話的な tweet の方が、対話行為を付与するデータとして適切だと考えたためである。

対話データと Twitter データの詳細は表 3 の通りである。

4.3 対話行為付与

各データに人手でアノートを行った。まず、対話データの各発話および Twitter データの各 tweet を一人が一文毎に分割する。その後、本研究とは無関係な第三者の四人をアノータとして用意し、二人が対話データを、別の二人が Twitter データをそれぞれ独立に、先行研究で用いた 34 種類の対話行為を付与した (各対話行為の詳しい定義は、先行研究 [15], [16] を参考にされたい)。

Twitter データは、in-reply-to で繋がっている発話を時間順に並べ、対話になっている状態で対話行為を付与した。リツイートは「RT」となっている箇所から後ろをすべて「その他」とした。リプライは、ユーザアカウント名となっている「@アルファベットの羅列」部分は「その他」とした。

対話データと Twitter データの対話行為の傾向を見るために、出現分布を比較した。図 2 は各対話行為ごとの出現分布を示す。自己開示、質問という対話行為は本来それぞれ 9 種のサブタグを持つが、これら対話行為は、上位タグにまとめて集計した。ただし、質問のサブタグが自問であるものだけ別にした。これは自問が他の相手の発話を陽に求める発話とは性質が異なるものであると考えたためである。また、繰り返しと言い換えは発話意図が似ていると考えられるため合算した。

図 2 から、対話データと Twitter データはほぼ同ような分布であることがわかる。つまり、Twitter データでも、対話データと同じような対話が行われているということがわかる。しかし、一部の対話行為では、分布の違いも見られた。例えば、Twitter の方では、質問や相槌、共感が減っている。これらの対話行為は、対話を円滑に進めるために重要な対話行為であるが、Twitter は、対話を円滑に進め

^{*1} <http://www.csie.ntu.edu.tw/~cjlin/liblinear/>

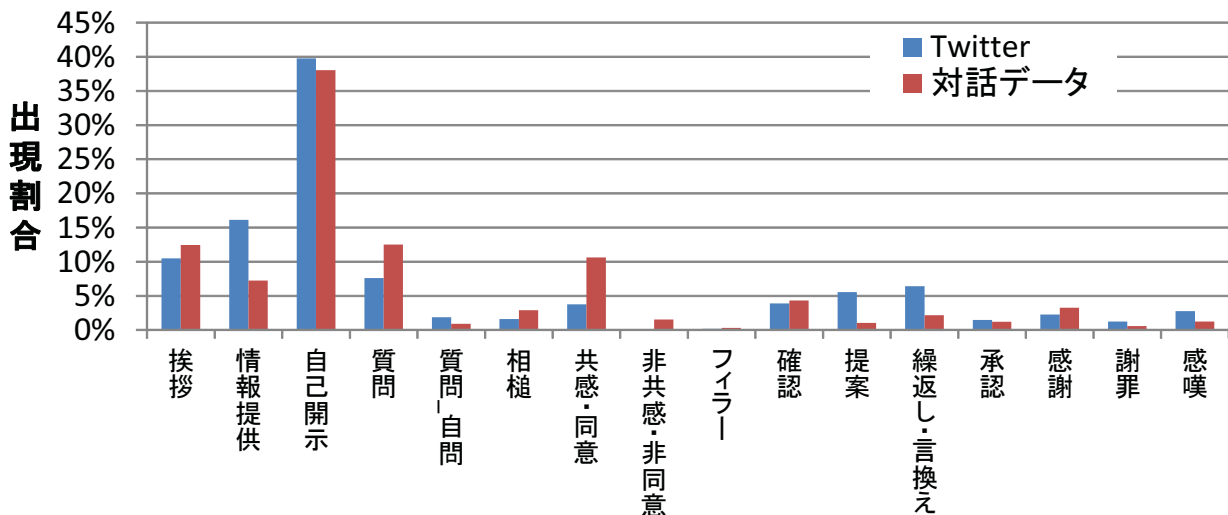


図 2 対話データの Twitter 内の対話行為の出現割合
Fig. 2 The proportion of dialogue act for each dataset.

る必要性は必ずしもないため、割合が減ったものと思われる。その代わりに、提案や繰返しが増えているようである。これは、

4.4 対話データを学習データに用いた実験

対話データから得られる特徴量が Twitter データにおいて、どの程度対話行為を付与できるかについて比較実験を行った。提案方法は、意味属性を用いた単語 unigram～7gram と文字 trigram～5gram を特徴量として用いた。ベースライン手法として、下記の二種類を用意した。

ベースライン 1 単語 unigram のみの特徴量として用いる手法。先行研究でよく使われている単語 N-gram の中でも最も単純なものである。先行研究では、このような単純な特徴量でも、対話データであれば、高精度で付与することができていると報告されている [10]。

ベースライン 2 文字 3-4gram と単語 1-2gram を用いる特徴量として用いる手法。ベースライン 1 より情報量を増やし、文末表現など対話行為を表す特徴的な表現を含めることができると考えられる。

4.4.1 学習データと評価データ

対話データは 55000 文のみ学習に用い、残りの 2262 文と、Twitter データ 2300 文をそれぞれテストデータとした。また、対話データの学習データとテストデータは異なるアノテータが作成したデータを用いることにした。なぜなら、今回、学習データに対話データ、テストデータに Twitter を用いる場合、アノテータが異なるためである。同じアノテータのデータを学習とテストの双方に用いた場合は、同じアノテータであるため、学習とテストは対話行為付与の精度の上限値が 100% となる。一方、異なるアノテータのデータを学習とテストに用いた場合は、異なるアノテータであるため、上限値は二者のアノテータの一致率

となる。そのため、精度を平等に比較できるように、学習データに対話データ、テストデータに対話データを用いる場合にも、異なるアノテータのデータを用いる。なお今回、「その他」が付与されたデータは除外されている。

4.4.2 結果

表 4 は Twitter をテストデータにしたときの精度と、対話データをテストデータにしたときの精度である。学習データはすべて同じ対話データである。上限値は、二人のアノテータの一致率と考えられるので、括弧内は、一致率を上限値であると考えたときの相対精度である。Twitter への対話行為付与において、ベースライン手法である「単語 unigram と提案手法」「文字 3-4gram 単語 1-2gram と提案方法」の間で、それぞれの正解数と不正解数に差があるか二項検定を行ったところ、有意な差が見られた ($p < 0.05$)。このことから、シソーラスと文字 N-gram を用いた提案方法はベースラインの手法に比べ、有意に優れており、上限値に対する相対精度も十分に高いことがわかる。

なお、対話データに対話行為を付与した際は、上記のような各手法間で有意な差は見られなかった。先行研究で扱われているような、初対面同士が、丁寧な言葉で話しているようなデータでは、単語 unigram といった少ない特徴量でも十分であるということが分かる。また、提案方法は、対話データにおいても精度を下げておらず、データの種類が変わっても頑健な手法であると言える。

4.5 Twitter データを学習データに用いた実験結果

対話データのみを学習データに用いても、提案方法で精度よく付与できることがわかったが、Twitter データを学習データとした場合に、どの程度の精度が実現できるかを調査した。特徴量は提案手法のものを用い、学習データとして Twitter を用いたときの精度を確認する。Twitter の

表 4 特徴量の比較結果. 上限値は, 二人のアノテータの一致率である. 括弧内は, 一致率を 100%の精度であると考えた際の一致率を表している.

Table 4 The comparison of the feature sets.

	単語 unigram	文字 3-4gram 単語 1-2gram	提案方法	上限値
Twitter	29%(48%)	31%(52%)	37%(62%)	60%(100%)
対話データ	44%(75%)	46%(78%)	45%(76%)	59%(100%)

表 5 学習データとテストデータに, それぞれ Twitter データと対話データを用いたときの精度.

Table 5 The result of the feature sets.

		テストデータ	
		Twitter	対話データ
学習データ	Twitter	33%	23%
	対話データ	37%	45%

学習データは 2000 文, テストデータに 300 文を用いた. 対話データはテストデータとして, すべて用いた. また, 学習データとテストデータは異なるアノテータが作成したデータを用いた.

表 5 は, 学習データに Twitter データを使った場合の実験結果である. データ数が少ないため, 対話データを用いたときと比較して対話行為付与の精度が低くなっており, また, Twitter の方が精度が高くなっている. 対話データを学習データに用いたときの結果は, 表 4 の提案方法を再掲したものである. こちらでは, 対話データをテストデータとしたときの方が, 精度が高い. 当然のことながら, 学習データとテストデータの種類が同じデータである方がよく, 二つのデータは違う特徴を持っている部分があることを示している. しかし, テストデータが Twitter の場合では, 学習データが対話データである方が精度が高い. これは, 学習データとして, 対話データの方が多いため, 対話データと Twitter データで共通の特徴を持っている部分が Twitter データだけのときに比べ, 多いためだと思われる. これらの結果から, 二つのデータには共通の特徴もあるが異なる特徴を持つ部分もあり, 共通の特徴部分は, 対話データからでも学習できていることがわかる. 今後は, 二種類のデータを混ぜて学習を行う方法を検討していきたい.

5. おわりに

本論文では, 対話データを学習データとしてマイクロブログに自動的に対話行為を付与する手法を検討した. マイクロブログを扱う上で, 多種多様な話題と言い回しに対応するために, シソーラスと文字 N-gram を保持する特徴量を提案した. 比較実験の結果, 提案手法の特徴量を用いた場合, 本手法は他の特徴量を用いた場合より, 有意に対話行為付与の精度が高いことがわかった.

今後は, 対話行為毎に付与する手法を変更したり, 少量のマイクロブログデータを活かして, 精度を向上させたい. また, 本手法を用いて自動的に対話行為を付与したマイクロブログデータを用いて, 非タスク指向型対話システムの発話を生成する研究を行いたいと考えている.

参考文献

- [1] Java, A., Song, X., Finin, T. and Tseng, B.: Why we twitter: understanding microblogging usage and communities, *In Proc. WebKDD/SNA-KDD workshop on Web mining and social network analysis*, pp. 56–65 (2007).
- [2] Gaonkar, S., Li, J., Choudhury, R. R., Cox, L. and Schmidt, A.: Micro-Blog: sharing and querying content through mobile phones and social participation, *In Proc. MobiSys*, pp. 174–186 (2008).
- [3] 真衣宮部, 英治荒牧, 麻子三浦: 東日本大震災における Twitter の利用傾向の分析, 情報処理学会研究報告. EIP, pp. 1–7 (2011).
- [4] Ritter, A., Cherry, C. and Dolan, B.: Unsupervised modeling of twitter conversations, *In Proc. NAACL*, pp. 172–180 (2010).
- [5] 別所史浩, 原田達也, 國吉康夫: リアルタイムクラウドソーシングと Twitter 大規模コーパスを利用した対話システム, 自然言語処理研究会 (SIG-NL), pp. 1–8 (2012).
- [6] Ritter, A., Cherry, C. and Dolan, W. B.: Data-driven response generation in social media, *In Proc. the Conference on Empirical Methods in Natural Language Processing*, pp. 583–593 (2011).
- [7] 稲葉通将, 神園彩香, 高橋健一: Twitter を用いた非タスク指向型対話システムの発話生成, 第 27 回人工知能学会全国大会 (オーガナイズドセッション) (2013).
- [8] Godfrey, J. J., Holliman, E. C. and McDaniel, J.: SWITCHBOARD: telephone speech corpus for research and development, *In Proc. ICASSP*, pp. 517–520 (1992).
- [9] Stolcke, A., Coccaro, N., Bates, R., Taylor, P., Van Ess-Dykema, C., Ries, K., Shriberg, E., Jurafsky, D., Martin, R. and Meteer, M.: Dialogue act modeling for automatic tagging and recognition of conversational speech, *Comput. Linguist.*, Vol. 26, No. 3, pp. 339–373 (2000).
- [10] 磯村直樹, 鳥海不二夫, 石井健一郎: 対話エージェント評価におけるタグ付与の自動化, 電子情報通信学会論文誌, Vol. 92, No. 11, pp. 795–805 (2009).
- [11] Ohtake, K.: Unsupervised Approach for Dialogue Act Classification., *In Proc. PACLIC*, pp. 445–451 (2008).
- [12] 垣鍔亮太, 南角吉彦, 李晃伸, 徳田恵一: 音声対話システムにおける発話文の自動クラスタリングに基づく応答選択 (話し言葉処理), 情報処理学会研究報告. SLP, 音声言語情報処理, No. 75, pp. 7–12 (2007).
- [13] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系, 岩波書店 (1999).
- [14] Fuchi, T. and Takagi, S.: Japanese morphological analyzer using word co-occurrence: JTAG, *In Proc. ICCL*, pp. 409–413 (1998).
- [15] Meguro, T., Higashinaka, R., Minami, Y. and Dohsaka, K.: Controlling listening-oriented dialogue using partially observable Markov decision processes, *In Proc. COLING*, pp. 761–769 (2010).
- [16] 目黒豊美, 南泰浩, 東中竜一郎, 堂坂浩二: 聞き役対話の分析及び分析に基づいた対話制御部の構築, 情報処理学会論文誌, Vol. 53, pp. 2787–2801 (2012).