

相手が求めるほめ言葉の自動生成

矢野 真由

Twitterなどのソーシャルネットワーキングサービス（SNS）を見ると、ほめて欲しいような投稿、共感を求める投稿を数多く見かける。SNS上に限らず、「ほめる」という行為は、人間関係を円滑に保つために非常に重要である。そして、何かほめて欲しいことがあったとき、「すごい」「いいね」など、一言で簡単にあしらわれるよりも、内容にまで踏み込んで、どこをどう良いと思ったのか、詳しくほめてもらった方が嬉しいというのは経験上明らかだろう。そこで、本論文では、人間同士の円滑なコミュニケーションに必要不可欠である「ほめ」に着目し、機械学習とルールベースの手法を併用することで、既存の対話システムよりも、より発話内容に踏み込んでほめることができるような対話システムの構築を行う。

機械学習を用いたほめ言葉の生成としては、Sequence to Sequence モデルを使用し、Twitterに投稿された、ツイートとリプライのペア 10 万件を用いて学習を行い、応答文を生成した。

ルールベースの手法を用いたほめ言葉の生成としては、まず、Twitterに投稿された 1,000 件の「ほめて欲しいようなツイート」と、「それに対するほめ言葉のリプライ」のペアに対し、「品詞割合分析」、「単語頻度解析」、「係り受け解析」、「関連研究に基づく考察」の 4 つの分析結果を行い、その結果をもとにテンプレートを作成した。そしてテンプレートに、ユーザーの入力から得た「話題語」「単語極性」「類義語」「対義語」を当てはめることで、ほめ言葉を生成する。

提案手法で生成されたほめ言葉に対し、「自然な文章か」、「ほめ言葉として適切か」、「ほめ言葉のバリエーションは豊富か」の 3 点について評価実験を行った結果、提案手法を用いると、約 70% の確率で文法的に正しい応答文を生成でき、そのうちの 65.7% が、効果的なほめ言葉となっていることが確認できた。

今後の課題としては Sequence to Sequence モデルの精度向上、形態素解析のみではなく、文中の単語の共起頻度なども使用した文章の評価、使用する単語極性辞書の語彙数を増やすことなどが挙げられる。

このような「ほめ」に焦点をあてた対話システムは、対話システムとしてのみでなく、「ほめるのが苦手な人にほめ言葉の例を提示する」といったような、対人コミュニケーションの円滑化のためのツールや、オンライン教材における講師から生徒への「ほめ」など、様々な分野への応用も期待できるだろう。

(指導教員 手塚太郎)