

Web マイニングを用いた文章中の行為者及び被行為者の感情推定

堂腰裕明[†] 小山聡[†] 栗原正仁[†] ミハウ プタシンスキ^{††} ジェプカ ラファウ^{†††} 荒木健治^{†††}

北海道大学大学院情報科学研究科 複合情報学専攻 表現系工学研究室[†]

北海学園大学ハイテク・リサーチ・センター^{††}

北海道大学大学院情報科学研究科 メディアネットワーク専攻 言語メディア学研究室^{†††}

1 はじめに

コンピュータの能力が向上するに伴い、その利用方法もますます拡大されてきている。コンピュータに関する基本的な知識を持たない人もコンピュータに触れることが増えたため、人とコンピュータをつなぐインタフェースにも、人とコンピュータ間におけるより柔軟なコミュニケーションを実現する能力が求められるようになっていく。

その一つの方向性として、文章表現から人の感情を読み取るというものがある。

文章中に記された動作や会話から、その動作主や会話主などの、いわゆる動作主体の感情を推定するといった研究が多く行われている[1][2]。コンピュータに人間の感情情報を読み取らせることで、読み取った感情情報に応じて、ユーザのニーズにより柔軟に対応可能になることが考えられる。

各々の動詞、形容詞などについて、その主体のどのような感情を表現するかという情報を、各動詞や形容詞ごとに考えるという観点においての研究が見られる。これに関して、ある動作や言動に付随して発生するものとしてその主体の感情がある一方で、主体の動作の対象となった被行為者の側にも、その動作の結果として何らかの感情が発生することが推察できる。例えば文中に「殴る」という動作が存在した場合、その行為者のみならず、被行為者の側にも何らかの感情が発生することが考えられる、といった具合である。

ここでは Web マイニングを用いて、文章中の行為表現の能動態および受動態と、感情を示す語のデータベース内の単語との共起関係を調べることで、行為者及び被行為者の感情を推定する手法を提案する。

2 関連研究

本研究は、以前の研究[3]において行った、童話などの物語文における登場人物の感情の推定について、新たな手法の一つとして提案するものである。

ブログの文章やチャットの話し言葉など、その

文章の話者が定まっておき、その話者の感情を推定する場合に比して、物語文の登場人物の感情推測は、複数の登場人物の関係性にも左右され、また、文中の登場人物の動作により関係性そのものも変化していく。

このため、一つの動作から、その動作の行為者の感情のみではなく、動作の対象となった被行為者の感情も同時に推測することで、物語文における登場人物の感情の読解をより効率よく行えると考えられる。

本研究において使用している『ML-Ask システム』[4]についても説明する。

ML-Ask は『感情表現辞典』[5]を利用して、感情をカテゴリ化している。感情をカテゴリ化するのは、人間の感情という情報を、コンピュータが扱いに優れた数値へと変換することで、コンピュータが感情という情報を扱うことが可能となるだろうという考えによる。

感情表現辞典では、「喜・怒・哀・怖・恥・好・厭・昂・安・驚」という 10 種類の感情各々の表現が収集されている。カテゴリ化に関しては、ブルチックの立体モデルや Ekman の表情に基く 6 分類など様々なものがあるが、施らの手法では、この 10 種類のカテゴリを用いて感情を分類している。感情表現辞典は日本語の言語表現を前提とした感情分類であり、日本語が関連する研究に適していると考えられる。本研究においてもこの分類を利用し、人間の感情をこれらの 10 種にカテゴリ化するものとする。

感情表現辞典の分類に基づき、ML-Ask システムでは、各々の感情状態を表す表現をデータベース化し、感情表現データベースとして用いている。本研究では、このデータベースを利用する。

3 提案手法

文章中に表れる語に対し、検索エンジンを用いて、ML-Ask の感情表現データベース内の各表現との and 検索を行い、得られた検索件数を記録する。

また、感情表現データベース内の各表現に対してそれぞれ単一で検索エンジンにかけ、同様に検索件数を得る。

これらの比から、文章中の語に感情表現データベース内のどの表現との共起が見られたか、また、

[†]dokoshi@complex.hokudai.ac.jp

^{††}札幌市北区北 14 条西 9 丁目北海道大学大学院情報科学研究科

どの感情カテゴリ内の表現との共起が最も頻繁に見られたかが観察可能である。ここでは検索エンジンとして、Bing の検索 api[6]を用いる。

これらの処理を、行為者と被行為者が存在すると判断できる「動作」を示す語において、その語の能動態及び受動態に施し、結果をデータベース化することによって、文章中の一つの表現から、その行為者と被行為者両方の感情を推測するシステムを提案する。

参考文献

- [1] 松本和幸, “会話文からの話者感情推定に関する研究”, 人工知能学会誌 24(1), pp120(2009)
- [2] 熊本忠彦, 田中克己, “Web ニュース記事を対象とする喜怒哀楽抽出システム”, 情報処理学会インタラクシオン 2005, pp25-26(2005)
- [3] 堂腰裕明, 小山聡, 栗原正仁, “シソーラス利用による童話文からの感情情報読み取りシステム” 情報処理学会全国大会講演論文集第 72 回平成 22 年(2), pp.361-362(2010-03-08)
- [4] Michal Ptaszynski, Pawel Dybala, Rafal Rzepka and Kenji Araki, “Affecting Corpora: Experiments with Automatic Affect Annotation System - A Case Study of the 2channel Forum -”, The Conference of the Pacific Association for Computational Linguistics (PACLING-09), September 1-4, 2009, pp. 223-228.
- [5] 中村明, “感情表現辞典”, 東京堂出版, 1993
- [6] BingDeveloperCenter.
<http://www.bing.com/developers>