

重文や二重否定を考慮した文の感情認識

桶本夏輝 長名優子

東京工科大学 コンピュータサイエンス学部

1 はじめに

人間は会話を行うときに話す声のトーンや話者の表情などから感情を読み取るということを普通の会話の中で行っている。しかし、文字ベースのコミュニケーションツールでは文章から感情を読み取ることしかできず、正しく感情を伝えられないこともある。また、対話システムにおいてもユーザの感情を正しく認識していないと適切な応答文が作成できず、会話そのものがちぐはぐになってしまうことも考えられる。非タスク指向型対話システムにおけるユーザとの会話を円滑に行う手法の一つとして、ユーザの嗜好や人間関係を用いた非タスク指向型のシステム [1] が提案されている。非この手法では、ユーザの発話内容から、形態素解析や構文解析を行い、深層格、極性、素性などの情報を抽出してユーザの嗜好を推定し、その情報をもとに応答文を作成して対話を実現している。また、文に対する感情認識の精度を向上させる手法として、重文や二重否定表現を正しく処理できるように考慮した構文解析の手法も提案されている [2]。この手法では、ML-Ask [3] を用いて感情認識を行う際に、特殊な表現である重文や二重否定表現を考慮することで感情認識の精度を向上させている。

本研究では、重文や二重否定表現を考慮した文の感情認識を提案する。

2 重文や二重否定を考慮した文の感情認識

ここでは、提案する重文や二重否定を考慮した文の感情認識について説明する。提案手法では、形態素解析器 MeCab、構文解析器 CaboCha を用いて、入力文解析を行う。その後、その情報を用いて、重文および二重否定文の判定を行う。その後、ML-Ask や自作する感情値辞書を用いて、感情推定を行う。

2.1 重文・二重否定の判定

Emotion Recognition of Sentences considering Compound Sentence and Double Negative Expression.
Natsuki Okemoto and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

重文とは2つ以上の文がつながって1つの文になっている文である。提案手法では、2つ以上の文が接続助詞を用いてつながっている重文を対象として判定を行う。二重否定文とは、1文に否定表現が2つ存在している文のことである。否定表現には、曖昧の意味を伴うものと強調の意味を伴うものの2種類が存在する。二重否定文には様々な表現方法が存在しているため、対応する否定文を絞るために類型化を行う。

2.2 感情推定

提案手法では、ML-Ask を用いた感情推定と fast-Text を用いて生成した感情値辞書を用いた感情推定の結果を合わせて判断する。

2.2.1 ML-Ask を用いた感情推定

ML-Ask [3] は、感情表現辞書 [5] に基づいて作成された感情表現データベースを用いて、Russel の感情円環モデルを実装した手法である。このシステムでは、入力されたテキストデータを文が感情的であるか判定する。感情的であった場合には、そのテキストデータに対して感情タイプを認識する。感情タイプは「喜・怒・昂・哀・好・怖・安・厭・驚・恥」の10種類の中から最も適していると判断された感情が出力される。

このシステムは、CVS (Contextual Valence Shifters) の概念を実装しており、否定と強意語の2種類のCVSを想定している。否定の場合には入力された単語の意味的な極性を変更している。また強意語の場合には、評価単語の意味的方向性を強化している。このシステムは108の構文否定構造を持つ否定タイプのCVSを組み込んでいる。さらに、このシステムはRusselの感情円環モデルを実装している。Russelの円環モデルとはRusselが提唱した感情が、快 ↔ 不快と覚醒 ↔ 沈静の2つの軸を持つ2次元の平面上の円環に配置されているとするモデルである。このモデルでは似ている感情は、平面上の近い位置に配置され、反対の感情は平面上の原点に対して点対称な位置に配置される。

提案手法では、入力文に対してML-Askを用いて感情推定を行う。

2.2.2 fastText を用いて生成した感情値辞書を用いた感情推定

fastText[4] を用いた感情値辞書の生成は以下の手順で行う。

Step 1: 感情表現辞書 [5] を参考に、喜・怒・昂・哀・好・怖・安・厭・驚・恥の 10 種類の感情を表す単語を感情ごとに複数個選択する。

Step 2: 学習済みの fastText モデルを用いて、Step 1 で選択した単語に対する分散表現を取得する。

Step 3: 感情値を求めたい単語の分散表現と Step 2 で求めた分散表現との類似を求め、感情ごとの類似度の平均値を算出する。

Step 4: Step 3 で求めた感情ごとの類似度の平均値に基づいて感情値を決定する。

2.2.3 生成した感情値辞書を用いて感情推定

感情値辞書を用いた感情推定は以下のような手順で行う。

Step 1: 感情値の算出

各感情の感情値を算出する。文 s の感情 e に関する感情値 E_e^s は

$$E_e^s = \sum_{w_i^s \in C_e} C^s(w_i^s) E_e(w_i^s) \quad (1)$$

で与えられる。ここで、 w_i^s は文 s の単語 i 、 C_e は感情 e の感情値を持つ単語の集合を表す。 $C^s(w)$ は文 s に単語 w に関する係数であり、これは曖昧表現や強調表現などを反映するのに用いる。また、 $E_e(w)$ は単語 w の感情 e に対する感情値を表している。この式では、文 s の感情 e に関する感情値を文に含まれる感情 e の感情値を二重否定表現の曖昧や強調などを反映する係数で重みづけして足し合わせたものとして表している。文中に否定表現が存在した場合には Russel の感情円環モデルに基づいて対になる感情が存在する場合には、反対の感情の感情値を使用する。ここでは、喜 ↔ 哀、好 ↔ 厭、安 ↔ 怖を対になる感情として扱うものとする。

Step 2: 感情の決定

式 (1) で求めた文 s に関する感情値に基づいて文に対応する感情を決定する。文 s の感情 e_s^{DIC} は

$$e_s^{DIC} = \begin{cases} \operatorname{argmax}_e E_e^s & (\max_e E_e^s \geq \theta) \\ \text{なし} & (\max_e E_e^s < \theta) \end{cases} \quad (2)$$

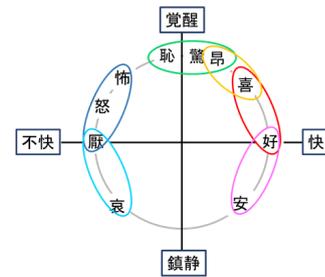


図 1: Russel の感情円環モデルと感情群

で与えられる。ここで、 θ はしきい値を表している。この式は感情値が最大のものがしきい値以上の場合には最も大きい感情値の感情を文の感情として扱うということを表している。感情値が最大のものがしきい値以下の場合には文の感情は「なし」となる。

2.3 最終的な感情推定結果の判定

提案手法では、最終的に ML-Ask を用いた感情推定と感情値辞書を用いた感情推定の結果を合わせて判断する。推定結果が一致していた場合には、そのまま結果を出力する。結果が一致していなかった場合でも、感情が矛盾していないと考えられる場合には両方の結果を出力する。ここでいう矛盾していないというのは、図 1 の Russel の感情円環モデルにおいて色つきの楕円で囲っている感情群を指す。

3 計算機実験

提案手法を用いて実験を行い、文の感情推定が行えることを確認した。

参考文献

- [1] 小林峻也, 萩原将文: “ユーザの嗜好や人間関係を考慮する非タスク指向型対話システム,” 人工知能学会論文誌, Vol.31, No.1, 2016.
- [2] 滝沢満, ラファウ ジェブカ, 荒木健治: “重文および否定形を考慮した感情認識システムの改善,” ファジィシステムシンポジウム講演論文集, 2013.
- [3] <http://arakilab.media.eng.hokudai.ac.jp/ptaszynski/repository/mlask.htm>, 「Michal Ptaszynski/Research」, (2022.01.05 参照).
- [4] P. Bojanowski, E. Grave, A. Joulin and T.Mikolov: “Enriching word vectors with subword information,” Transactions of the Association for Computational Linguistics, Vol.5, pp.135–146, 2017.
- [5] 中村明: 感情表現辞典, 東京堂出版, 1993.