

研究背景と目的

文の極性を判定する際、極性辞書を用いて単語の極性を同定するが、極性が判定できない文や、極性が矛盾する文が存在する。
ゴルフ場レビューの中で極性の矛盾が生じる文と極性判定できない文について、正しく極性を判定する。

コースはとても良かったです。→コースはとても良~~か~~ったです。→Positive

コースは難しくて良かったです。→コースは難しくて良~~か~~ったです。→極性が矛盾

極性辞書[1] に記載されている単語		極性辞書[1]に記載されていない単語
文の極性が矛盾しない単語	文の極性が矛盾する単語	文の極性が決定できない単語
• コースも難しく、苦勞しました。 • 非常に戦略的なコースで難しいです。 • コースも長くて楽しかったです。	• コースは難しく、楽しいです。 • コースも難しくて面白い。	• 距離が長く、グリーンはしっかり手入れされています。 • グリーンがもう少し早けれればと思いました。

提案手法

極性辞書[1]

レビュー文を入力

初期極性の付与

文脈一貫性に基づく極性修正

評価情報抽出

極性の出現割合に基づく極性修正

コースの難しさも楽しい物です。

コースの難~~し~~さも~~楽~~しい物です。
Negative Positive

コースの難~~し~~さも~~楽~~しい物です。
Negative Negative

コースの難~~し~~さも~~楽~~しい物です。
Positive Positive

コースの難しさも楽しい物です。
→ (コース, 難しい),
(コース, 楽しい)

コースは広いです。

コースは~~広~~いです。
?

距離は短~~い~~ けど, ~~良~~い所に有って面白かった
? 逆接
距離は短~~い~~ けど, ~~良~~い所に有って面白かった
Negative Positive
距離は短~~い~~ けど, 良~~い~~所に有って面白かった
Negative

係り受け関係から評判を抽出
OBが~~多~~く、~~難~~しいコースに感じました。
→(OB, 多い), (OB, 難しい), (コース, 難しい)

属性	品詞パターン	例
属性	名詞 (一般, 固有名詞, サ変接続)	コース, OB
	形容詞 (自立)	多い, 難しい
表現	動詞(非自立)+形容詞 (自立)	読みやすい, 悪くない

表3. 評価情報極性の出現回数

属性	表現	Positive	Negative	Unknown	Total
コース	難しい	748	629	1604	2981
コース	楽しい	823	158	0	981
コース	広い	211	175	394	780

$$eval(w) = \frac{pos(w) - neg(w)}{total(w)}$$

$eval(w) > 0 \dots Positive$

$eval(w) < 0 \dots Negative$

Pos(w) ... Positiveと判定された評価情報wの回数

neg(w) ... Negativeと判定された評価情報wの回数

total(w) ... Positiveと判定された評価情報wの回数

表4. 評価情報と修正後の極性

属性	表現	極性
コース	難しい	Positive
コース	楽しい	Positive
コース	広い	Positive

実験結果

• 楽天GORAレビューデータ[4]を使用

• レビュー数の多い順30コース(約17,000文)を対象に実験

• 文脈一貫性に基づいて極性を判定

• 極性辞書[1]に定義されていない表現を調査

図3. 評価情報に基づく極性推定結果

• “フェアウェイも広く、ナビも便利で楽しめました。”
→Negativeと判定

• 極性辞書[1]では「便利」がNegativeだと判定されている

• 極性辞書[1]のすべての値を利用したことが誤判定の原因だと考えられる

表6. 極性判定できない評価情報と文

属性	表現	極性判定できない文
フェアウェイ	広い	フェアウェイは広く、OBも少ないので、飛ばし屋の男性には好評でした。
距離	短い	フェアウェイも広く距離も短い気がしました。
風	強い	きれいなゴルフ場ですが、風が強くて大変でした。

• 今後は他のレビューについても極性の判定を行う

参考文献

[1] 高村大也, 乾孝司, 奥村学, "スピンモデルによる単語の感情極性抽出", 情報処理学会論文誌, Vol.47 No.02 pp. 627-637, 2006.
[2] 那須川哲哉, 金山博, "文脈一貫性を利用した極性付評価表現の語彙獲得", 情報処理学会研究報告自然言語処理 (NL), vol. 2004, no. 73 (2004-NL-162), pp. 109-116, 2004.
[3] 平山拓央, 湯元高行, 新居学, 高橋豊, "属性評価モデルに基づく商品評価の抽出と提示", DEIM Forum 2011, F2-5, 2011.
[4] 楽天グループ株式会社 (2010): 楽天GORAデータ. 国立情報学研究所情報学研究データリポジトリ. (データセット).
<https://doi.org/10.32130/idr.2.3>

謝辞

本研究では国立情報学研究所のIDRデータセット提供サービスにより楽天グループ株式会社から提供を受けた「楽天データセット」(https://rit.rakuten.com/data_release/)を利用しました。感謝申し上げます。