

情緒推定のための名詞句の評価極性

竹本 翔[†] 徳久 雅人[†] 村田 真樹[†]

鳥取大学大学院工学研究科 情報エレクトロニクス専攻[†]

1 はじめに

情緒推定のために、用言の語義に注目した情緒属性付き結合価パターン辞書の構築が行われ [1], 情緒が成立するための判断条件の改良が行われた [2]. 判断条件は、用言の語義だけでは不足する情緒成立条件を補うものであり、情緒推定の精度に影響する. 判断条件の充足判定には名詞句の評価極性値を用いる方法 [3] があるが、「A の B」型名詞句の評価極性を算出する際、名詞 A, B の *SO-score*[4] の単純な足し合わせで判定したため評価極性判定精度が悪かった. また、状況による評価極性の変化を考慮していない (例えば「冷たいかき氷を食べる」ことは、夏は嬉しいが冬はそれ程でもない). 従って、状況ごとの評価極性が必要となる.

そこで本稿では、名詞句の種類、判断条件、状況を考慮した評価極性を得るために、*SO-score* 算出用の極性表現とコーパスを可変化することを提案する.

2 パターン辞書を用いた情緒推定

2.1 基本構造

情緒属性付き結合価パターン辞書のレコードの1つを表 1 に示す. 日本語パターンと意味属性制約は用言の語義を表す. それ以外の5項目は情緒属性を表し、そのうち判断条件は情緒原因に比べて用語の語義に不足する意味を、情緒原因は情緒が生起する原因を、情緒名は生起する情緒の候補をそれぞれ表す.

入力文に対する情緒推定では、まずこの辞書に入力文を照合し、日本語パターンが適合し、かつ意味属性制約が充足するレコードを検索する. 次にそのレコードの判断条件の真偽を判定し、判断条件が真ならば情緒名などの情緒属性を出力し、偽ならば情緒《なし》を出力する.

本稿では、名詞単語変数 N のかわりに、修飾語句を含むことのできる名詞句変数 NP を用いる.

表 1 情緒属性付き結合価パターン辞書の一部

◆日本語パターン: $N1$ は/が $N2$ を食べる
◆意味属性制約: $N1$ (人 動物) $N2$ (食料 生物)
◆判断条件: 生理・近 ($N1, N2$) ◆情緒原因: 〈五感 (味覚)〉
◆情緒主: $N1$ ◆情緒対象: $N2$ ◆情緒名: 〈好ましい〉

2.2 判断条件と評価極性の関係

判断条件には生理・近, 生理・離のように、接近と乖離の関係を表したものがある. この判断条件の充足判定に、判断条件の第二引数 $NP2$ の評価極性を用い、表 2 のように判定する. 例えば、 $NP2$ の評価極性が好評のとき、生理・近の真理値は T となる.

判断条件の判定は、情緒主の個別事情を配慮する場合と、一般的な情緒主を想定する場合とに分かれる. 例えば、「生理・近 (竹本翔, ストレートのウイスキー)」は個別事情を知りながら解く場合で、個人ごとのデータベースを要する. 「生理・近 (ϕ , ストレートのウイスキー)」は一般の人を想定する場合で、*SO-score* で解くことを試みる. 本稿は後者の位置付けである.

表 2 判断条件の真理値表

判断条件	$NP2$ の評価極性	
	好評のとき	不評のとき
生理・近	T	F
生理・離	F	T
心理・近	T	F
心理・離	F	T
目標実現・近	T	F
目標実現・離	F	T
対人関係・近	T	F
対人関係・離	F	T
対人関係 (上下)・近	T	F
不要	T	T

3 評価極性算出方法の提案

3.1 *SO-score* 算出のための拡張

複数の単語に対する評価極性を算出するため、*SO-score* 算出式の変数をそれぞれ表現集合 T , 好評表現集合 P , 不評表現集合 N に拡張する.

$$SO-score(T) = PMI(T, P) - PMI(T, N) \quad (1)$$

$$PMI(X, Y) = \log_2 \frac{p(X, Y)}{p(X)p(Y)} \quad (2)$$

$$p(\{x_1, x_2, \dots, x_n\}) = \prod_{i=1}^n p(x_i) \quad (3)$$

$$p(\{x_1, \dots, x_n\}, \{y_1, \dots, y_m\}) = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \quad (4)$$

ここで、表現集合 T は、「A の B」型名詞句や、「形容詞 A 名詞 B」型名詞句の場合は、 $T = \{A, B\}$ となる. 好評表現集合 P , 不評表現集合 N は、名詞の種類、判断

Polarity of noun phrases to infer emotions

[†]Department of Information and Electronics, Graduate School of Engineering, Tottori University

条件, 状況の3つで極性表現対知識ベースを参照して求める(3.2節). $p(x)$, $p(x, y)$ はフレキシブルなコーパス(3.3節)から求まる x , y の出現確率である. 評価極性は図1の流れで算出する.

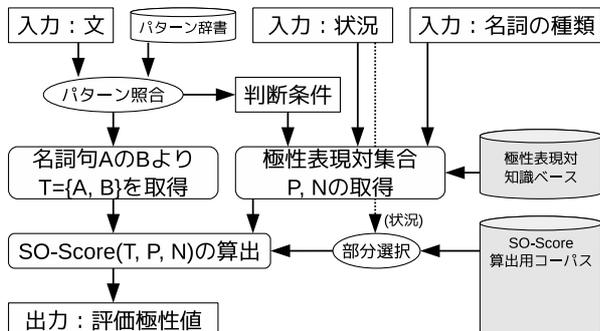


図1 評価極性を算出するための処理のフロー

3.2 極性表現対集合の取得

SO-scoreを算出する極性表現対には「良い」、「悪い」のような極性表現対がよく使われる. 本稿では名詞の種類, 判断条件, 状況の3つの情報で極性表現対知識ベース(4章で構築)を参照し, 極性表現対を決定する.

3.3 SO-score算出用コーパスと部分選択

SO-scoreで共起頻度を得るためのコーパスには2012年のブログコーパス1年分約4千万文を用いる. 状況を「夏」とすると7, 8月分の約7百万文に, 「冬」とすると12, 1月分の約6百万文にそれぞれ限定できる.

4 極性表現対の知識ベース化

SO-Scoreの好評/不評表現集 P/N を収録する知識ベースを構築する.

本知識ベースは, 名詞の種類, 状況, および, 判断条件に応じて好評/不評表現を提供する. 本稿では, 衣食住の名詞, 「夏/冬」の状況, および, 「生理, 目標実現, 対人関係」を想定して表現集を手作業で作成する.

ここで, 「想定」とは, 表現を発想するためのきっかけである. 例えば, 夏に食品を誰かにプレゼントするならば「対人関係」の想定は重要なので, 贈り物は, 「涼しい, 美味しい」の他に「高級」も考え, そして, 「不味い」や「粗末」というものを避けるように考える.

まずは, こうして直感的に表現集を作成する. 次のように条件とともに列挙できる: $(P, \text{夏}) = \{\text{涼しい}\}$, $(P, \text{食}) = \{\text{美味しい}\}$, $(P, \text{対人関係}) = \{\text{高級}\}$, $(N, \text{夏}) = \{\text{熱い}\}$, $(N, \text{食}) = \{\text{不味い}\}$, $(N, \text{対人関係}) = \{\text{粗末}\}$. 次に, 条件の重なりや一般化を考慮しながら構成する. 「夏」と「食」を同時に想定するような表現には「あっさり」などがあり, 作成の際に考慮に入れた. また, 「美味しい」は, 「夏」と「冬」に共通するものなの

で, 状況には「一般」というカテゴリを追加した.

以上の結果, 好評表現として105件, 不評表現として105件を作成した.

5 実験

5.1 目的と手法

提案手法により, 名詞句の評価極性が正しく推定できる割合を求める. 提案手法では, 次の2つの機能を有する評価極性の算出を行う.

- f1. 極性表現対選択機能: 名詞句の種類, 判断条件, 状況に基づき極性表現対集合を取得
- f2. コーパス絞り込み機能: 状況に応じてフレキシブルにコーパスを部分選択

一方, ベースライン手法は, 極性表現対集合を「良い」と「悪い」のみとし, f2を行わないものである.

5.2 結果

被験者5名に, 名詞句, 状況, 判断条件, および, 評価極性を列挙してもらったところ110問が得られた. 各手法による評価極性との一致率を求めた(表3).

提案手法は, 「熱い鍋」は「夏」に不評, 「冬」に好評と出力した. ベースラインでは, 一辺倒であったので, 提案手法の効果が確認できた.

表3 実験結果

手法	一致数	出題数	一致率
提案手法 (f1+f2)	83	110	0.75
比較手法 (f1のみ)	83	110	0.75
比較手法 (f2のみ)	82	110	0.75
ベースライン手法	78	110	0.71

6 おわりに

名詞句の評価極性の判定は情緒推定の精度に影響するので, 本稿では, 名詞句の種類, 状況, および, 判断条件(概して名詞句に対する動作の分類)に応じて評価極性を判定する方法, すなわちSO-score算出用の極性表現とコーパスを可変化する方法を提案し, 実験により有効性を確認した. 今後の課題は情緒推定との連携である.

参考文献

- [1] 田中努, 徳久雅人, 村上仁一, 池原悟: “結合値パターンへの情緒生起情報の付与”, 言語処理学会第10回年次大会発表論文集, pp.345-348, 2004.
- [2] 野口和樹, 徳久雅人, 村上仁一: “情緒推定における状況の対称性を考慮するためのパターン辞書の拡張”, 言語処理学会第19回年次大会, pp.256-259, 2013.
- [3] 竹本翔, 徳久雅人, 村田真樹: “SO-scoreとパターン辞書を用いた情緒推定”, 言語処理学会第18回年次大会, pp.771-774, 2012.
- [4] Turney, P. D.: “Thumbs Up or Thumbs down? Semantic Orientation Applied to Unsupervised Classification of Reviews”, ACL2002, pp.417-424, 2002.