

情緒状態と直結した情緒反応を表す用言の結合価パターン辞書の構築

Construction of valency pattern dictionary representing directly emotional reaction

黒住亜紀子
Akiko Kurozumi

徳久雅人
Masato Tokuhisa

村上仁一
Jin'ichi Murakami

池原悟
Satoru Ikehara

鳥取大学 工学部 知能情報工学科
Dept. of Knowledge and Info. Eng., Tottori University

1 はじめに

テキストからの情緒推定において、情緒を表す用言に着目した手法が考えられている。松本ら [1] は、日本語彙大系 [2] のうち、「感情動作」、「感情状態」に属する結合価パターン全てに対して情緒推定用の情報を付与し、辞書として使用している。一方 [3] では、上記 2 つのカテゴリに属する結合価パターンのうち、情緒を明示するもの（直接表現）に対して情緒情報を付与し辞書としているが、情緒による反応（反応表現）は対象外とされている。

感情のカテゴリに属している結合価パターンの中に、情緒状態を表す場合や情緒反応を表す場合が混在しているが、一樣に情緒情報を付与して正しい推定が行えるかどうか懸念される。

そこで本稿では、まず、[3] で反応表現に分類された結合価パターンに対して情緒情報の付与を行い、辞書を拡張する。次に、直接表現、反応表現それぞれのパターンにおける推定性能を実験的に調査する。

2 情緒属性付き結合価パターン辞書の拡張

2.1 反応表現への情緒属性の付与

反応表現に分類されている結合価パターンは 104 件ある。これに対して、「情緒主」、「情緒対象」、「情緒名」（本稿ではこれらを「情緒属性」と呼ぶ）の付与を行う。情緒名は《喜び》、《悲しみ》、《好ましい》、《嫌だ》、《驚き》、《期待》、《恐れ》、《怒り》、《なし》の 9 種類とする。

図 1 に反応表現に対する情緒属性の付与の例を示す。「泣く」という用言は、人物 $N1$ が $N2$ に対して大きく分けて《喜び》の情緒状態で泣く場合、または、《悲しみ》、《恐れ》、《嫌だ》、《怒り》のいずれかの情緒状態で泣く場合の 2 つがあると考えられるので、2 つの情緒属性を付与する。(ii) のように情緒名が複数付与されている場合、「選択」か「混合」のフラグを付与する。「選択」は複数の中から適切な情緒を選択する必要があることを示し、「混合」は同時生起を示す。

日本語彙大系に存在する項目	
泣く(なく) (2)	31 感情動作 23 身体動作 (動作 受身不可)
$N1$ が $N2$ で/に 泣く	[$N1(4人)$ $N2(*)$]
追加項目	
(i) 情緒主: $N1$, 情緒対象: $N2$, 情緒名: 《喜び》	
(ii) 情緒主: $N1$, 情緒対象: $N2$, 情緒名: 《悲しみ》, 《恐れ》, 《怒り》, 《嫌だ》*1 (選択)	

図 1 反応表現に対する情緒属性の付与の例

2.2 付与結果

反応表現に分類されていた結合価パターン 104 件を分析した結果、4 件は分類の誤りであると考えられるため除外し、100 件のパターンに対して情緒属性を付与した。このうち、40 件のパ

ターンは情緒選択が必要であった。中でも、8 種類（《なし》除く）の情緒の中からどれかを選択する必要のあるパターンが 9 件であった。また、Positive, Negative の相反する情緒属性を持つパターンが 28 件あった。

3 情緒推定実験

3.1 実験の目的

直接表現のパターン、および、反応表現のパターンを辞書として次のように定義する。

辞書 A: 直接表現のパターン (1,030 件)

辞書 B: 反応表現のパターン (100 件)

それぞれの辞書が情緒推定の対象とする文に対してどの程度の性能を持つか評価する。また、それぞれにおいてパターンに適合した節に対する評価を行い、辞書への情緒付与の精度を調査する。

3.2 実験対象

実験には日記コーパス [4](1,615 節) を使用する。[4] は、節の単位で文が分割されており、節ごとに情緒分析の経験者 6 名と初心者 1 名によって情緒タグが付与されている。

[4] から、情緒状態を表す節、および、情緒反応を表す節を抽出したところ、155 節を得た。本稿では情緒推定の正解を、経験者 6 名のタグの多数決で決定する。しかし、人間による推定においても曖昧性のある節は、明確な正解を決定できないため対象外とする。

そこで、4 名以上の一致があり、かつ上位 2 位までの情緒を正解タグとする。この条件によると、最終的に 141 節となった。このうち、辞書 A でカバーすべき節が 127 節、辞書 B でカバーすべき節が 14 節であった。

3.3 情緒推定の方法

情緒推定は、プログラムによるパターン照合によって行う。本辞書の結合価パターンを、パターン照合用に加工して照合に使用する。

まず、入力文の形態素解析結果に適合するパターンを検索する。その際、結合価パターンに定義されている変数の名詞意味属性制約によって、適合パターンの絞り込みを行う。絞り込み後、回答候補となるパターンが複数ある場合は、人手で最適パターンを選択し、出力とする。図 1 の (i) と (ii) のように複数の情緒属性が付与されている場合は、さらに人手で最適な情緒属性を選択する。しかし、(ii) のように複数の情緒タグが付与されている場合は、付与されている全てのタグを出力情緒として出力する。

*1 《悲しみ》や《恐れ》は泣くという反応に直結しやすい情緒である。また、子供などが強く拒否する際、《怒り》や《嫌だ》という情緒を持っている状態で泣くことが考えられるため、このように付与した

3.4 情緒推定の例

日記中の1節に対する情緒推定の例を示す。本稿で作成した反応表現のパターンに適合した例である。適合パターンの出力情緒名が《期待》,《恐れ》となっている。出力の情緒名が正解タグと等しく,推定が成功している。

入力文: すごく緊張する。

正解タグ: 《期待》,《恐れ》

適合パターン: 「N1が緊張する」

情緒主: N1, 情緒対象: その他,

情緒名: 《期待》,《恐れ》(混合)

出力: 情緒主: 話者, 情緒対象: その他,

情緒名: 《期待》,《恐れ》

3.5 評価方法

評価はそれぞれ,再現率,適合率,一致率,節単位の正解率で行う。

$$\text{再現率} = \frac{\text{正解タグと出力情緒の一致数}}{\text{正解タグの総数}}$$

$$\text{適合率} = \frac{\text{正解タグと出力情緒の一致数}}{\text{出力情緒の総数}}$$

$$\text{一致率} = \frac{\text{正解タグと出力情緒の一致数} \times 2}{\text{正解タグの総数} + \text{出力情緒の総数}}$$

節単位の正解率

$$= \frac{\text{出力情緒が正解タグと1つ以上一致した節数}}{\text{実験に使用した節数}}$$

4 実験結果

4.1 情緒推定の性能

それぞれの辞書の情緒推定の評価値を表1,2に示す。また,()内はそれぞれ実際の数を示す。辞書A,Bともに初心者に劣る。特に反応表現の適合率は低い。これは,付与された情緒名の曖昧さによって複数の出力を行うことが原因と考えられる。

表1 辞書Aの情緒推定における評価値

推定者	再現率	適合率	一致率	節単位の正解率
辞書A	0.55 (78/143)	0.51 (78/152)	0.53	0.58 (74/127)
初心者	0.68 (97/143)	0.68 (97/143)	0.68	0.76 (97/127)
経験者 平均	0.64~0.85 0.75	0.64~0.83 0.73	0.70~0.79 0.74	0.80~0.94 0.88

表2 辞書Bの情緒推定における評価値

推定者	再現率	適合率	一致率	節単位の正解率
辞書B	0.56 (9/16)	0.38 (9/24)	0.45	0.57 (8/14)
初心者	0.63 (10/16)	0.71 (10/14)	0.67	0.71 (10/14)
経験者 平均	0.69~1.0 0.75	0.73~0.93 0.73	0.71~0.91 0.74	0.79~1.0 0.93

4.2 情緒属性付与の精度

辞書A,Bそれぞれに対して,パターンに適合した節における評価値を表3,4に示す。また,()内はそれぞれ実際の数を示す。辞書Aについては,初心者と同程度またはそれ以上の結果となり,パターンへの情緒属性の付与は正しく行われていると考えられる。辞書Bについては,再現率はやいので正解の出力は初心者以上となっているが,曖昧な出力をするために適合率は初心者に劣っている。

表3 辞書Aのパターンに適合した節における評価値

推定者	再現率	適合率	一致率	節単位の正解率
辞書A	0.81 (78/96)	0.72 (78/109)	0.76	0.88 (74/84)
初心者	0.69 (66/96)	0.72 (66/92)	0.70	0.79 (66/84)
経験者 平均	0.66~0.88 0.76	0.67~0.92 0.78	0.70~0.81 0.76	0.79~0.94 0.89

表4 辞書Bのパターンに適合した節における評価値

推定者	再現率	適合率	一致率	節単位の正解率
辞書B	0.82 (9/11)	0.45 (9/20)	0.58	0.80 (8/10)
初心者	0.55 (6/11)	0.60 (6/10)	0.57	0.60 (6/10)
経験者 平均	0.73~1.0 0.84	0.75~0.90 0.82	0.76~0.88 0.83	0.80~1.0 0.90

5 考察

実験において反応表現に適合した例を次に示す。

例: みんな泣いていた。

(正解)《悲しみ》

(出力)《悲しみ》,《恐れ》,《怒り》,《嫌だ》

図1で示した「N1がN2で/に泣く」というパターンに適合した例である。今回の実験では,情緒属性(i),(ii)の選択は適切に出来るものと仮定し,(ii)を出力した。しかし,実際にこの例の1文だけでは判断することはできない。[1]では,話者にとってのN2のイメージ値を条件に,出力情緒を決定している。しかし,この例ではN2が省略されている。さらに,(ii)の選択の後,4つの情緒名から適切な情緒を選択する必要があるが,N2に依存する条件だけでは情緒の選択は難しい。

この例は,「親しい友人との別れ」について書かれた日記文の一部である。「別れ」という情緒生起の原因を示す事象と,「泣く」という情緒反応の両方によって,《悲しみ》であると推測できる。このように,情緒生起原因と情緒反応の関係性によって,適切な情緒を選択できる可能性があると考えられる。

6 おわりに

本稿では,反応表現における情緒属性付き結合価パターンを100件作成した。次に,直接表現のパターンと反応表現のパターンそれぞれに対して実験を行った。その結果,直接表現は辞書の範囲であれば初心者と同程度以上の推定が出来るということが分かった。反応表現は,情緒の選択の必要性があるということ,および,情緒を絞り込むための規則や機構が必要になることが分かった。

参考文献

- [1] 松本, 三品, 任, 黒岩: “感情生起事象文型パターンに基づいた会話文からの感情推定手法”, 自然言語処理, Vol.14, No.3, pp.239-271, 2007.
- [2] 池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林: “日本語語彙大系”, 岩波書店, 1997.
- [3] 黒住, 徳久, 村上, 池原: “結合価パターン辞書からの情緒を明示する用言の知識ベース化”, 言語処理学会第13回年次大会発表論文集, pp.39-42, 2007.
- [4] 徳久, 松浦, 村上, 池原: “情緒タグ付き日記コーパスの作成と評価”, 第5回情報科学技術フォーラム, 一般講演論文集, 第2分冊, pp.235-236, 2006.