

情緒タグ付き日記コーパスの作成と評価

Construction and evaluation of diary text corpus with emotion tags

徳久 雅人[†] 松浦 大樹[†] 村上 仁一[†] 池原 悟[†]Masato Tokuhisa[†] Taiki Matsuura[†] Jin'ichi Murakami[†] Satoru Ikehara[†]

1 はじめに

最近、インターネット上の掲示板や blog などは、幸福な話題や残念な話題など感情的な内容を多く含んでいることから、流行や事件に関する情報を得る手がかりとして注目されている。しかし、現在の研究対象の多くは製品の評判や顧客から企業への要望を抽出・解析することであり、blog などによく見られる日記を対象として喜怒哀楽などの情緒を直接扱う研究は多くない[1]。その理由の一つに、言語表現から情緒を区別することが不安定であり、研究対象として扱いにくいことが挙げられる。

そこで、本稿では日記を対象とした情緒タグ付きコーパスを構築し、情緒タグ付与の不安定さを調査し、安定性を得るために必要なタグ付与の作業数について考察する。

2 日記コーパスの構築

2.1 情緒タグの種類

本稿では、情緒タグの大別として次に述べる9分類系、5分類系、3分類系を用意する。9分類系は、複雑な情緒が基本的な情緒の合成で説明できる点に着目し Plutchik の分類[2]を参考にして、《喜び》、《悲しみ》、《好ましい》、《嫌だ》、《驚き》、《期待》、《恐れ》、《怒り》、《なし》の9種類に情緒を分類したものである。手作業でのタグ付与に使用する。ただし、英語と日本語では感情語の概念に差があるため、複雑な情緒を表すタグを付与する際、必ずしも Plutchik の示す組に従う必要はないものとする。

5分類系としては、9分類系の微妙な差異を吸収してタグの分析をするために、《P》、《N》、《A》、《S》、《なし》を用意する。《喜び》、《好ましい》、《期待》は、《P》に対応するなど、9分類系から自動的に決定することができる。その対応関係は、日本人の感情認識において、類似して認識されること[3]に着目して定めた。

3分類系は、[4]など音声対話における情緒分類で使用されているため、安定性の評価で比較するために使用するもので、《Positive》、《Negative》、《なし》とする。

2.2 付与対象のテキスト

本稿では、同一の著者が多方面の情感にまたがって記述した日記を対象にタグを付与することを目指すので、[5]の日記の日本語文を対象とする。この文献には、「ワクワクしたこと」や「イライラしたこと」のように情感に基づくテーマ分けがなされている。

情緒の関わる事態をとらえるために、1つの事態表現として、1つの「節」、すなわち、「格要素と述語の組」を単位とする。そして、その単位で情緒タグを付与することにする。

節単位の基準として、連用中止節、接続詞相当表現で区切られた節、に着目する。引用節や連体修飾節は主節と分割しない。引用節や連体修飾節の表す事態は、文脈の表す時点と異

なる時間や場所の事態であるためである。

以下に節単位に分解した例を示す([5]より引用)。

- 例1 1) 今朝スーパーで買ったサバはとても新鮮で、
2) 脂がのっけていて、
3) すごくおいしかった。

※ 1)の連体節に関して、この文脈で、日記の作者が満足している事態は、今朝でもなく、スーパーに於いてでもない。

2.3 付与の方法

情緒タグの付与は、大学生7名が行う。これら作業者を、a 者～g 者と呼ぶ。タグの付与は、次の基準で行うこととする。

- 日記の作者の情緒を考えること
- 節単位に1個以上のタグを付与すること
- 文脈(文の前後)を考慮して付与して良い

2.4 付与の結果

付与対象の日記(286話/約1,000文)から1,642個の節が得られた。7名による9分類系のタグの総数は13,165個となった。コーパスの一部を表2に示す。表中のa～gは作業者名であり、誰がどのタグを付与したかを表す。

3 コーパスの安定性

3.1 作業者間のタグの一致

次に述べる「一致率」と「 κ 値」によりタグ付与の安定性を評価する。まず、作業者7名から2名を選ぶ組み合わせを作り、それぞれの一致率から、最大値と最小値を求める。次に、全ての2名組からの一致タグ数と総タグ数により総合的な一致率とする。

$$\langle \text{一致率} \rangle = \langle \text{一致タグ数} \rangle * 2 / \langle \text{両者の総タグ数} \rangle * 100 (\%)$$

主観的な判断で付与されるタグの一致を、偶然による一致を排して評価するために、Cohenの κ 値が使用される[6]。本稿では1つのタグ付与の対象に複数のタグ付与を許したため、複数のタグが同時に一致することを一致の条件として計算する。複数付与のため厳密には正しい方法とはならないが、関連研究との参考値として求める。

評価結果を表1にまとめる。音声対話における3分類系情緒のタグging[4]での一致率・ κ 値の81.75%($\kappa=0.624$)と比べ、日記コーパスでは一致率・ κ 値がともに低かった。

9分類系と5分類系は現在のところ直接比較できるコーパスが存在しないため、今後の関連研究のための参考値となる。

表1 作業者間のタグの比較

方法	分類	総合	最大	最小
一致率	9	50.3(%)	58.9(%)	40.9(%)
	5	61.5(%)	68.5(%)	51.7(%)
	3	70.3(%)	77.8(%)	63.3(%)
κ 値	9	0.307	0.388	0.230
	5	0.431	0.554	0.345
	3	0.517	0.628	0.431

[†] 鳥取大学工学部, Tottori Univ.

3.2 正解タグと作業者との間のタグの一致

7名の付与したタグが最も多く重なる情緒タグを正解タグとする。同数の場合は、それら複数の情緒タグを正解タグとする。正解タグと各作業者のタグを比較して、作業者の精度を評価する。

9分類系の場合、各一致率(%)は、59.2, 64.3, 64.3, 66.8, 68.1, 68.1, 69.0 となった。このうち1名の作業精度が、残りの者と比べてやや目立って低かった。なお、表2のg者である。

4 正解タグ決定のための人数

多数決制による正解情緒タグの決定が、情緒タギングにおいては問題のない方法であるが、より規模の大きいコーパスを構築する上では、作業者は少ないほうがコスト面において望ましい。そこで、多数決制での人数をどの程度減らすことができるのかを考察する。

4.1 9分類系の正解タグについて

7名によるタギングの結果から決定した正解タグを「基準正解タグ」とする。その7名から多数決制の構成人として n 人を選ぶとき、 ${}_{7}C_n$ 通りの組ができる。各組ごとに多数決で正解タグを決定すると、各組による「仮正解タグ」が得られる。「基準正解タグ」と「仮正解タグ」の一致率を求めると、構成人数が少ないほど一致率が低下し、ばらつきが生じると予想される。なお、 $n = 1$ のときは、第3.2節で示したものと同一である。

図1にその結果を示す。次のことが分かった：

- 一致率 80%で正解タグを得るには、中央値と最大値が80%を超える点の人数である4名が必要である。中央値以上の精度を狙うので、やや熟練した作業者が必要である。

4.2 5分類系と3分類系の正解タグについて

5分類系においては、多数決制が3人のとき中央値が83%、最大値が85%となり、6人のとき中央値が93%となった。9分類系と人数差が1名しかないので、9分類系の代わりに5分類系のタグ付与を人手で行うことは、コスト面での利点が無い。

3分類系においては、既に1人の時点で中央値が81%、最大値が83%となり、6人のとき中央値が95%となった。3分類系は2名程度でも安定したタグ付与の見込みがある。

5 おわりに

情緒タグを付与した日記コーパスを構築し、タグ付与の安定性を評価した。さらに多数決制で正解タグを決定するために必要な人数について考察した。具体的には、コーパスの規模は約

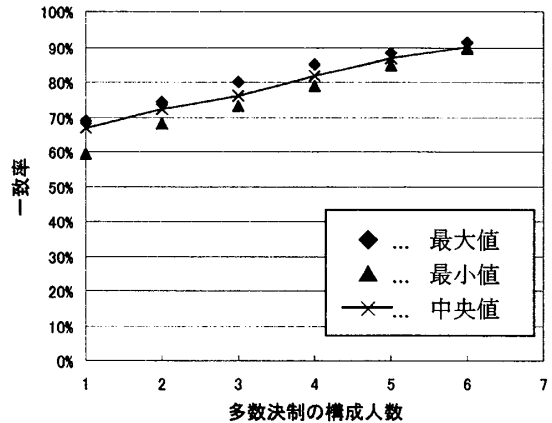


図1 多数決人数と正解タグ決定精度の関係

1,000文/1,642節、7名による情緒タグが13,165個となった。3分類系の一致率は70.3% ($\kappa=0.517$)であり、9分類系はさらに低く50.3% ($\kappa=0.307$)であった。本コーパスにおいて多数決制による正解情緒タグの決定は、9分類系の場合やや熟練した作業者4名で、また、3分類系の場合2名で、ある程度の精度(80%)が得られることが分かった。

以上により、本稿では、情緒タグ付き日記コーパスの構築における情緒タグの安定性を定量的に評価し、安定した情緒タグ付与のための作業者数に関する知見を得た。

謝辞 本研究は科研費(若手B:17700151)の下で行いました。

参考文献

- [1] 奥村学, 乾孝司: blogマイニングと評判分析, 言語処理学会第12回年次大会チュートリアル資料, pp.79-96, 2006.
- [2] Plutchik, R.: The multifactor-analytic theory of emotion, *Psychology*, 50, pp.153-171, 1960.
- [3] 重野純: 感情認知における視聴覚情報の統合と文化の影響, 視聴覚研究会資料, 35(8), pp.507-512, 2005.
- [4] Litman, D., Forbes, K.: Recognizing emotions from student speech in tutoring dialogues, *Automatic Speech Recognition and Understanding Workshop*, 2003.
- [5] 石原真弓: 英語で日記を書いてみる, ベレ出版, 2002.
- [6] Cohen, J.: A coefficient of agreement for nominal scales, *Education and Psychological Measurement*, 20(1), pp.37-46, 1960.

表2 コーパスの一部(事態表現は[5]より引用)

#	事態表現	3分類系		Positive			Negative			なし	
		5分類系	9分類系	P			N			A	
				喜び	好ましい	期待	悲しみ	恐れ	嫌だ	怒り	S
1	お父さんに床の張り替えを手伝うように言われ								acdef	g	b
2	えーっ、そんなことでせつかくの日曜日を無駄にしたいよーと思っただけど、					af	e		abcdeg		
3	仕方なく起きて								abcdefg		
4	手伝った。								cef	g	abd
5	最初は面倒だと思っただけど、						g		abcdef		
6	コツがつかめてきたら			ef	g						abcd
7	結構楽しいと思えた。	abe		cdfg							
8	今日は台所の床が終わり		a	e							bcdfg
9	来週の日曜日は廊下をする予定。			aceg				d			bf