

談話役割に基づくクラス制約規則を利用したレビュー文の意見分類

浅野 翔太 乾 孝司 山本 幹雄

筑波大学大学院 システム情報工学研究科

asano@mibel.cs.tsukuba.ac.jp {inui,myama}@cs.tsukuba.ac.jp

1 序論

ソーシャルメディアの普及によって、個人の Web 上での活動が盛んになっており、Web 上には個人の意見や感想が大量に蓄積されている。また、これら意見や感想データを利活用する技術として、評判分析に関する研究が盛況であり、近年では、文書単位ではなく文やフレーズのような粒度の細かな分析や、意見の分類軸として評価極性 (positive/negative) 以外の観点を考慮する試みがなされている [5]。しかし、これらの新しい課題は文書の評価極性を推定するような古典的な課題に比べると、未だ発展段階にあると言え、新たな解析モデルや素性情報の検討が進められている (e.g. [8], [4])。

本研究では、レビュー文の意見分類課題に対して、レビュー内の各文がもつ談話役割を考慮する手法を提案する。レビューの主目的は感想や意見を述べることであるが、実際には、冒頭箇所に唐突に感想や意見を述べるのではなく、その前に挨拶やお礼などを述べるのがよくある。また、自分の抱いた意見を述べる前に、その意見を抱くに至った経験や状況説明を述べることも多い。また、好意的な意見を述べた後には、対象商品のリピート使用を示唆するようなコメントを残すこともある。そこで本研究では、まず、このようなレビュー文書の構成要素である各文がもつ挨拶や状況説明といった談話役割の情報と書き手の対象商品に対する評価情報との対応関係を調査した。その後、調査から得られた知見を元に、談話役割を考慮した意見 (評価) 文の分類手法を提案し、評価実験を通して、その有効性を検証する。

談話役割はドメインに依存すると考えられる。本研究では、楽天データ公開¹において公開された楽天トラベルの施設レビューから 1,000 件のレビューを抽出して調査および評価実験に用いた。

2 談話役割と評価情報

レビュー中の文がもつ談話役割と評価情報の対応関係を調査した。まず、談話役割と評価情報のそれぞれの分布状況を表 1 に示す。なお、調査対象は施設レビューに含まれていた 4,309 文であり、この文の認定にあつ

表 1: 談話役割と評価情報の各分布

談話役割	文数	割合 (%)
宿泊事情	321	7.45
挨拶	164	3.81
感想	2,542	58.99
体験	202	4.69
要求	218	5.06
再泊意向有	271	6.29
再泊意向無	8	0.19
個別事情	117	2.72
情報追加	37	0.86
その他	118	2.74
(マルチラベル)	311	7.22
計	4,309	

評価情報	文数	割合 (%)
褒め (positive)	1,979	45.93
苦情 (negative)	740	17.17
ニュートラル	101	2.34
要求	230	5.34
評価なし	930	21.58
その他	20	0.46
(マルチラベル)	309	7.17
計	4,309	

ては句点等に基づく簡単なヒューリスティック規則を適用した。談話役割クラスは、大沢ら [10]、安藤ら [9] を参考にし、評価情報との関連が希薄であると考えられるものを統廃合して作成した。評価情報クラスは、標準的な分類に概ね従っているが、「褒め (positive)」と「苦情 (negative)」のどちらの評価極性も持たない文に対しては、意見を述べた「ニュートラル」と意見を述べていない「評価なし」に細分類している。また、どちらのクラスセットに対しても、複数の分類が該当する場合はマルチラベルとして区別して集計した。

表 1 によると、談話役割では約 6 割の事例が「感想」であり、次いで「宿泊事情」、「再泊意向有」、「要求」が多いことがわかる。また、評価情報では「褒め」があ

¹<http://rit.rakuten.co.jp/rdr/index.html>

表 2: 談話役割と評価情報の対応関係

	褒め	苦情	ニュートラル	要求	評価なし	その他	(マルチラベル)	計
宿泊事情	15	1			305			321
挨拶	13				150	1		164
感想	1,656	553	88	3	32	7	203	2,542
体験	41	116	8		22	3	12	202
要求				218				218
再泊意向有	73				198			271
再泊意向無		8						8
個別事情	7	8			96	4	2	117
情報追加	14	5			18			37
プラン名					19			19
その他	8	6	1		80	4		99
(マルチラベル)	152	43	4	9	10	1	92	311
計	1,979	740	101	230	930	20	309	4,309

よそ半数を占め、次いで、「評価なし」が多かった。

次に、談話役割と評価情報の対応関係を調査した。調査では、ある文の評価情報に着目し、3種類の文(同一文、1つ前の文、1つ後ろの文)について、談話役割の分布状況を調べた。ここでは、紙面の都合上、同一文における集計結果のみ表2に示す。表中の空欄は該当する文が0件であったことを示す。

表2から、談話役割毎に対応する評価情報には大きな偏りがあることがわかる。例えば、談話役割が「宿泊事情」や「再泊意向有」の場合、ほとんどの評価情報は「評価なし」となり、残りは「褒め」となる。また、談話役割が「再泊意向無」の場合はすべての文の評価情報が「苦情」であった。また、集計表は割愛するが、ある文の評価情報とその前文の談話役割を見ると、評価情報として「苦情」か「要求」を述べている文は、その前文の談話役割が「再泊意向無」となるなどの傾向が明らかになった。

この調査結果は、談話役割は評価情報クラスの推定にとって有益であることを示唆している。そこで、次節において、談話役割の情報を考慮した評価クラスの推定手法について検討する。なお、本節の調査で用いた談話役割および評価情報は、著者らのうちの1名が人手で付与したものである。この信頼性を評価するために、もう1名の別の作業にも同様の作業を行ってもらい、作業間での判断の一致率を評価したところ、談話役割で82%、評価情報で92%であった。

3 談話役割を考慮した評価クラス分類

本節では、談話役割を考慮する手法として、2つの手法(談話素性法と談話制約法)を提案する。談話素性法では談話役割を選好情報として利用し、談話制約

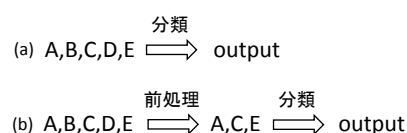


図 1: 談話制約法の概念図

法では談話役割を制約情報として利用する。

3.1 談話素性法

評価クラス分類の手続きとして、SVM[7]のような分類器を使用することを仮定する。この場合、談話役割を考慮する単純な手法として、談話役割の情報を素性として取り込むことが考えられる。本稿では、後述する標準素性セットに談話役割ごとのバイナリ素性を加える手法を検討する。

3.2 談話制約法

図1の状況(a)がもともと解くべき評価クラスの分類問題であるとする。この時、状況(b)のように、分類対象クラスを前処理で絞り込むことを考える。談話制約法では、前節の調査結果を反映させた後述のクラス制約規則を使って対象クラスの絞り込みを実現する。本研究では、談話役割と評価情報の対応に現れる顕著な偏りを元にしてクラス制約規則を作成する。例えば、談話役割が「宿泊事情」である場合、ほとんどの評価情報は「評価なし」か「褒め」のいずれかになることがわかっている。この知見から、

- if 対象文の談話役割 = 「宿泊事情」
then 分類対象クラス = { 評価なし, 褒め }

というクラス制約規則が作成できる。規則の条件部には、対象文の他にその前後文脈がもつ談話役割を指定することも可能である。また、条件部が無矛盾であれば、複数の規則を組合せて新たな規則を作成することもできる。このように複数の規則を組合せることで絞り込み機能を強めることができる。

4 評価実験

4.1 各手法の詳細設定

評価実験を通して、前節で述べた談話役割を考慮する手法の有効性を検証する。ここでは、談話役割を考慮しない分類手法をベースライン法として、これと提案手法の分類正解率を比較する。なお、マルチラベルを持つ文は、評価情報が曖昧であるため、評価データは除外した。ただし、学習データには含まれているため、マルチラベルと分類されたものは全て誤りと判定した。

実験では、各手法において SVM にペアワイズ法を適用して多値分類を実現し、カーネルには RBF カーネルを用いた。各パラメータは、学習データの中で 5 分割交差検定を用いてグリッドサーチを行って求めた。

• ベースライン法

談話役割を考慮せず、標準素性のみを利用して学習したモデルを利用する。標準素性は、対象文の単語の Bag of Words、異なり単語数、positive 単語が negative 単語より多いか、対象文のレビュー中での位置等とする。

• 談話素性法

標準素性に談話素性を加えて学習したモデルを利用する。

• 談話制約法（規則の緩和適用）

対象文にクラス制約規則を適用し、分類対象となる評価クラスを制限する。得られた分類対象クラスにはベースライン法を適用し、最終出力を得る。ここで、条件部で考慮する前後文脈に応じ、複数の規則が適用可能となる場合がある。このような場合には、次の優先順位に従って規則を適用する。

1 位 条件部に対象文とその前後 1 文についての談話役割が指定されている規則

2 位 条件部に対象文のみの談話役割が指定されている規則

適用できる規則が存在しない場合は、そのままベースライン法を適用する。また、実際には、規則の適用によって本来分類すべきクラスが誤って分類対象クラスから除外される可能性がある。これに対し、すべての規則の実行部には、前節の説明で明示的に選ばれたクラス以外にそれ以外のクラスをまとめた「他」クラスを常に含むように作成し、それにあわせて分類モデル

も学習する。もし、優先順位が 1 位の規則を適用し、分類結果が「他」となった場合は、それを最終出力とせず、優先順位が 2 位の規則を適用して同様の処理を再度実行する。2 位でも「他」となる場合は、ベースライン法の結果で最終出力を置換する。

• 談話制約法（規則の全適用）

対象文にクラス制約規則を適用することで分類対象となる評価クラスを制限する。得られた分類対象クラスに対してはベースライン法を適用する。複数の規則が適用可能となる場合は、それら全ての規則を並行に適用し、分類も並行に実行する。その後、分類結果の多数決に基づき、最終出力を得る。ただし、1 位タイとなった場合は、1 位タイのクラス群に制約したと見なし、その制約のモデルで再分類する。なお、これらの分類結果が「他」となった場合は、ベースライン法の結果で最終出力を置換する。

• 談話素性&制約法（規則の緩和適用）

談話制約法（規則の緩和適用）において、ベースライン法を適用する箇所を談話素性法で置き換えた手法。

• 談話素性&制約法（規則の全適用）

談話制約法（規則の全適用）において、ベースライン法を適用する箇所を談話素性法で置き換えた手法。

4.2 その他の設定

素性の抽出にあたって、形態素解析には MeCab-0.98[3] を用いた。また、SVM の実装物には Libsvm-3.14[1] を用いた。学習・評価用データは、1,000 件のレビューから無作為に選んだ 750 レビュー（3,190 文）であり、残りは開発用データとして用いた。学習・評価用データ中の各文をレビュー単位で 10 分割し、交差検定を実施した。クラス制約規則の作成において、今回は、2 節の調査結果のうち、ある談話役割に対して 10% 以上の偏りをもつ評価クラスへ絞り込むよう、クラス制約規則を作成した。ただし、学習データの中で 5 分割交差検定を行い、各規則を適用した場合における分類正解率がベースライン法よりも劣っている場合は、その規則は採用しないこととした。なお、評価実験において、評価クラスのうち「その他」は「評価なし」へ併合した。

4.3 実験結果

実験結果を表 3 に示す。ベースライン法と比較して t 検定を実施し、有意水準 1% で有意差が認められたものに“+”印を付けた。また“*”印と“***”印は、談話素性法と比較して有意水準 5% 及び 1% で有意差が認められた結果である。

提案手法では談話役割の情報が必須である。表 3 のうち、“Oracle”の列は人手で談話役割を与えた場合の結果であり、ベースライン法と比較して全ての手法で

表 3: 実験結果 (正解率)

手法	Oracle	SVM
ベースライン	74.98	
談話素性	81.57 ⁺	75.59
談話制約 (緩和適用)	77.62 ⁺	74.68
談話制約 (全適用)	80.11 ⁺	75.76 ⁺
談話素性&制約 (緩和適用)	83.15 ^{+,**}	74.34
談話素性&制約 (全適用)	82.28 ^{+,*}	76.54 ^{+,**}

表 4: 談話素性&制約法 (全適用) のクラス別評価

分類クラス	正解率	ベースラインとの差
褒め (positive)	88.38	+ 1.79pt
苦情 (negative)	61.40	+ 3.70pt
ニュートラル	20.00	± 0.00pt
要求	40.80	+ 1.14pt
評価なし・その他	77.09	- 0.28pt

有意な結果が得られた。“SVM”の列は標準素性で学習された SVM 分類器を用いて自動推定した談話役割 (正解率 79.33%) を与えた場合の結果である。こちらは、談話制約法 (全適用) 及び談話素性&制約法 (全適用) のみ有意な結果となった。特に最も良い談話素性&制約法 (全適用) では、ベースライン法と比較して 1.56pt の向上がみられ、談話素性法と比較しても 0.95pt 向上していた。

表 4 に談話素性&制約法 (全適用) のクラス別評価を示す。談話役割を用いることで、全体の分類性能を落とさずに「褒め」「苦情」「要求」の分類性能が向上している。実例として (感想・苦情) (感想・苦情) (要求・要求) という文脈で、要求直前の「苦情」が新たに正解した。これは、要求の前の感想は「苦情」になりやすく、このような談話役割の情報を有効に活用出来た事が性能向上に繋がった。

5 関連研究

Wang ら [8] は、接続詞などの単語と修辞関係の対応関係規則を独自に作成し、その規則に従って得られた文間の修辞関係を意見文分類に利用した。Voll ら [2] や Taboada ら [6] は、RST に基づく文間の修辞関係を解析することで文書中の核となる文を判定し、それらの重要度を上げて意見分類をする手法を提案している。これらの先行研究では文間の論理的な関係に注目している。一方で、本研究で採用している談話役割は書き手の意図を反映した意味的な情報である。また、先行研究では文間関係が重視されているが、本研究では文単体に注目しており、文間関係はあくまで個々の文単体が持つ談話役割の連鎖として間接的に考慮される。

6 結論

本研究では、談話役割と評価情報との対応関係に注目し、談話役割を考慮した意見文の分類手法を提案した。その結果、談話役割を考慮しない手法と比較し、分類正解率が最大で 1.56pt 向上し有意な結果となった。以上により、談話役割が評価情報の分類に有効であることがわかった。しかし、現状の談話役割の推定手法は、分類クラスが異なっていることを除けばベースライン法と全く同様であり、まだ改善の余地がある。例えば、評価情報の推定を先に行い、対応関係から談話役割を推定する。その後、その談話役割を用いて評価情報を再推定するという手法が考えられる。また、この操作を結果が収束するまで繰返す手法も考えられる。これらの検討は、今後の課題としたい。

謝辞

評価実験にあたり、楽天トラベル株式会社から施設レビューデータを提供して頂きました。深く感謝いたします。

参考文献

- [1] C.-C. Chang and C.-J. Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 27, 2011.
- [2] Voll Kimberly and Taboada Maite. Not all words are created equal: Extracting semantic orientation as a function of adjective relevance. In *Proceedings of the 20th Australian Joint Conference on Artificial Intelligence. Gold Coast, Australia.*, pp. 337–346, 2007.
- [3] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. Applying conditional random fields to Japanese morphological analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp. 230–237, 2004. <http://mecab.sourceforge.net/>.
- [4] Ryan McDonald, Kerry Hannan, Tyler Neylon, Mike Wells, and Jeff Reynar. Structured models for fine-to-coarse sentiment analysis. in *Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL)*, 2007.
- [5] Bo Pang and Lillian Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, Vol. 2, No. 1-2, pp. 1–135, January 2008.
- [6] Maite Taboada, Kimberly Voll, and Julian Brooke. Extracting sentiment as a function of discourse structure and topicality. *Simon Fraser University School of Computing Science Technical Report*, 2008.
- [7] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, 1995.
- [8] Wu Fei Wang, Yunfang, and Qiu Likun. Exploiting discourse relations for sentiment analysis. In *Proceedings of COLING 2012*, pp. 1311–1320, 2012.
- [9] 安藤まや, 関根聡. レビューには何が書かれているのか?, 2013. NLP 若手の会.
- [10] 大沢裕子, 郷垂里沙, 安田励子. Web サイトにおけるクチコミの苦情と返答 - 「宿泊予約サイト」を対象に -. 言語処理学会第 16 回年次大会論文集, pp. 322–325, 2010.