

文脈一貫性を利用した極性付評価表現の語彙獲得

那須川 哲哉 金山 博
日本 IBM(株) 東京基礎研究所

概要

組織・製品などに関して好不評を示す表現、及びそれが好評を示すか不評を示すかの極性を、文書データから教師なし学習で獲得する手法を示す。ある対象に関して評価を述べる内容を記述する際には、好評もしくは不評の意見を列挙することが多く、好不評の極性を反転させる際には「けれども」「しかし」といった接続表現で明示することが多い。本手法では、この『評価表現の文脈一貫性』を利用し、「満足する」「不満だ」などの種表現の周辺文脈から評価表現の候補とその極性を抽出し、各候補の文書データ全体における分布から評価表現としての妥当性を判定する。得られた評価表現を種表現に追加する操作を再帰的に実行することで、好不評の極性付きの語彙を、少数の種表現から自動構築することが可能になる。デジタルカメラと映画に関する掲示板のデータで実験したところ、各々の分野に特徴的な評価表現を高い精度で抽出することができた。

Acquisition of Sentiment Lexicon by Using Context Coherence

Tetsuya Nasukawa Hiroshi Kanayama
IBM Research, Tokyo Research Laboratory
nasukawa@jp.ibm.com hkana@jp.ibm.com

Abstract

This paper presents an unsupervised learning method to acquire sentiment expressions that indicate either favorable or unfavorable opinion toward a specific subject. Our approach is based on a characteristic of sentiments that multiple sentiment expressions often appear near each other in texts with the same polarity of favorability, unless the change is explicitly indicated by a clue term such as an adversative conjunction. By taking advantage of this characteristic, our unsupervised method extracts sentiment expressions and their polarities from a corpus starting from a very small set of seed expressions and analyzing their neighboring expressions in the corpus. In our experiment on discussion board messages about digital cameras and movies, we could acquire a good set of sentiment expressions relevant to each domain.

1. はじめに

インターネットを中心にネットワーク上でアクセス可能なテキストが増大している結果、新しい自然言語処理技術の需要が高まってきている。その一つが評判分析の技術である。掲示板や BLOG などに記述された各種製品、映画、旅行先、音楽などに関する様々な情報は、消費者にとって、何を買うかどこへ行くかを検討する上での貴重な判断材料となり、また企業などの提供者には、商品をどう向上させるべきかといった様々な知見を与えてくれる。近年、このような情報が、人手では網羅的に目に通せないほど膨大な量に達しており、それを十分に活用する技術として、評判分析の技術が注目を集めてきている [1][13]。評判分析の中心的な機能は、対象となる製

品や組織名などに関する情報を収集し、その対象に関して

- 良いと書かれているか、悪いと書かれているか
- 良いこととして何が挙げられているか、悪いこととして何が挙げられているか
- 特徴は何と書かれているか

といった内容を集約することであり、そのために重要なのが、何が良いこと(好評)で何が悪いこと(不評)かという評価を示す表現(以降、**評価表現**と記す)である。

評価表現は、「良い」「悪い」といった形容詞に限らない。「満足する」「落胆させる」といった動詞や、「絶品」「欠陥」といった名詞なども含まれる。その上、「寿命が長い」なら良いが「待ち時間が長い」なら悪いというように、複合的な表現をもって良し悪しが表

現される場合もあり、多種多様である。また、同じ表現の良し悪しが分野によって反転する場合もある。例えば、「怖い」という表現は、一般的には否定的な（不評の）意味で使われることが多いが、遊園地のジェットコースターやお化け屋敷に関しては肯定的な（好評の）意味で使われることが多い。また、「眠くなる」という表現は、寝具に関しては好評の表現であるが、講演などに関しては不評の表現となる。このように多種多様でしかも分野に依存する評価表現を内観によって人手で定義するのは困難であり、対象分野の実文書から実例に基づいて定義することが望ましい。

本稿では、評価表現、及びそれが好評を示すか不評を示すかという極性を、大量の文書データから自動的に抽出する課題に取り組む。このような評価表現には、一連の文脈で連続して出現し易く、その際には反転を明示されない限り同じ極性を取り易いという傾向が見られる。この性質を利用し、種となる評価表現の周辺文脈から新たな評価表現の候補を抽出した上で、文書データ全体での分布を考慮して高い抽出精度を実現する手法を提案する。また、実験データに基づいてその妥当性を示す。

2. 関連研究

評判分析には、

- A) ある対象に対して、文章全体として好評と不評のどちらを示しているかを判定するタイプ。例えば「この映画は非常に良い」「この製品には満足している」といった総合的な判断を下す[2][5][14]
- B) ある対象に関して、何が良く何が悪いかを抽出するタイプ。例えば、ある製品に関して「使い勝手が良い」「デザインはまいち」といった肯定的な記述内容と否定的な記述内容を認識して、各内容を抽出する[6][9][13][15]

という二つのタイプが存在し、評判分析のための評価表現は、このタイプによって異なってくる。

A) のタイプの場合、好不評を示す評価表現は、あくまでも最終的な判定を下すための材料であって、その表現自体を必ずしも最終出力に含める必要がない。そのため、好不評表現が人間の直感とは一致しない表現であっても構わない。このタイプでは、映画や製品に関して五つ星のようなレビュー結果が付けられた正解データを用い、文書中に出現する単語の N グラムに対する好不評との相関度を機械学習させるものが多い[2][5]。判別精度が高くなるのは 1 グラムもしくは 2 グラムを対象にして機械学習させた場合で、N グラムに文法的属性を加味しても精度向上には寄与しないという実験結果が得られている。

[5]で具体的に肯定的な要素として抽出された語を見ると、not や back など、直感的には評価表現とは認識し難い表現も多い。但し、工藤ら[7]が提案しているような、可変長の構造を対象として分類に寄与する素性を選択する手法を好不評の分類に適用した場合は、重みの高い素性が人間の直感に近づく可能性が高い。また、[14]では、「形容詞と動詞の組み合わせ」というように、ある程度文法的性質を考慮して、組み合わせを限定した 2 グラムを表現の単位にしている。さらに、この表現に対し、インターネットの検索エンジンを用いて、各表現が文書中で poor 及び excellent の近傍で共起している度合いから、語の好評度、不評度を測っている。この手法では、限定的な正解データからの機械学習と異なり、インターネット上の膨大な文書を知識源として、好評もしくは不評との相関が高いとみなされる表現を抽出できる。さらに、典型的な不評表現 (poor) 好評表現 (excellent) との近傍共起性を考慮する点で、本稿で提案する手法との共通性がある。しかし、2 グラムに限定していることや、分野依存性を考慮していない点などから、得られる評価表現は、我々の目的としている評価表現とは異なってくる。

B) のタイプの評判分析では、評価表現自体が出力の中心となるため、評価表現として直感的にも妥当な表現の抽出が求められる。従って、このタイプの評判分析で用いる表現が、本稿の対象とする評価表現と一致する。しかし、既存の手法[6][9][13][15]では評価表現の辞書を人手で構築している。

評判分析という応用よりも、評価に関する表現やその極性を学習する手法に焦点を当てた研究としては、文中の並列構造[8]や、因果関係を示す構造[4]を利用する試みが存在する。

[8]では、文書中で and, or, but の接続詞で結ばれた並列句で共起している評価表現（論文中では形容詞）の好不評極性に関して、and と or は同じ極性を、but の場合は逆の極性を取るという仮定に基づいて極性を学習する。例えば、“simple and well-received” という表現からは simple と well-received の両者が同じ極性を取るという仮定に基づいて、simple が肯定的であれば well-received も肯定的であると推定できる。[4]では、『肯定的な出来事は肯定的な出来事を（否定的な出来事は否定的な出来事を）引き起こす』という仮定を、因果関係の出来事対に適用することで、初期の辞書からブートストラップ的に肯定的もしくは否定的な出来事を推定していく手法を提案している。これらの手法は、特定の構造を持った文のみが対象となることから、表現を抽出できる対象が限定される上、対象の構造が

正確に解析されている必要がある。また、文単位での処理のため、同文中で共起している表現しか学習の対象にならないという制約がある。

文脈一貫性という手法的な観点での関連研究に関しては、文脈内での語の使われ方の一貫性を応用して文脈単位で処理を行う試みはまだ少ない。語義や係り受けの曖昧性解消や代名詞の照応への適用[11]の他には、同一人物が記述した文章における表現の一貫性を利用した同義表現の獲得への適用[9]を挙げることができる。

3. 文脈一貫性を利用した語彙獲得

本稿では、評価表現の特性を利用して、同一文内に限らず隣接文も対象に含め、より広い範囲から、新たな評価表現を学習する手法を示す。

3.1. 評価表現の文脈一貫性

一般的に、ある対象に対して「良い」「悪い」「満足した」という記述をする際には、どのように良い(悪い)か、なぜ満足したかといった好評(不評)の内容を示す表現を併記することが多い。その際、

- 軽くて小型で便利。しかも長持ちするので気に入っている。
- 迫力満点で見ごたえがありました。大満足です。

のように好評を述べる表現が連続する場合と、

- 使い難い上に不具合があって、がっかりです。
- 主人公に共感できず、腹立しい行動が多くて期待はずれです。

のように不評を述べる表現が連続する場合があり、好評と不評の表現を並べる場合は、

- 解像度は落ちるんですけど、小さくて軽いので使っています。
- 役者の演技は素晴らしかった。しかし、ストーリーがパッとしなかった。

のように、好不評の極性の切り替えを「けど」や「しかし」といった逆接の接続表現で明示することが多い。

すなわち、文書中に評価表現が存在すると、その周囲に評価表現の連続する文脈(以降、**評価部脈**)が形成されることが多く、その中では、明示的されない限り、好不評の極性が一致する傾向がある。これを本稿では**評価表現の文脈一貫性**と呼び、以降、好評の評価表現が連続する文脈を**好評文脈**、不評の評価表現が連続する文脈を**不評文脈**と呼ぶ。

この文脈一貫性が実際にどの程度の確度で成立しているかをインターネット上の掲示板のデータで調査した結果を表 3.1 に示す。

表 3.1: 好評表現「満足する」と隣接する述語の極性

述語位置	各極性の述語の数					
	間に逆接の表現を挟まない			間に逆接の表現を挟む		
	好評	その他	不評	好評	その他	不評
前文	25	9	1	0	1	6
同文	18	10	2	0	2	5
後文	18	4	1	0	2	4

ここでは、典型的な好評表現である「満足する」を主節とする文において、それと隣接する表現と考えられる以下の3つの場合に、それらが「好評」「その他」「不評」のそれぞれと判定される場合を数え上げた。

前文: 「満足する」一つ前の文の主節

同文: 「満足する」と並列の用言句¹

後文: 「満足する」の一つ後の文の主節

これらについて、逆接の表現を挟む場合と挟まない場合に分けて調査した。逆接の表現を挟む場合は、「前文」ならば「満足する」の文の先頭に、「後文」ならば次の文の先頭に、逆接の接続詞(「しかし」「だが」など)が付いている場合であり、「同文」ならば、対象となる述語と「満足する」が「ものの」「けれど」などの逆接の接続助詞で結ばれている場合を指す。

結果を見ると、逆接の表現が無い場合には、隣接表現の69%が好評の表現であり、残りのうち26%は評価表現以外の表現であった。また、逆接表現を挟む場合は、その75%が不評の表現であり、好評の表現となる場合はなかった。接続詞で結ばれた形容詞の極性に注目した Hatzivassiloglou らの手法[8]では、英語の場合、極性を判定させる接続詞としては“but”のみを扱っているが、本稿の手法では他のさまざまな言語現象を扱うため、考慮すべき接続表現が多くなる。5節の実験では、接続表現は人手により準備したが、評価表現のリストをもとに順接・逆接の接続表現を自動抽出することも可能である。

この調査結果より、隣接表現が必ずしも評価表現とは限らないものの、評価表現である場合の極性の一致性に関しては、高い確度で成立していることが認められた。成立しないケースとしては、形式ばらない掲示板のくだけた文体であるがゆえに、反転を示す接続表現が省略されている場合と、

¹順接、逆接、累加、因果関係を表す接続助詞(11種)を伴って、または用言の連用形によって用言句が結ばれている場合を考えた。

-このレストランはサービスが良い上に美味しいけれど、低料金でお薦めです。

のように、通常は併存しないと考えられる好意的(もしくは批判的)性質が併存している場合などがある。従って、この文脈一貫性が必ずしも成立するとは限らないが、現実的には、大半の場合成立しているという傾向が見られる。

3.2. 文脈一貫性を利用した語彙獲得の基本的なアイデア

評価表現の文脈一貫性が成立すると仮定すると、何らかの表現を評価表現として与えた場合、文書中で出現するその評価表現の周囲に、同じ極性を持つ評価の表現が存在する可能性が高い。従って、種となる評価表現の周囲の表現を、同じ極性を持つ評価表現の候補として抽出することができる。もちろん、候補となった表現が必ずしも評価表現であるとは限らない。例えば、「満足する」が好評を「がっかりだ」が不評を示す評価表現として与えられた場合、

- この製品を試してみて、とても満足しました。
- この製品を試してみて、がっかりです。

における「試してみる」のように、評価には直接関与しない非評価表現が周囲に存在しても決して不自然ではない。しかし、このような非評価表現は好評文脈にも不評文脈にも出現し得るのに対し、真の評価表現であれば、その表現と同じ極性の評価文脈に出現する確度が高くなると考えられる。従って、文書データ全体から、種表現と隣接する表現を候補表現として抽出した後、各候補表現が、好評文脈、不評文脈、及びそれ以外の文脈にどの程度の割合で分布しているかを調べることによって、最終的に真の評価表現と解釈すべきかを判定する。評価表現と判定されたものは、新たに種表現として、文書データにおけるその隣接表現から、別の評価表現を抽出するために用いることができる。

この処理の概要を図 3.1 に示す。入力として文書データと種表現(ここでは「満足する」という好評表現)の二つが与えられる。文書データ中で種表現の周囲(評価文脈)から評価表現の候補(C1~C6)を抽出し、各候補が好評文脈、不評文脈、それ以外の文脈に出現する頻度を求め、その結果から、C1及びC5を好評表現、C4を不評表現と判定している。次のステップでは、この結果を新たな種表現(S1, S4, S5)として、種表現を拡張した上で同様の処理を繰り返すことにより、さらに新たな評価表現を抽出することになる。

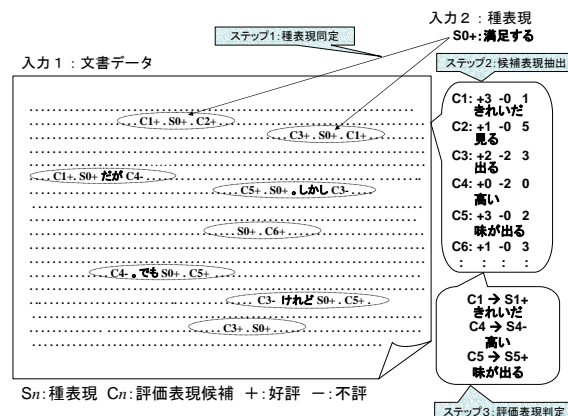


図 3.1: 文書データにおける種表現の周囲文脈からの候補表現抽出及び評価表現判別の概要

4. 語彙獲得アルゴリズムの実装

前節のアイデアに基づいて、教師なし学習による語彙獲得を実装した。処理の流れを図 4.1 に示す。

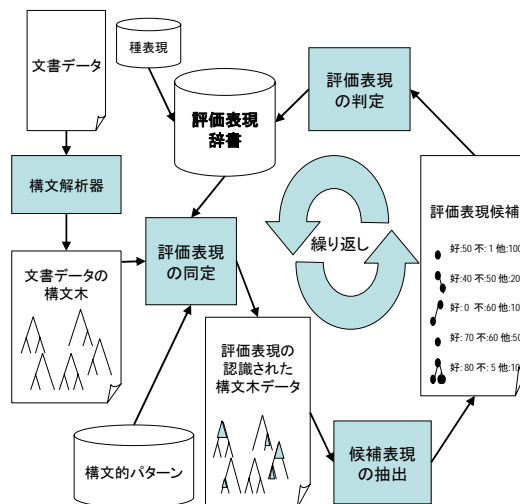


図 4.1: 処理の流れ

提案する語彙獲得の手法は、次の3つの操作を再帰的に繰り返す。最初に、種表現を含む評価表現辞書を用いて、文書データ中の全文から、評価に関する表現を、極性付きで同定する。これは、評判分析システム自体の動作に相当する。4.1 節で、構文的パターンと評価表現辞書を用いてこのような表現を同定する方法を解説する。

次に、同定された評価表現に隣接する表現を、評価表現の候補として集める。その際、前節の観察結果に基づき、通常は好評・不評の表現が文脈においてその極性を保つと仮定するが、逆接表現により極性が反転する現象も考慮する。この方法を 4.2 節にて述べる。

最後に、候補表現が実際に評価表現となるかどうかを判定する。候補表現の文書データ中での出現

頻度と、その極性が一致する確率によって判定し、基準を満たしたものについて、当該表現を新たな評価表現の辞書に追加する。詳細は 4.3 節に記す。

以上の3つの操作を繰り返すことにより、評価表現辞書のエントリ数を漸増させていく。従って、種として与えられる少数の評価表現の辞書から、多くの語彙を得ることができる。特に、初期状態が例えば「満足する」という好評の極性を持つ1語だけであっても、以下に述べる3つの操作を繰り返すことにより、好評の表現と不評の表現の両方を得ることができる。

4.1. 評価表現の同定

我々の方法で用いる評価表現の語彙の例を示す。

- | | |
|--------------|------------------|
| (1) 素晴らしい | [+] ² |
| (2) 嫌う | [-] |
| (3) 長い〈寿命:が〉 | [+] |

(1) と (2) は、それぞれ形容詞・動詞が単体で極性を決定する例であり、これらのエントリにより「この性能は素晴らしい」「私はこの電池を嫌う」といった表現を、それぞれ好評・不評と判定することができる。以後これを単純エントリと記す。また、(3)は複合エントリの例で、形容詞「長い」はガ格に「寿命」を取る時のみ好評を表すことを記述するものである。「その製品は寿命が長い」などにマッチする。

評価表現の同定には、これらの用言に関する語彙エントリのほかに、様々な構文的パターンを用いる。まず、極性を持つ用言に「Pない³」「Pということはない」等の否定を表す助動詞・文末表現が続く場合には、Pの持つ極性を反転させたものを出力する。また、条件節・疑問文や「Pしたい」といった表現は評価表現としない。基本的には文の主節(日本語の場合、文末の用言句)のみを同定対象とするが、接続助詞「て」「し」などで同定対象の用言句に係る部分を再帰的に調べる、「Pと思う」の場合は「P」の部分と同定対象とする、といった操作をしている。また、用言の格解析⁴、語の正書化なども行っている。これらには構文木の断片を変換する日英翻訳のアルゴリズムを利用している[6]。

4.2. 候補表現の抽出

同定された評価表現に隣接する表現を評価表現の候補として収集する。3.1 節でみたように、同定された評価表現の前の文・次の文の主節の用言句、および同一文中の並列用言句を隣接表現として考え

る。これらの隣接表現のうち、評価表現の同定の際と同様に、対象とするパターンを選択や極性の反転を行い、単独の用言と、用言に最も近い格助詞句を含んだ複合表現を、候補表現とする。以下に例を示す。

- デジタルカメラなど不要だと思っていました。ところが、画像がきれいで、とても満足しました。何も文句を言えません。

下線部の「満足する」が、既に好評の表現として語彙に含まれており、評価表現として同定されたとする。この時、同一文の並列句から「綺麗だ」「綺麗だく画像:が>」⁵が、それぞれ単純エントリ・複合エントリの候補として、好評の属性付きで抽出される。また、前文との間に逆接の表現「ところが」があるため、「不要だ」「不要だくデジタルカメラ:が>」⁶が不評表現の候補、さらに、次の文の「言える」「言える<文句:を>」は、その否定形が好評表現であると仮定されるため、不評表現の候補として抽出される。

4.3. 評価表現の判定

本手法では好評文脈・不評文脈中で極性が一致するという仮定を用いるが、3.1 節の末尾で述べたような例外のほか、構文解析の誤り、皮肉・殴り書きなどの特殊な記法などに起因して、抽出された候補表現にノイズが含まれることは避けられない。そこで、得られた候補表現が評価表現であるかどうか、その極性はどちらかを統計的に判定する必要がある。

単純エントリの候補について、次の基準 1~3 を満たすものを評価表現であると判定する。

1. 出現頻度が一定(次節の実験では 10) 以上
2. 候補表現が極性付きで抽出された場合のうち、好評ないし不評の極性が一定(次節の実験では 0.9) 以上の割合で一致
3. コーパス全体における、候補表現である用言の出現のうち、当該極性(2.で割合が高い側の極性)の文脈で出現した数が十分に多い

上記の 3. に関しては、Brent らによる動詞の格フレームの判定方法[3]を利用する。具体的には、まず、全ての用言に対して、

$$p = \frac{\text{好(不)評文脈での頻度}}{\text{全体での頻度}} \quad (0 \leq p \leq 1)$$

を計算する。それぞれの用言に対して与えられた p

² [+], [-] は、それぞれ「好評」「不評」の極性を示す。

³ 以下、「P」は用言句を表す。

⁴ 副助詞(「は」「も」など)のガ格やヲ格への割り当て、受動・使役などの場合の格変換など。

⁵ 表記の揺れ「きれいだ」「綺麗だ」は、語の正書化により同一視される。

⁶ 格解析により、この場合の副助詞「など」はガ格に相当すると判定されている。

を、0 から 1 を n 個に分割した bin $B_1 \sim B_n$ にプロットしていく。すなわち、 m 番目の bin B_m の頻度 d_m は、上記の p の値が

$$\frac{m-1}{n} \leq p < \frac{m}{n}$$

となるような用言の数である。 $d_1 \sim d_k$ の分布が最も二項分布に近づくような k を求め、 k よりも小さな bin に入るような用言はノイズであり、 $p > k/n$ となる用言を評価表現であるとみなす。

複合エントリの場合は、上記の条件 1 及び 2 のみで判定する。条件 1 の出現頻度の閾値は単純エントリの場合より低く(次節の実験では 3)設定している。

以上の判定方法の特徴として、以下の3点を挙げることができる。

- I. 条件 2 で好評・不評の極性を振り分け、かつ条件 3 で評価表現か否かを振り分けている。
- II. 用言単独で極性が定まらない語(「長い」など対象や分野に依存する語や、「ある」などのいわゆる light verb)が自然に省かれる。このようなものは、複合エントリに対して極性を定める。逆に、用言単独で好不評の極性がほぼ定まる語(「素晴らしい」など)を単純エントリとして判定することにより、冗長な語彙を増やさずに、評価表現辞書の被覆率を上げることができる。
- III. 条件 3 における値 p は、学習時に用いる、既に極性が付された評価表現辞書の語彙数に大きく依存するため、再帰的な学習の際に一定の割合を閾値とすることができない。この問題を、「ノイズの出現確率の分布は正規分布になる」という性質を用いて解決している。

5. 実験結果

本手法の有効性を調べるため、デジタルカメラ及び映画に関するインターネット上の掲示板(日本語)のデータから評価表現を実際に抽出した。種表現から評価表現を抽出し、抽出された表現を種表現に加えて新たな評価表現を抽出するという繰り返しを、デジタルカメラのデータに対しては7回まで、映画のデータに関しては5回まで実行した。

実際に抽出された結果を表 5.1 及び表 5.2 に示す。複合表現に関しては、実際には 4.1 節の (3) の例のように、格フレームの形で抽出されるが、表中では全て終止形の表層表現の形で出力している。

各表において、得られた評価表現のうち、() を含む表現(例えば、「(音楽、発想)が良い」)は、同じ述語に対して格内容が異なる表現を結合して表現したものであり、実際には、() 内を展開してできる複数の表現(例えば「音楽が良い」と「発想が良い」)が抽出されたことを示している。

表 5.1: デジタルカメラに関する掲示板からの自動抽出

入力	
種表現	満足する(好評)
文書データ	書き込み数: 17,625 文数: 101,292
1回目の抽出で得られた評価表現	
好評表現	購入する
不評表現	気になる
2回目の抽出で得られた評価表現	
好評表現	安い, 綺麗だ, 参考になる, 手に馴染む
不評表現	迷う
3回目の抽出で得られた評価表現	
好評表現	気に入る, S 45 が良い
不評表現	画質が落ちる, ノイズが増える, 手間が掛かる, 起動が遅い, ノイズが多い
4回目の抽出で得られた評価表現	
好評表現	デザインを好きだ, 選ぶ, 欲しい
不評表現	ノイズが出る, 困る
7回の抽出で最終的に得られた評価表現	
好評表現	購入する, 安い, 綺麗だ, 参考になる, 手に馴染む, 気に入る, (S45, 場合, 使い勝手)が良い, デザインを好きだ, デザインが好きだ, 選ぶ, 欲しい, 買う, 写真を撮る, 使える, 凄いい, 勉強になる, 手に入る, (問題, ぶれ, 不満)がない, 嬉しい, 楽しい, 仲間が増える, イイと思う, これに決める, 楽しみにする, 拍手を送る, 楽しさがある
不評表現	気になる, 迷う, 画質が落ちる, ノイズが増える, 手間が掛かる, 起動が遅い, ノイズが多い, ノイズが出る, 困る, 右に傾く, 迷う

両方のデータにおいて、人手により作成した評価表現と自動抽出結果との比較を行った結果、デジタルカメラのデータにおいては、再現率が 21.2%、適合率が 87.5%であり、映画のデータにおいては再現率が 44.4%、適合率が 93.5%であった。

両者の自動抽出結果を比較すると、大半の表現が異なっており、分野の特徴を見ることができる。例えば、「怖い」「悲しい」「泣ける」という表現は、通常は不評表現と捉えられる可能性が高いと思われるが、映画のデータ(表 5.2)では、ホラー映画や感動的な作品の評価に用いられることから、好評表現として抽出されている。

映画のデータを対象とした表 5.2 において、最終的に得られた不評表現に「涙が止まる」という表現が含まれている。これは、その否定形である「涙が止まらない」が好評文脈に頻出した結果、それが評価表現として抽出される過程で、肯定形に変換された上

表 5.2: 映画に関する掲示板からの自動抽出

入力	
種表現	満足する(好評), 面白い(好評), つまらない(不評)
文書データ	書き込み数: 73,444 文数: 284,966
1回目の抽出で得られた評価表現	
好評表現	好きだ, 可愛い, カッコいい, カッコ良い, 大好きだ, 怖い, 感動する, (緊迫感, 緊張感)がある, (音楽, 発想)が良い, 印象に残る
不評表現	飽きる
2回目の抽出で得られた評価表現	
好評表現	楽しい, 笑える, 素敵だ, どきどきする, 上手い, 旨い, 美しい, 引き込む, 強い, 演技が凄い, ファンになる, 味を出す
不評表現	
3回目の抽出で得られた評価表現	
好評表現	素晴らしい, 感動する, 印象的だ, 魅力的だ, 綺麗な, 泣ける, 笑う, 泣く, 気に入る, 楽しむ, 悲しい, 似合う
不評表現	悪い, 退屈だ, 読める, 退屈する
5回の抽出で最終的に得られた評価表現	
好評表現	面白い, 好きだ, 可愛い, カッコいい, カッコ良い, 大好きだ, 怖い, 感動する, (緊迫感, 緊張感, 迫力, 夢)がある, (音楽, 発想, 映画, ところ, テンポ, 演技, これ, 所, 感じ, キャラ, 話, 展開, 曲, 歌)が良い, (本当, 非常, 最高)に良い, (印象, 心, 耳)に残る, 楽しい, 笑える, 素敵だ, どきどきする, 上手い, 旨い, 美しい, 引き込む, 強い, (演技, 映像, ところ)が凄い, (ファン, 気分, 勉強)になる, 味を出す, 素晴らしい, 感動する, 印象的だ, 魅力的だ, 綺麗な, 泣ける, 笑う, 泣く, 気に入る, 楽しむ, 悲しい, 似合う, 嬉しい, 買う, 思い出す, 生きる, 光る, 圧倒する, 頑張る, (音楽, 演技)がよい, 心に響く, 度が高い, 胸が熱い, 衝撃を受ける, 好感が持てる, 個性が出る, 嵌まる, 表現する, キャストが豪華だ, 雰囲気を出し出す, (良さ, 味)が出る, 記憶にある
不評表現	飽きる, 悪い, 退屈だ, 読める, 退屈する, 苦手だ, 涙が止まる, どうかと思う

で極性を反転して、不評表現となったものである。それに対し、表 5.1 の最終抽出結果の好評表現に「(問題, ぶれ, 不満)がない」が存在するのは、この「ない」が助動詞ではなく形容詞であり、肯定形として扱われているためである。

6. まとめ

評判分析などへの応用を目的とした評価表現の語

彙獲得を目的とし、評価表現の文脈一貫性という特徴を利用して、種となる評価表現と文書データから、新たな評価表現を極性付きで自動的に抽出する手法を提案した。実装システムでデジタルカメラ及び映画に関するデータに本手法を適用した結果、インターネット上から抽出した生の文書データと、ごく少数の種表現から、人手を一切介さずに、各々の分野に適切な評価表現を高い精度で獲得することができた。この結果から、『評価表現の周囲には、複数の文にまたがって評価表現が連続する評価文脈が形成され易く、その際、反転が明示されない限り、極性が一致し易い』という評価表現の文脈一貫性が実際に高い確度で成立していることが認められた。

本手法では、抽出対象データにおいて、並列関係を示す構造や因果関係を示す構造といった特別の文法構造を前提とせず、同一文内だけでなく隣接文も抽出対象に含めることができる。そのため、従来研究よりも広い範囲の事象を使って語彙の獲得ができる。抽出される表現の性質は、種として与える表現の文法的属性や構造、好不評の属性に限定される事が無く、複合表現を含む様々な表現を適切なレベルで獲得することができた。例えば、デジタルカメラに関するデータでの実験結果では、「仲間が増える」は好評表現として、「ノイズが増える」は不評表現として抽出されている。

実装アルゴリズムでは、述語を対象とした評価表現の抽出を行っているが、「ある」「多い」や「ない」を述語とする複合表現から、その格内容を好不評の名詞として抽出することが可能である。例えば映画のデータの実験結果における「(緊迫感, 緊張感, 迫力, 夢)がある」から「緊迫感」、「緊張感」、「迫力」、「夢」を映画の分野において好評を示す名詞とし、デジタルカメラのデータの実験結果における「(問題, ぶれ, 不満)がない」から、「問題」、「ぶれ」、「不満」をデジタルカメラの分野において不評を示す名詞として解釈することができる。

実際のデータの中には、好不評表現が逆接の表現を挟まずに隣接しているケースが存在する。さらに、接続助詞の「が」のように、逆接と単純接続の両方に使われる語が存在するため、好不評表現の候補を抽出する際に、その極性の判定が必ずしも正確に行われるわけではない。しかし、最終的な好不評表現としての妥当性は、候補となる各表現の文書データ全体における出現傾向から判定されるため、候補抽出時の多少のノイズは無視することができるという、頑健な仕組みになっている。自動的に抽出された実験結果の中には、必ずしも評価表現とはいえない可能性のある表現も含まれるが、好評表現と判

断した中に不評表現が含まれているといった、実用に大きな支障をきたすような、極性が逆転する誤りは存在しない。また、評価表現と判定する際の閾値を上げることで、再現率と適合率という観点からの適合率を上げることが可能である。

実験の結果得られた評価表現は、分野の特徴を如実に示しており、ある対象分野でどのような内容が好意的もしくは否定的に受け止められるかという知見を与えてくれる点でも価値のあるデータであり、今後様々な応用が期待されている。

参考文献

- [1] AACL Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications, 2004.
- [2] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up? Sentiment Classification using Machine Learning Techniques. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 79-86. 2002.
- [3] Michael R. Brent. Surface cues and robust inference as a basis for the early acquisition of sub-categorization frames. *The Acquisition of the Lexicon*, pages 433-470. 1994.
- [4] 乾孝司, 乾健太郎, 松本祐治 「出来事の望ましさを判定を目的とした語彙知識獲得」, 第 10 回言語処理学会年次大会, 2004.
- [5] Kushal Dave, Steve Lawrence, and David M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th International World Wide Web Conference. 2003.
- [6] Hiroshi Kanayama, Tetsuya Nasukawa, and Hideo Watanabe. Deeper sentiment analysis using machine translation technologies. In Proceedings of 20th International Conference on Computational Linguistics. (to appear) 2004.
- [7] 工藤拓, 松本裕治: 部分木を素性とする Decision Stumps と Boosting Algorithm の適用, 情報処理学会自然言語処理研究会. No.158-009. 2003.
- [8] Vasileios Hatzivassiloglou and Kathleen R. McKeown. Predicting the semantic orientation of adjectives. In Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL, pages 174-181. 1997.
- [9] Akiko Murakami and Tetsuya Nasukawa: Term Aggregation: Mining Synonymous Expressions using Personal Stylistic Variations. In Proceedings of 20th International Conference on Computational Linguistics. (to appear) 2004.
- [10] Tetsuya Nasukawa and Jeonghee Yi. Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In Proceedings of the Second International Conference on Knowledge Capture (K-CAP). pages 70-77. 2003.
- [11] 那須川哲哉: 頑健な文脈処理のパラダイム -- 文脈依存性を考慮した自然言語処理--, 人工知能学会, Vol. 12, No. 1, 1997.
- [12] 立石健二, 石黒義英, 福島俊一: インターネットからの評判情報検索, 情報処理学会自然言語処理研究会, No.144 - 011, pp.75-82, 2001.
- [13] 立石健二, 石黒義英, 福島俊一: インターネットからの評判情報検索, 人工知能学会誌, Vol.19, No. 3, pp. 317-323, 2004.
- [14] Peter Turney. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), pages 417-424. 2002.
- [15] Jeonghee Yi, Tetsuya Nasukawa, Razvan Bunescu, and Wayne Niblack. Sentiment Analyzer: Extracting Sentiments towards a Given Topic using Natural Language Processing Techniques. In Proceedings of the Third IEEE International Conference on Data Mining (ICDM), pages 427-434. 2003.