

# 好不評文脈を応用した自然言語処理

那須川 哲哉 金山 博 坪井 祐太 渡辺 日出雄  
日本アイ・ピー・エム株式会社 東京基礎研究所

## 1. はじめに

ネット上の各種掲示板や blog に加え、顧客とのコンタクト記録を含む各種報告書など、多種多様な意見や感想を記した文書データが自然言語処理の対象として着目されてきている[1]。意見や感想がどのような表現で示されているのか、その表現が特定の対象に関して肯定的な内容(好評)なのか否定的な内容(不評)なのかを認識する処理は、多様なアプリケーションへの応用という観点からの有用性に加え、自然言語処理の研究対象としても、面白い問題を豊富に含んでいる。

基本的に、好評か不評かは意味的な問題であり、その認識には、好評を示す表現(好評表現)と不評を示す表現(不評表現)の定義が必要となる。また、「AはBより優れている。」のように、好評もしくは不評(好不評)を示す表現(ここでは「優れている」と評価対象(ここでは「A」または「B」と)との格関係によって、好評か不評かが異なる場合があるため、好不評の認識には、単なる共起性の解析ではなく、構文構造の解析が必要となる[6]。さらに、対象を直接評価するのではなく、対象の属性内容(例えば、「自動車」という対象における「走行性能」や「外観」)を評価する機会が多いので、対象の属性を把握して、それに対する評価を分析することで、対象に関連した評価内容の認識の網羅性を上げることができる[10]。このように、評価内容の認識は、語の意味と語間関係の正しい分析という点で機械翻訳に近いレベルの処理を必要とし、実際に、機械翻訳のモジュールを転用することで、システム構築作業の面から低コストで比較的高い精度の処理を実現できる[3]。

好不評の分析においては、機械翻訳などの自然言語処理と同様、省略や照応、文法的不適格文への対応といった、文脈参照を含む深い処理を必要とする場合が出てくる。また、好不評の表現は「良い」「悪い」といった形容詞に限らず、「満足する」といった動詞や、「好物」のような名詞なども含まれる。その上、単語レベル(例えば「長い」だけ)では好不評を表現しなくても、「寿命が長い」で好評を、「待ち時間が長い」で不評を示すように、複合表現になって初めて好不評を示す場合もある。さらには、何が良く何が悪いかという基準は分野や人によって異なるため、好不評表現の定義は容易でない。

以上のように、好不評の分析には、様々な難しさが見いだされる反面、文章中での好不評表現の使われ方には、特徴的な性質が存在する。第一に、好不評表現の出現する前後では、複数の好不評表現が連続し、その際に、好不評の極性が一致するという傾向が見られる。例えば、

「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、好評の表現が連続したり、

「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、不評の表現が連続したりする場合が多い。さらに、好評の表現と不評の表現が混在する際には、

「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。」

「役者の演技は素晴しかったのに、ストーリーがバツとらなかった。」のように、「けど」「のに」「しかし」「だが」といった、反転を示す接続表現が用いられるという傾向が見られる。本稿では、この性質を**好不評表現の文脈一貫性**と呼ぶ。第二に、例えば「満足する」と「満足しない」のように、程度の差こそあれ、否定形によって、好評から不評へ、不評から好評へと、好不評表現の極性が反転するという性質も備えている。

本稿では、以上のような好不評表現の性質を利用することで可能になる処理を示し、実験結果から、その妥当性と有用性を示す。例えば、ごく少数の好不評表現を種として、特定の分野の文書データから、その分野においてはどのような表現が肯定的もしくは否定的内容を示すかを学習できる[7]他、肯定的もしくは否定的に捉えられる内容の時代や地域による特性から、人々の嗜好と、その変化を分析することなどが可能になる。

## 2. 好不評文脈の認識

前述した好不評表現の文脈一貫性が成り立つとすれば、好評表現の周囲には複数の好評表現を含む好評文脈が、不評表現の周囲には複数の不評表現を含む不評文脈が形成され、好不評表現の周囲に反転を示す接続表現が存在すれば、その接続表現を挟んで好評文脈と不評文脈が隣接すると考えられる。インターネット上の掲示板データにおける調査では、同文内のみならず、句点をまたいだ前後の文にわたって、実際にこの仮定を裏付ける結果が得られている[7]。小規模な実験ではあるが、好評表現の「満足する」が存在すると、それに隣接する述語は、反転の接続表現を挟まない限り、69%が好評表現で、26%は(「試す」「較べる」のように)直接評価に関わらない表現であり、不評表現が隣接する割合は5%にすぎなかった。反転の接続表現を挟む場合は、75%が不評表現で、25%は評価に無関係な表現であり、好評表現が隣接するケースは見られなかった。

したがって、種となる好不評表現を定義し、文章中で、その種表現が存在する場所を特定できれば、その周囲に好不評文脈が形成されている可能性が高いと推測することができる。そして、推測された好不評文脈中に出現している表現は好不評を示す可能性が高い。評価とは無関係の表現が含まれる可能性も十分に高いが、そのような表現は、好評文脈と不評文脈の両方に出現し得るため、大量の文書データにおいて、多くの好不評文脈を認識することができれば、その中で、好評文脈に偏って出現する表現は好評表現、不評文脈に偏って出

現する表現は不評表現である可能性が高いと考えられる。また、評価に無関係な表現は文書全体に分布し、好不評表現は好不評文脈に偏って分布する可能性が高いと考えられる。このような考え方に基づいて、図1に示すように、

- ステップ1) 種表現を認識して好不評文脈を推測
- ステップ2) 好不評文脈中に出現する表現を抽出
- ステップ3) 好不評文脈に偏って出現する表現を認識することで、好評表現と不評表現を推測する仕組みを構築した。

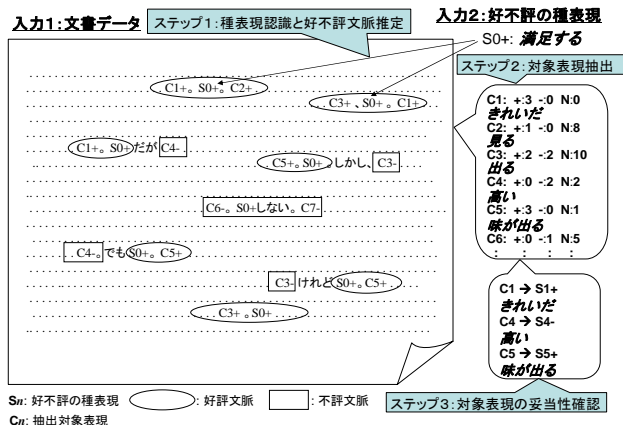


図1 文書データにおける種表現を起点とした好不評文脈の認識と好不評文脈からの評価表現の推測

### 3. 好不評表現の自動獲得

前節のステップ3において、好評文脈もしくは不評文脈へ出現が偏る度合いが十分に高い表現を、好評もしくは不評の表現と確定して、その新たな好不評表現を種表現に追加したうえで、ステップ1からステップ3を再度繰り返すことができる。ステップ3で新たな好不評表現が追加される限り、この操作を繰り返すことにより、好不評表現を漸進的に獲得することができる[7]。

実際に、好不評文脈を推測し、好不評表現を獲得する際に、構文解析を適用した上で、

- ・「～ならば良い」といった条件節中の表現や、「良いのか」といった疑問文、「良くしたい」といった希望・要望を示している文に含まれる表現を対象から外す
- ・「良くない」「良いということはない」というような否定を示す助動詞や文末表現などを伴う場合には好不評の極性を反転させる
- ・ノイズを避けるため、出現頻度の高い表現のみを対象とする
- ・「長い」では好不評を特定できず「寿命が長い」「待ち時間が長い」のような複合表現で好不評を示す場合に対応するため、「この電池は寿命が長い」のような表現からは、「長い」単体と、「寿命が長い」のようにそれに係る表現(格内容)を組み合わせた複合表現の両方を候補として抽出する

といった様々な工夫を加えた結果、表1及び表2のような結果が得られている。これらの表に示されているのは、出力から比

較的良好なものを選択した結果ではなく、上記アルゴリズムで自動的に得られた表現の全てである。人手による選別は一切行われていない。

表1 デジタルカメラに関する掲示板の約10万文から「満足する」一語を種に自動抽出した評価表現

入力	
種表現	満足する(好評表現)
文書データ	書き込み数:17,625 文数:101,292
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, きれいだ, 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

表2 映画に関する掲示板からの自動抽出結果

入力	
種表現	満足する(好評表現), 面白い(好評表現), つまらない(不評表現)
文書データ	書き込み数:73,444 文数:284,966
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, かつこいいい, かつこ良い, 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, きれいだ, 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, 雰囲気を出さず, (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

表1と表2を比較すると、分野の特徴が良く表現された好不評表現が精度良く抽出されている。表1における、「選ぶ」「写真を撮る」のように、単純に好評表現として良いか疑問を感じる表現も含まれているものの、好評表現の中に不評表現と考えられるものが含まれるというような、好不評の極性が逆になるレベルの誤りは見られない。

入力としては、ごく少数の種表現と(特別なタグ付けなどの前処理が一切施されていない)生の文書データを与えるのみで、しかも、データに依存した特殊なルールの設定や、前処理

や後処理も一切適用せずに、表1や表2に見られるような結果が得られるのは、好不評文脈の特徴と十万文を超えるデータ量を活かす事で、誤字脱字といった原文における問題や構文解析エラーなどに起因するノイズが自ずと排除されるためと考えられる。<sup>1</sup>

表1の出力の中には、複合表現のレベルで極性が確定した表現として、「仲間が増える」が好評に、「ノイズが増える」が不評になっているのを見ることができる。また、デジタルカメラのデータにおける実験では、頻度が低いために好不評表現として最終出力に含まれなかった表現の中に、「ピントが遠景に合う」や「ピントが風景に合う」が好評文脈中の肯定的な内容として存在し、「ピントが金網に合う」は不評文脈中の否定的な内容として存在した。

#### 4. 嗜好分析

好不評文脈を十分な精度で認識でき、好不評文脈に含まれる表現から、肯定的な表現と否定的な表現を実際に抽出できることから、人が何を好意的に捉え、何を否定的に捉えているかという嗜好を推測することが可能になると考えられる。

嗜好という観点で対象としたい表現は、「良い」「悪い」「好き」「嫌い」「美味しい」「不味い」といった純粋に好不評を示すような評価の表現とは異なる。例えば、ビールに関するコメントで、「苦くて美味しい」「苦くて不味い」となる場合に、「苦い」が肯定的に捉えられたり否定的に捉えられたりと意見が分かれるように、時代や時期といった時間軸、もしくは、年齢や性別といった筆者の属性によって好不評の極性が変化する表現が嗜好分析の対象となる表現である。

基本的には、2節のステップ1から3を処理して得られる、肯定的な内容(好評表現)と否定的な内容(不評表現)のうち、時間軸や筆者属性の観点で捉えた場合に偏りが生じる表現を嗜好表現と捉えると、その嗜好表現の出現傾向を分析する事によって嗜好を捉えられる可能性が出てくると考えられる。

以上の考えに基づき、デジタルカメラに関する掲示板への書込データを利用して、好評とされる内容と不評とされる内容の出現頻度が2002年と2004年でどの程度異なるかを実験した。

種表現としては、分野への依存度が少ない汎用的な好不評表現を手で定義した辞書を利用した。好評表現が1,417表現、不評表現は2,202表現からなる。好不評表現を認識するエンジンは[3]に記述されたエンジンを用いており、そこで利用している構文パタンの数は153であった。解析したデータと、そこから出てきた評価表現の数は以下の通りである。

2002年のデータは10月及び11月に書込まれた約2万6千件(25,659件)の発言から成る約14万文(144,316文)であり、その中で辞書中の評価表現が認識された件数は延べ17,176件(好評10,415件、不評6,761件)であった。その他の(辞書に

含まれない)表現が好不評文脈から抽出された延べ件数は8,226件(好評4,756件、不評3,470件)であった。

2004年のデータは、5月及び6月に書込まれた約3万6千件(35,669件)の発言から成る約21万文(212,365文)であり、辞書中の評価表現が認識された件数は延べ26,384件(好評16,598件、不評9,786件)であった。辞書以外の表現が好不評文脈から抽出された件数は、延べ13,216件(好評7,876件、不評5,340件)であった。

こうして好不評文脈中で認識された表現のうち、2002年もしくは2004年のデータにおいて、好評か不評のどちらかへの偏りが大きく、しかも、2002年と2004年のデータにおける出現頻度に差が見られる表現から、代表的な表現(辞書中の評価表現も一部含む)を手で選択し、まとめた結果を表3に示す。

表3 デジタルカメラに関する掲示板における2002年と2004年のデータから自動抽出した好不評表現の比較

好不評表現 (嗜好表現 評価表現*)	2002年の データ(書込数 25,659件)に おける件数	2004年の データ(書込数 35,669件)に おける件数
<b>好評が増加</b>		
液晶が大きい	0 (+0, -0)	12 (+12, -0)
撮影を楽しむ*	1 (+1)	7 (+7)
操作が簡単だ	0 (+0, -0)	5 (+5, -0)
価格が安い	3 (+3, -0)	8 (+8, -0)
レスポンスが良い*	7 (+7)	12 (+12)
<b>不評が増加</b>		
シャッタースピードが遅い	1 (+0, -1)	11 (+4, -7)
ノイズが出る*	2 (-2)	7 (-7)
三脚穴がない	0 (+0, -0)	4 (+0, -4)
<b>好評が減少</b>		
サイズが小さい	5 (+3, -2)	1 (+1, -0)
画質が良い*	9 (+9)	5 (+5)
<b>不評が減少</b>		
起動が遅い	5 (+1, -4)	0 (+0, -0)
ピントが甘い	3 (+0, -3)	1 (+0, -1)
光学ズームがない	3 (+0, -3)	3 (+2, -1)
画質が悪い*	7 (-7)	6 (-6)

表3において、評価表現として示されているのは、辞書中で定義されている表現であり、嗜好表現として示されているのは、それ以外の、好不評文脈から抽出した表現である。辞書中で定義されている評価表現の中には「～を楽しむ」「～が良い」のように、基本的には「～」の部分に何が来ても良いという表現が含まれているため、「画質が良い」や「レスポンスが良い」などは「が良い」の部分だけが、評価表現として示されている。また、評価表現に関しては、その表現によって好不評文脈の極性が決まるため、出現件数の全てが同じ極性を取る。

表3の結果からは、デジタルカメラの技術の進歩に従い、ユーザーが好ましいとして求める、もしくは避けようとする内容が変化している様子が見られる。

<sup>1</sup> このようなノイズは、出力中の多様性を高め、異なりを増やす反面、雑多な原因によるため、同じ出力表現に集約されることは少なく、低頻度にとどまることから、高頻度の出力に着目することで、対象から外すことができる。

## 5. 関連研究

好不評表現の極性の一貫性を利用した研究としては、[9][11]を挙げることができるが、これらは、同一文書内での極性が全て一致しているという仮定を適用しており、反転の接続表現によって極性が反転していたり、(例えば数十ページにわたる様な)大きな文章中で多様な好不評文脈が分散していたりするようなケースでは、処理精度が低くなる。

また、反転の接続表現による好不評の極性の反転に関しては、並列構造中の *but* を介した等位レベルの表現の極性が逆になることが指摘されている[4]が、並列構造のみが対象となっている。さらに、意見を示す表現の抽出は、(意見を示す)既知の表現と同文中に出現しやすいという性質を利用した研究[8]も存在するが、文単位の処理である。

好不評表現の学習という観点では、因果関係を表現した文において、肯定的な内容が肯定的な内容を、否定的な内容が否定的な内容を、各々引き起こすという関係から、連鎖的に学習する手法が提案されている[2]。因果関係を正確に認識できれば、比較的確度の高い学習が可能になると考えられるが、構文構造の正確な解析が前提であるため、学習対象の文が、因果関係を含み適切な解析が可能な文に限定される。さらに誤りが発生した場合のノイズを排除する仕組みが自動化できていない。

また、評価表現が、対象名・属性表現・評価表現の組合せで出現する傾向にある点に着目し、対象名・属性表現・評価表現が共起して出現するパターンを定義した上で、既知の表現と未知の表現の組合せが定義パターンに合致する場合に未知の表現を学習する手法が提案されている[5]。パターンを文単位で適用するため、制約が強いパターンはあてはまる文が少なく、制約の弱いパターンによる獲得結果はノイズが大きくなるといった難しさがあり、出力結果は人手で判別する必要がある。

いずれの研究も、複数の文からなる局所的な好不評文脈という捉え方は行っておらず、また、嗜好表現のように時間や個人差によって好不評の極性が異なるような現象は扱っていない。

## 6. おわりに

好不評を示す表現が二極性をもち、さらに同極の表現が連続して出現し易い上、極性が反転する場合にはそれを示す表現を伴うという性質を利用した処理を示した。大量のデータを処理し、高頻度のケースに対して、偏りが大きい部分に着目することで解析エラーなどによるノイズを排除した精度の高い結果を得ることができた。

現行の自然言語処理システムは、殆どが短文単位、もしくは文書単位での処理を基本としている。その原因として、意味的な結束性を持った一塊の文脈の切り出しが困難であることが挙げられる。本稿で示した好不評文脈の認識のように、単文単位でなく、複数の文にまたがった単位の認識と、その処理が可能になることで、意味的結束性の強い文脈内における照応表現の用いられ方や、好評文脈と不評文脈の出現パターンの特徴を分析するなど、文脈処理研究の新たな展開が期待できる。

## 参考文献

- [1] AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications, 2004
- [2] 乾孝司, 乾健太郎, 松本裕治. 出来事の望ましき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
- [3] Hiroshi Kanayama, Tetsuya Nasukawa, and Hideo Watanabe. Deeper Sentiment Analysis Using Machine Translation Technology. In Proceedings of the 20th International Conference on Computational Linguistics, pages 494-500, 2004.
- [4] 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.
- [5] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. テキストマイニングによる評価表現の収集. 情報処理学会研究報告, 自然言語処理研究会, 2002-NL-154, pp.77-84, 2003.
- [6] Tetsuya Nasukawa and Jeonghee Yi. 2003. Sentiment Analysis: Capturing Favorability Using Natural Language Processing. In Proceedings of the Second International Conference on Knowledge Capture (K-CAP). pages 70-77.
- [7] 那須川哲哉, 金山博. 文脈一貫性を利用した極性付評価表現の語彙獲得, 情報処理学会第 162 回自然言語処理研究会, 2004-NL-162 (16), pages 109-116, 2004.
- [8] Ellen Riloff and Janyce Wiebe. Learning Extraction Patterns for Subjective Expressions. In Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 105-112, 2003.
- [9] 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.
- [10] 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.
- [11] Hong Yu and Vasileios Hatzivassiloglou. Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences. In Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 129-136, 2003.