

[DRAFT]

本記事は雑誌「自然言語処理」Vol.13, Num.3, pp.201–241 に掲載されているものの DRAFT 版です。

テキストを対象とした評価情報の分析に関する研究動向

乾 孝司[†]

奥村 学^{††}

インターネットが普及し、一般の個人が手軽に情報発信できる環境が整ってきている。この個人の発信する情報には、ある対象に関するその人の評価等、個人の意見が多く記述される。これらの評価情報を抽出し、整理し、提示することは、対象の提供者である企業や、対象を利用する立場の一般の人々双方にとって利点となる。このため、自然言語処理の分野では、近年急速に評価情報を扱う研究が活発化している。本論文では、このような現状の中、テキストから評価情報を発見、抽出および整理、集約する技術について、その基盤となる研究から最近の研究までを概説する。

キーワード: 評価, 評判, 意見, 感情

A Survey of Sentiment Analysis

INUI TAKASHI[†] and OKUMURA MANABU^{††}

In these days, people can easily disseminate the information including their personal evaluative opinions for some products and services on the Internet. The massive amount of their information is beneficial for both product companies and users who are planning to purchase and use them. Because their information is mainly presented as textual form, in the research field of natural language processing, many researchers have devoted themselves to developing techniques for exploring, extracting, mining, and aggregating the opinions and sentiments. This sort of techniques are commonly called *sentiment analysis*. In this paper, we survey and present the research efforts of sentiment analysis from its fundamentals to the state-of-the-art methods.

KeyWords: *sentiment, affect, reputation, opinion, emotion*

1 はじめに

1.1 背景

インターネットの普及により、インターネット上に膨大でかつ多種多様なテキスト情報が蓄積されるようになって久しい。インターネット上の膨大なテキスト情報を扱うための技術とし

[†] 日本学術振興会特別研究員, Research Fellow of the Japan Society for the Promotion of Science

^{††} 東京工業大学 精密工学研究所, Precision and Intelligence Laboratory, Tokyo Institute of Technology

て、テキスト検索、自動要約、質問応答等さまざまな知的情報アクセス技術に関する研究が活発化しているが、同様にインターネット上の多様なテキスト情報のうち、これまであまり研究対象とされてこなかったものを扱うための技術も研究が活発化してきている。

これまで研究対象とされてきたテキスト情報は、新聞記事、学術論文に代表されるように、事実を記述するものがほとんどであった。それに対し、チャット、Web 掲示板、Weblog 等の普及、利用者の増大に示されるように、インターネット上では、一般の個人が手軽に情報発信できる環境が整うとともに、個人の発信する情報に、ある対象に関するその人の評価等、個人の意見が多数記述されるようになってきている。

この個人の評価に関する情報（評価情報）をテキスト中から抽出し、整理し、提示することは、対象の提供者である企業やサイト運営者、また、対象を利用する立場の一般の人々双方にとって利点となる。このため、自然言語処理の分野では、近年急速に評価情報を扱う研究が活発化している。2004 年春には AAAI のシンポジウムとして評価情報を扱う最初の会議が開催された (Qu, Shanahan, and Wiebe 2004)。国内でも、2004 年度の言語処理学会年次大会では、評価情報の抽出に関連する研究報告が数多く見られた。

そこで本解説論文では、テキストから評価情報を発見、抽出および整理、集約する技術について、その基盤となる研究から最近の研究までを概説することを目的とする。上述したように、この研究領域ではここ数年で爆発的に研究が増大しているが、それらの研究を体系的に整理、概説する解説論文はいまだなく、研究の現状、あるいは今後の方向性を見極めるのに研究者が苦労しているのが現状である。本解説論文がその一助となれば幸いである。

1.2 テキスト評価分析とは？ –本論文で扱う問題領域–

個人の記述する「意見」と言われるものにはさまざまなものが存在する。意見を下位分類するなら、少なくとも以下のようなものがその範疇に含まれることになる。

- 評価を記述するもの、
- 要望、要求、提案の表明、
- 不安、懸念、不満、満足等の感情を表すもの、
- 認識、印象を述べるもの、
- 賛否の表明。

本解説論文では、このうち「評価を記述するもの」を対象とする研究を主に扱う。この分野でのこれまでの研究の多くは、以下の問題を解いているという風に要約できる：

表 1 のような、ある対象の評価を記述しているテキスト断片に対して、その評価が、肯定的な評価（たとえば「良い」）であるか、あるいは、否定的な評価（例えば「悪い」）であるかを推定する。

表 1: 評価を伴うテキスト例

(Morinaga, Yamanishi, Tateishi, and Fukushima 2002) 中の Table 1 から一部を抜粋して再録。

cellular phone A is my favorite.

I am a cellular phone A user, even though it is said to be inconvenient in some ways.

I feel a little unsatisfied with cellular phone C because it has fewer functions than other models.

I'm satisfied with my present phone -cellular phone E-.

You can only download five melodies to cellular phone C, so I recommend cellular phone B.

本稿では、このような評価に関する分析をテキスト評価分析と呼び、テキスト評価分析を取り巻く諸研究の現状を紹介する。

この問題は、もう少し具体的には、肯定的な評価 / 否定的な評価の 2 値分類として定式化されることが多い。また、問題は、テキスト断片の粒度によって、次の 3 つに大別できる。

- 語句レベル
- 文レベル
- 文書レベル

例えば、表 1 は文レベルでの 2 値分類である。

言うまでもなく、このテキスト断片の粒度ごとに問題の性質は大きく異なる。それぞれの詳細については、3 節で述べる。

1.3 用語の整理

背景思想の違いの影響などもあり、テキスト評価分析で利用される用語は各研究者間で統一されているとは言い難い。そのため、しばしば同一概念が論文間において異なった用語で参照されている。

本稿では、個人の評価に関する情報を評価情報、評価情報の良い / 悪いに関する軸を評価極性と呼ぶ。ある評価情報が良い評価をもつことを肯定極性をもつと呼び、逆に悪い評価をもつことを否定極性をもつと呼ぶ。また、肯定極性が否定極性をもつ評価情報がテキスト内で記述された表現を評価表現と呼ぶ。

表 2 に、本稿での用語に対応する、紹介論文において使用される代表的な用語を示す。表 2 の評価極性値とは、肯定極性と否定極性の間を連続的に捉え、各評価極性の強さを数値化したものである。評価極性値は、 $[-1, 1]$ の範囲の実数値として与え、正側が肯定極性、負側が否定極性に割り当てられることが多い。

表 2: 用語の対応

本稿での用語	紹介論文において使用される代表的な表現
評価情報	sentiments, affect parts of opinions, reputation, 評判
評価極性	semantic orientations, polarity, sentiment polarity
肯定極性 (肯定)	positive, thumbs up, favorable, desirable, 好評
否定極性 (否定)	negative, thumbs down, unfavorable, undesirable, 不評
評価極性値	semantic orientation score, SO-score
評価表現	sentiment expression, word with sentiment polarity

1.4 本論文の構成

本論文の構成は以下の通りである。まず、2 節では、テキスト評価分析の題材となるテキストデータについて述べる。3 節では、テキスト評価分析を支える各要素技術に関する諸研究を紹介する。続く 4 節では、テキスト評価分析の応用研究を紹介し、5 節で、テキスト評価分析に関連するその他の話題を紹介する。最後に、6 節で、テキスト評価分析で今後取り扱うべき課題を述べ、7 節で本論文をまとめる。

2 テキスト評価分析の題材となるテキストデータ

1 節で述べたように、この研究領域が活発化することになった 1 つの要因は、インターネット上に、対象となるテキスト集合が豊富に蓄積されるようになったことが挙げられる。しかし、評価分析の題材となるテキストは、Web 上にしか存在しないわけでもないし、Web 上に蓄積される以前には存在しなかったわけでもない。アンケート調査等における自由回答アンケートのテキストや、企業等のカスタマーサポートセンターに蓄積されている「お客様の声」等は、まさに評価分析の題材として以前から分析の対象となっていた。

評価分析の題材となるテキストは、大きく次の 2 つに分類することができる。

- 意見の収集、集約が目的となっているテキスト
 - 社会調査等による自由回答アンケート
 - カスタマーサポートセンターにおける「お客様の声」
 - レビュー（以下は、テキスト評価分析で利用される代表的なレビューサイト）
 - * Rotten Tomatoes (<http://www.rottentomatoes.com/>)
 - * Epinions.com (<http://www.epinions.com/>)
 - * Amazon.com (<http://www.amazon.com/>)

* Amazon.co.jp (<http://www.amazon.co.jp/>)

- 潜在的に意見を含むテキスト
 - － チャット
 - － Web 掲示板
 - － Weblog

前者は、まさにユーザ(個人)に意見を述べてもらうことを目的にして収集されたテキストである。そのため、テキスト中に意見が含まれる割合は比較的大きく、また、テキストの性質としても、ユーザが書いたものを校正したりしている場合等もあり、比較的良質であることが多い。さらに、意見の対象となっている内容が特化されており、話題が限定的であることも多い。

一方、潜在的に意見を含むテキストでは、テキスト中の意見部分の割合は前者に比べると小さくなり、また、Web上のテキストであるチャットや掲示板内の記述は、言語処理の精度が著しく劣化する可能性があるほど、テキストとしては質が悪いことが多い。Weblogのテキストは、それらに比べれば、比較的テキストの性質は良いと言える。また、テキストの内容は、掲示板の一部で話題をスレッドごとに限定している場合を除けば、テキスト中で雑多なことを記述している可能性が高いと言える。

本稿では、このような性質をもつ一連のテキストデータを合わせて評価文書と呼び、評価文書から評価情報を抽出し、整理し、あるいは提示する技術について紹介していく。

3 テキスト評価分析の要素技術に関する諸研究

テキスト評価分析を支える要素技術に関する研究は、注目するテキスト断片の粒度によって次のように分けられる。

- 評価表現辞書の構築に関する諸研究(3.1節)
- 評価情報を観点とした文書分類に関する諸研究(3.2節)
- 評価情報を含む文の抽出に関する諸研究(3.3節)
- 評価情報の要素組の抽出に関する諸研究(3.4節)

各要素技術の関係を図1に示す。また、各要素技術で用いられる手法の概観を表3にまとめる。

3.1節で述べる評価表現辞書とは、評価表現とその表現がもつ評価極性の組(例:「良い-肯定」)の集合である。この辞書はその他の要素技術において基礎的な知識として利用される。3.2節から3.4節までの残りの話題では、それぞれ文書、文、要素組という単位で評価情報とその評価極性を求める課題を主に扱う。3.4節で述べる評価情報の要素組の抽出に関する研究では、文書や文より細かい語句単位の粒度で評価情報を扱う。

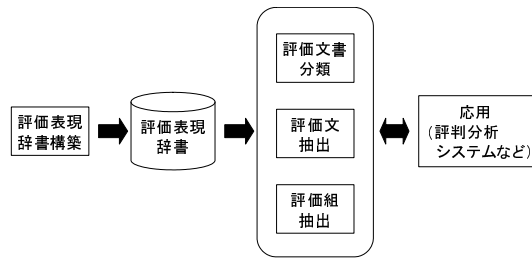


図 1: テキスト評価分析の要素技術間の関係

いずれの研究も互いに関連をもっているが、特に、3.1 節と 3.4 節で扱う話題はいずれも語句レベルの話題であるという点で似ている。しかし、3.1 節での話題は文脈情報とは独立な単語単体の評価極性に焦点が当たっているが、3.4 節では実際の文脈情報を含めた評価極性に焦点が当たっている点に注意して読み進めてほしい。

表 3: テキスト評価分析を支える要素技術の概観

要素技術	アイデア・手法	必要な資源	備考（残された課題）	代表的な論文
評価表現辞書の構築 (3.1 節)	語彙ネットワーク 共起情報 周辺文脈	国語辞書・シソーラス 大規模コーパス -	エントリのない語句への対応 大規模コーパスの入手方法 適切な周辺文脈の検討	(Kamps et al. 2004) (Turney 2002) (那須川 金山 2004)
評価情報を観点とした 文書分類 (3.2 節)	評価表現の比率 教師あり機械学習	評価表現辞書 教師信号つき文書集合	評価表現辞書への依存性 有効な素性情報の検討	(Turney 2002) (Pang et al. 2002)
評価情報を含む文の抽出 (3.3 節)	3.2 節と同様	3.2 節と同様	文単位に特有な特徴の考慮	(Yu and Hatzivassiloglou 2003) (Nigam and Hurst 2004)
評価情報の要素組の抽出 (3.4 節)	要素抽出 関係抽出 動的極性判定	- 教師信号つき事例集合 -	要素, 要素組の定義 -	(小林他 2005) (飯田他 2005) (鈴木 2005)

3.1 評価表現辞書の構築に関する諸研究

評価表現辞書とは，評価表現とその表現がもつ評価極性の組（例：「良い - 肯定」）の集合である．評価表現辞書の構築に関する研究は，構築に主に利用する情報によって，大きく3つに分けられる．以下，順に紹介する．

- 語彙ネットワークを利用した手法（3.1.1 節）
- 共起情報を利用した手法（3.1.2 節）
- 周辺文脈の情報を利用した手法（3.1.3 節）

3.1.1 語彙ネットワークを利用した手法

語彙ネットワークを利用した評価表現辞書構築の手法は，既存の語彙知識の情報を利用して，評価表現の評価極性を求める．もう少し具体的には，まず，シソーラスや国語辞書の情報を基にして，語彙をノードとする語彙ネットワークを用意する．そして，評価極性情報をネットワーク上で伝播させることによって，語彙ネットワーク中のすべてのノード（語彙）の評価極性を求める．

ここではまず，(Kamps et al. 2004) と (Hu and Liu 2004a) の研究を紹介する．これらの研究は，いずれも「類義関係にある語彙の評価極性は一致しやすい」という仮定に基づいてアルゴリズムが構成されている．

(Kamps et al. 2004)

Kamps et al. (Kamps et al. 2004) は，WordNet (Fellbaum 1998) の情報を利用して形容詞の評価極性を判定する手法を提案している．WordNet 中の形容詞は，類義関係 (synonymy) のリンクで結ばれ，ネットワーク構造を成している．Kamps らは，このネットワーク内の隣接情報を利用する．

Kamps らの基本的なアイデアは，まず，肯定極性を代表する語と否定極性を代表する語と選定する．彼らの評価実験では，「good」と「bad」をそれぞれ選んでいる．そして，評価極性を判定したい形容詞 t がネットワーク内において「good」と「bad」のどちらに近いかを考え，「good」に近ければ t を肯定極性とみなし，逆に「bad」に近ければ t を否定極性とみなす．

具体的には，形容詞 t の評価極性値 $SO-score(t)$ は，次の式で求められる；

$$SO-score(t) = \frac{d(t, bad) - d(t, good)}{d(good, bad)} . \quad (1)$$

ここで， $d(t_i, t_j)$ は，2つの形容詞 (t_i と t_j) 間の最短経路長を示す．この式のもとで， $SO-score(t)$ の値が正であれば肯定極性であると判定し，負であれば否定極性であると判定する．

(Hu and Liu 2004a)

Hu et al. (Hu and Liu 2004a) も WordNet を使用して形容詞の評価極性を判定する手法を提案している。Hu らは、Kamps らが注目していた類義関係の情報に加え、反義関係 (antonymy) の情報も利用している。

Hu らの手法では、まず、少数の、評価極性が既知である語集合 (Hu らの評価実験では 30 語) を種表現として用意する。そして次に、種表現から出発し、順番に WordNet 内の類義関係と反義関係を辿る。この時、類義関係を辿った場合は類義関係の辿り元にある形容詞と同じ評価極性を割り当て、反義関係を辿った場合は反義関係の辿り元にある形容詞と逆の評価極性を割り当てる。この操作を辿り先がなくなるまで繰り返すことで、WordNet 内の形容詞の評価極性を求めた。

以上の 2 つの研究は、共に形容詞に注目している。WordNet 内に含まれる形容詞以外の評価表現の分布については、(Strapparava and Valitutti 2004) が参考になる。

次に、WordNet ではなく、コーパスから語彙ネットワークを作成している (Hatzivassiloglou and McKeown 1997) の手法を紹介する。

(Hatzivassiloglou and McKeown 1997)

Hatzivassiloglou et al. (Hatzivassiloglou and McKeown 1997) は、コーパス中の接続詞の情報を手がかりにして語彙ネットワークを構築し、そこから形容詞の評価極性を判定する手法を提案した。これは「順接関係の接続詞で結ばれる形容詞は同じ評価極性をもつ。逆に、逆接関係の接続詞で結ばれる形容詞は異なる評価極性をもつ」という考え方に基づいた手法である。

まず、コーパスから「形容詞 - 接続詞 - 形容詞」という品詞の並びをもつ語系列を抽出する。そして、形容詞をノードとし、語系列内の 2 つの形容詞の間にリンクを張ることで語彙ネットワークを作成する。この時、リンクには、接続詞のタイプ情報などを利用して、リンクが結んでいる 2 つの形容詞の評価極性が一致するか否かに関する情報が同時に付与される。例えば、接続詞が「and」であれば評価極性が一致、「but」であれば評価極性が不一致という情報をリンクに付与する。次に、クラスタリング・アルゴリズムを用いてノードをまとめ上げることによって、語彙ネットワークを 2 つのグループに分割する。この時、片方のグループには肯定極性となる形容詞、もう片方のグループには否定極性となる形容詞が入るようにクラスタを構成する。そして、最終的に各グループの評価極性を定めることで、語彙ネットワーク内の全ての形容詞の評価極性を求める。

ここまでで紹介した 3 つの辞書獲得手法は、いずれも語彙ネットワーク内の局所的な情報から新たな評価極性を決定しており、語彙ネットワーク全体の情報を十分に活用できていない。

Takamura et al. (Takamura, Inui, and Okumura 2005) は、この問題点に対し、spin glass (Chandler 1987; Mackay 2003; Inoue and Carlucci 2001) の枠組みを利用することによって、語彙ネットワーク全体の大域的な情報を活用するモデルを提案している。また、Takamura らは、先行研究での知見から、語彙ネットワークの作成でも工夫を凝らしており、国語辞書の情報をベースにして、シソーラス、コーパスの情報を加えた語彙ネットワークを作成している。国語辞書の情報を評価極性辞書構築に利用するという着想自体は (小林, 乾, 乾 2001) が初めに提案したものである。

3.1.2 共起情報を利用した手法

共起情報を利用した評価表現辞書構築の手法では、肯定極性をもつ典型的な表現 (「good」や「excellent」) と否定極性をもつ典型的な表現 (「bad」や「poor」) を種表現として始めに用意しておき、種表現と共起する比率に従って語句の評価極性を判定する。この手法は「肯定極性をもつ語句の周辺文脈には肯定極性をもつ語句が現れやすく、否定極性をもつ語句の周辺文脈には否定極性をもつ語句が現れやすい」という考え方に基づく。

先の Hatzivassiloglou らの手法でも「形容詞 - 接続詞 - 形容詞」という共起情報が利用されていたが、そこでは、共起情報は語彙ネットワークを通して、間接的に形容詞の評価極性の判定に利用された。ここでは、より直接的に共起情報を単語の評価極性に反映させる (Turney 2002) の手法を紹介する。

(Turney 2002)

Turney (Turney 2002) は、コーパスから得られる共起情報を利用して語句の評価極性値を判定した。この手法は、国語辞書の見出し語や、WordNet のエントリ情報を利用しないため、見出し語単位やエントリ単位だけでなく、複数語からなる句に対しても評価極性値を判定できる。

ある評価表現 t の評価極性値 $SO-score(t)$ は式 (2) で算出される；

$$SO-score(t) = PMI(t, \text{“excellent”}) - PMI(t, \text{“poor”}) . \quad (2)$$

ここで、PMI(pointwise mutual information)(Church and Hanks 1989) は、2つの語句間の共起を測る尺度であり、任意の語句 a, b の間の PMI は、

$$PMI(a, b) = \log_2 \frac{p(a, b)}{p(a)p(b)} \quad (3)$$

で計算される。つまり、評価極性値を判定したい語句 t が肯定極性を示す代表的な語「excellent」と共起しやすければ、 $SO-score(t)$ は正に大きい値をとり、否定極性を示す代表的な語「poor」と共起しやすければ、逆に負に傾く。

Turney の手法は、国語辞書やシソーラスなどの言語資源を一切必要としないため手軽さがあるが、信頼性の高い共起情報を得るには巨大なコーパスを必要とすることに注意しなければならない。元論文では、World Wide Web 空間の文書全体をコーパスと見立て、検索エンジンを用いて語句 a, b 間の共起情報 $PMI(a, b)$ を得ている。同一著者による (Turney and Littman 2002, 2003) では、(Turney 2002) の考えを推し進め、種表現を複数用いた手法や、PMI の代わりに潜在的意味解析 (semantic latent analysis; LSA) (Seerwester, Dumais, Furnas, Landauer, and Harshman 1990; Landauer and Dumais 1997) を用いた手法も検討している。

3.1.3 周辺文脈の情報を利用した手法

周辺文脈の情報を利用した評価表現辞書構築の手法では、まず、評価極性が既知である少数の種表現を幾つか用意する。そして、辞書構築アルゴリズムでは、繰り返し過程の中で、種表現から評価表現を順次増やしていくブートストラップ的な戦略をとる。

一般に、ブートストラップ法では、注目している対象（ここでは評価表現）とその周辺文脈情報を交互に学習させる。現在までに、注目する周辺文脈情報の違いによって、幾つかの辞書構築手法が提案されている (小林他 2001; 乾, 乾, 松本 2004; 中山, 江口, 神門 2004; 小林他 2005)。多くの場合、周辺文脈には特定の言語パターンが想定される。ここでは、「文脈一貫性」という非常に汎用性の高い概念に基づいて周辺文脈を特定した (那須川・金山 2004) の手法を紹介する。

(那須川・金山 2004)

那須川ら (那須川・金山 2004) は、評価表現の周辺文脈に関する以下の仮定に基づき、ブートストラップ的に評価表現を収集した。

「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が形成されることが多く、その中では、明示されない限り、肯定 / 否定の極性が一致する傾向にある。」

つまり、ある評価表現 t の周辺文脈に注目した場合、「しかし」や「～だが」などの逆接関係を導く表現が存在しなければ、文脈中にある評価表現の評価極性は t の評価極性と等しくなり、逆に、逆接関係を導く表現が存在すれば、それ以降の極性は t の極性から反転すると考える。

まず、評価極性が既知である少数の評価表現を種表現として用意し、上記の仮定に従って、文書内からブートストラップ的に評価表現候補を収集する。複合的な評価表現に対応するため、用言と同時に、用言に最も近い格助詞句を加えた複合表現を評価表現候補として選んでいる。考慮する周辺文脈は、評価表現 t と同一文内の節、前文の主節、後文の主節としている。

種表現が「満足する - 肯定」である場合に抽出される評価表現候補の例を (1) に示す。

(1) a. 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。

b. 評価表現候補 ; 「綺麗だ - 肯定」「綺麗だ<画像 : が> - 肯定」「不要だ - 否定」「不要だ<デジタルカメラ : が> - 否定」「言える - 否定」「言える<文句 : を> - 否定」

この例では、「ところが」が逆接を導く表現である。「言える」「言える<文句 : を>」が肯定極性ではなく否定極性となっているのは極性反転子（文末の「～ません」）の効果による（極性反転子の詳細については、3.2.1 節の Kennedy et al. (Kennedy and Inkpen 2005) の箇所で述べる）。

上記の評価表現候補の作成法では、実際には評価極性をもたない評価表現候補も多く登録してしまうことになる。そこで、そのような評価表現候補を最終的な評価表現として登録することを避けるために、ある一定の基準を満たす評価表現のみ評価表現辞書に登録する。基準には、評価表現の出現頻度情報や、各評価表現に対してアルゴリズムから求められた評価極性の割合の情報などが利用される。

3.1.4 まとめ

評価表現辞書の構築に関する研究を、語彙ネットワークを利用した手法（3.1.1 節）、共起情報を利用した手法（3.1.2 節）、周辺文脈の情報を利用した手法（3.1.3 節）の3つに分類し、各手法の代表的な論文を紹介した。

現在の状況では、各手法はそれぞれに欠点がある。語彙ネットワークを利用した手法は、語彙ネットワークを作成する際に国語辞書やソーラスなどの既知の言語資源を利用することが多いが、この場合、新語など、既知の言語資源にエントリのない語句への対応は困難である。共起情報を利用した手法は、信頼性の高い共起情報を得るために必要な巨大なコーパスをどのように用意すればよいかという課題が残る。また、現状の周辺文脈の情報を利用したブートストラップに基づく手法では被覆率が低いという問題がある。

評価表現辞書構築の精度評価には、The General Inquirer (Stone, Dunphy, Smith, and Ogilvie 1966) (URL: <http://www.wjh.harvard.edu/~inquirer/>) がしばしば利用される。これは、元はテキスト内容分析 (Stone et al. 1966) のために作成された言語知識データであり、単語毎に幾種類かのラベルが付与されている。その中の“Positiv”ラベルと“Negativ”ラベルの情報が主に辞書構築の精度評価に利用される。The General Inquirer を用いた評価表現辞書構築の手法の比較については、(Takamura et al. 2005) や (高村, 乾, 奥村 2005a) に記述がある。

評価表現辞書は、実際の文書中の文脈とは独立した静的な知識である。しかし、後述する3.4 節の内容とも関連するが、ある表現の評価極性は文脈に依存して変化することがある。評価表現辞書のエントリとして許容される範囲は、個々の応用ごとに個別に決定しなければならない。

3.2 評価情報を観点とした文書分類に関する諸研究

本節では、評価情報を観点とした文書分類に関する研究を紹介する。評価情報を観点とした文書分類とは、ある評価文書が肯定極性か否定極性のいずれの極性をもつかを判定する課題である。この分類課題では、前節で述べた評価表現が主要な情報として利用される。以下本稿では、既存のトピックに基づく文書分類との混乱を避けるために、評価情報を観点とした文書分類を特に評価文書分類と呼ぶ。

評価文書分類を実現することは、次のような状況において有益な情報を提供する。例えば、デジタルカメラ-Aに関する評価文書に関して、肯定極性をもつ評価文書のみをまとめることで、そのデジタルカメラの優れている点を把握することができる。この情報をもとにして、デジタルカメラ-Aの潜在的な購入者は、実際に購入を決断するかも知れない。また、否定極性をもつ評価文書のみをまとめることで、現時点でのデジタルカメラ-Aの欠点が把握できる。デジタルカメラ-Aの開発者は、この情報をもとにして改良を施すことができる可能性がある。

評価文書分類の手法は、(Turney 2002)を起源とする評価表現の比率に基づく手法(3.2.1節)と、(Pang et al. 2002)を起源とする機械学習に基づく手法(3.2.2節)に大きく分かれる。以下、順に紹介する。

3.2.1 評価表現の比率に基づく手法

評価表現の比率に基づく評価文書分類の手法では、評価文書中に現れる評価表現に注目し、肯定極性をもつ評価表現と否定極性をもつ評価表現の出現比率に従って、評価文書全体の評価極性を求める。つまり、肯定極性をもつ評価表現が否定極性をもつ評価表現に比べて多く出現している評価文書を肯定極性をもつ評価文書であると判定する。逆に、否定極性をもつ評価表現が肯定極性をもつ評価表現に比べて多く出現している評価文書を否定極性をもつ評価文書であると判定する。

以下ではまず、(Turney 2002)の研究を紹介する。

(Turney 2002)

Turney の分類手続きは3つのステップから構成される。

- ステップ1 評価文書に含まれる評価表現を評価極性値付きで抽出する。これには、3.1.2節で述べた手法が利用される。
- ステップ2 抽出された評価表現の極性値の平均値(平均極性値)を求める。
- ステップ3 平均極性値の符号に応じて、評価文書全体の評価極性を決定する。平均極性値が正であれば肯定極性とし、そうでなければ否定極性とする。

Turney は評価極性値を求める語句に制限を設けており、「low fees」のような、形容詞（あるいは副詞でも可）を含む句に対してのみ評価極性値を求めている。これは、「形容詞は文書内に含まれる意見や評価を特定する際のよい指標となる」(Hatzivassiloglou and Wiebe 2000; Wiebe 2000; Wiebe, Wilson, and Bell 2001) という先行研究の指摘に基づく。また、句を選択しているのは、指標の文脈や領域依存性へ対応するためである。例えば、「unpredictable」という形容詞を考えた場合、自動車の評価文書内では「unpredictable steering」のように否定極性となるが、映画の評価文書内では「unpredictable plot」のように、逆に肯定極性となる。このような状況では、形容詞「unpredictable」単体は、肯定的な評価文書と否定的な評価文書のどちらにも出現し、評価文書の評価極性を判定する際のよい指標とはならない。その一方で、「unpredictable steering」や「unpredictable plot」といった「unpredictable」の前後文脈を追加した語句は、評価極性が現れる文脈を適切に捉えており、よい指標になる。特に、複数の評価対象（自動車、映画など）が含まれる評価文書集合を扱う状況では、単語単体ではなく句を利用することが望ましいと考えられる。

Turney の手法は、3.1.2 節で述べた評価極性値の計算式が対数オッズに対応することに注意すると、ナイーブベイズ分類器に基づく文書分類と接点をもつことがわかる。Beineke et al. (Beineke, Hastie, and Vaithyanathan 2004) らは、ナイーブベイズ分類器の観点から Turney の手法を捉え直し、Turney の手法の拡張を行っている。

次に、評価文書分類に有効な素性情報について検討した研究を紹介する。ここでは、先述の Turney の手法に基づいたアプローチを採用したものを取り上げる。機械学習に基づく評価文書分類における素性情報に関する検討については 3.2.2 節で紹介する。

(Taboada and Grieve 2004)

Taboada et al. (Taboada and Grieve 2004) は、(Turney 2002) の手法の拡張版を用いて¹、評価表現の出現する位置情報が評価文書分類に与える影響について調査している。この研究の背景には「評価文書中において、書き手の主要な意見は評価文書全体に均等に現れるのではなく、特定の部分に集中して現れる」という仮定がある。

Taboada らは、評価表現が出現する位置に応じて、各評価表現がもつ評価極性値を修正することで、位置情報の有効性を検討した。具体的には、まず、文書内の位置ごとに人手で重みを設定する。そして、評価表現が出現する位置に応じて、人手で定めた重みを評価極性値に乗じることによって評価極性値を修正する。出現位置に対する重みの設定を変更させながら、重みの設定と分類精度との関係を調査したところ、評価文書の後半 2/3 の位置に現れる評価表現へ

¹ (Taboada and Grieve 2004) の基本的な分類手続きは (Turney 2002) と同じであるが、分類手続きの 3 ステップ目が拡張されている。

の重みを最も高くした場合に最も良い分類精度が得られたと報告している。

江崎ら(江崎, 松井, 大和田 2005)も日本語の Weblog 記事を対象にして, 形容詞の出現位置を考慮した同様の実験を実施している。実験結果によると, Weblog 記事の前方側に現れる評価表現への重みを最も高くした場合に最も良い分類精度が得られたと報告している²。

次に, 極性変化子が評価文書分類に与える影響について調査した Kennedy et al. (Kennedy and Inkpen 2005) の研究を紹介する。単語の中には, それ単体は評価極性を持たないが, 評価極性をもつ単語を修飾することでその評価極性を変化させるものがある。このような評価極性を変化させる単語のことを極性変化子 (contextual valence shifter) と呼ぶ (Polanyi and Zaenen 2004)。極性変化子には, 評価極性を肯定から否定, あるいは否定から肯定に反転させたり, 極性の強さを変化させるものがある。前者を反転子 (negations), 後者を強調子 (intensifiers) と呼ぶ。代表的な反転子の例としては「not」や「never」などがある。また, 強調子の例としては「very」や「deeply」などがある。

(Kennedy and Inkpen 2005)

Kennedy et al. (Kennedy and Inkpen 2005) は極性変化子に注目し, 極性変化子が評価文書分類に与える影響を検証した。トピックに基づく文書分類と比べて, 評価文書分類では極性変化子, 特に, 極性反転子の処理に注意を払う必要がある。例えば, 形容詞「good」は, 肯定極性を示す代表的な単語であるが, これに「not」が付いた「not good」は, 逆の評価極性, すなわち否定極性を示す。評価文書分類を精度よくおこなうには, この特性をうまく素性情報として取り込むことが重要になる。

Kennedy らは, 評価文書分類の手法として (Turney 2002) の手法を用い, 次のようにして極性変化子の情報を取り込んだ。

- 評価表現と同一の節内に極性反転子がある場合, 評価表現の評価極性を反転させる。ある評価表現の評価極性が肯定極性であれば否定極性に, 否定極性であれば肯定極性にする。
- 評価表現と同一の節内に極性強調子がある場合, 強調子に応じて評価表現の評価極性値を増減させる。強調の変化子であれば評価極性値を増やし, 抑制の変化子であれば評価極性値を減らす。

評価実験を通して, 極性変化子を考慮した方が, 極性変化子を考慮しない場合よりも分類精度が向上することを確認した。

² (Taboada and Grieve 2004) と (江崎他 2005) の実験結果には大きな相違がある。これには, 対象としている評価文書やその評価文書の使用言語など, 幾つかの要因が絡んでいるだろう。

3.2.2 教師あり機械学習に基づく手法

ここでは、教師あり機械学習に基づく評価文書分類を扱った研究を紹介する。これまで、トピックに基づく文書分類では、さまざまな機械学習手法が適用されており、高い分類精度を達成している (Sebastiani 2002)。このような背景を考えれば、評価文書分類に機械学習手法を適用することは自然な流れであると言える。

まず、評価文書分類に初めて機械学習を適用した Pang et al. (Pang et al. 2002) の研究を紹介する。

(Pang et al. 2002)

Pang et al. (Pang et al. 2002) は、トピックに基づく文書分類で有効であった教師あり機械学習に基づく分類手法が、評価文書分類にも有効であるかどうかを実験的に検証した。用いた分類器は、トピックに基づく文書分類でもしばしば適用される、ナイーブベイズ分類器 (Mitchell 1997)、最大エントロピー法に基づく分類器 (Berger, Pietra, and Pietra 1996)、サポートベクトルマシン分類器 (Vapnik 1995) である。学習に利用する素性情報には、単語 uni-gram、単語 bi-gram など、トピックに基づく文書分類で一般的に利用される情報のみを利用している。

評価実験から次のような知見を得た。

- 評価表現は形容詞となることが多いが、単語 uni-gram として、評価文書中のすべての単語を利用した場合の結果は、形容詞のみを利用した場合の結果よりも高い精度を得た。このことは、形容詞以外の単語が、評価文書の評価極性を判定する際のよい指標となっていることを示している。これについては、(Salveti, Lewis, and Reichenbach 2004) でも同様の報告がされている。
- 単語 uni-gram 素性を利用して評価文書分類を行った場合の精度は、同一の学習手法と同一の素性情報を用いたトピックに基づく文書分類の精度よりも低かった。このことは、評価文書分類は、トピックに基づく文書分類よりも難しく、単純な単語 uni-gram よりも複雑な情報が必要であることを示している。

Pang らが評価実験の際に作成したデータセットは、共通の評価実験用データとして多くの研究者に利用されている。このデータセットは以下の URL から入手できる。

Pang's movie review data ;

<http://www.cs.cornell.edu/people/pabo/movie-review-data/>

(Pang et al. 2002) の結果を受けて、評価文書分類に有効な素性情報についての検証がこれまでにされている。Mullen et al. (Mullen and Collier 2004) は、サポートベクトルマシン分類器を用いた評価文書分類において、(Turney 2002) や (Kamps et al. 2004) の手法によって得られた評

価表現の評価極性値を素性情報に用いた。また、評価表現の位置情報も利用しており、評価対象を表す語に近い位置にある評価表現が評価文書分類に有効に働くことを示した。

Pang et al. (Pang and Lee 2004) は、「事実文に含まれる評価は対象に関する書き手の評価とは関係がなく、意見文に含まれる評価のみが対象に関する書き手の評価と関係する」と仮定して、文の主観性 (subjectivity) に注目した。Pang らは、分類の第 1 段階として、評価文書中の文を意見文と事実文に分け、意見文のみを抽出する。その後、抽出された意見文集合のみを対象として、(Pang et al. 2002) の機械学習手法を適用して評価文書を分類した。

Matsumoto et al. (Matsumoto, Takamura, and Okumura 2005) は、サポートベクトルマシン分類器を用いる場合の素性情報として、単語 uni-gram, 単語 bi-gram に加えて、語の系列および語間の依存構造木の情報を利用した。また、Bai et al. (Bai, Padmanand, and Airolodi 2004) は、ナイーブベイズ分類器で仮定される語間の統計的独立性の仮定を排除するために、ベイジアンネットワークの一つである Markov Blanket Directed Acyclic Graph (Pearl 1988) を用いた。現在のところ、Pang's movie review data を用いた評価実験の結果の中では、(Matsumoto et al. 2005) や (Bai et al. 2004) が、その他の手法と比べて高い精度を達成している (実際の数値については各論文を参照されたい)。

ここまでで紹介した評価文書分類の諸研究では、必ず肯定極性が否定極性のどちらかに分類できる評価文書を扱っている。これは、しばしば利用される Pang's movie review data がそのように構成されていることに一因がある。

しかしながら、現実には、このようにすべての文書が必ず肯定極性が否定極性のどちらかに分類できる状況でないことも多い。次に紹介する Gamon の研究 (Gamon 2004) は、カスタマーサポートセンターに届けられた文書を扱った。Gamon のデータは、Pang's movie review data などと比べて、断片的で、短いコメントが多く、そのため、本質的に評価極性が決定できない文書も含んでいる。

(Gamon 2004)

Gamon (Gamon 2004) は、カスタマーサポートセンターに届けられた顧客レビューを題材にして、評価文書分類を行った。

Gamon は、まず、データの特性を知るための事前調査を行っている。収集された顧客レビューにはあらかじめレビューを書いたユーザが付与した評点 (rating) が付与されている。200 サンプルを人が肯定極性が否定極性かの 2 クラスに分類したところ、約半数の 117 件しか正確に分類できなかった。残りのサンプルは、評価極性の判定に必要な情報が含まれていなかったり、評価極性をもつような内容が書かれていなかったり、肯定と否定の評価極性が混在していた。この結果から、Gamon のデータは、全ての文書が必ずしも評価極性の判定に適した文書と

なっているわけではなく、ノイズを多く含んでいることがわかる。

Gamon は、評価文書分類にサポートベクトルマシン分類器を用いる際、上記のようなノイズを多く含む文書の分類に有効な素性を明らかにするために幾つかの種類の素性情報を利用した。まず、分類に利用する素性情報を、浅い言語解析のみで獲得できる素性（表層素性）と、獲得するために深い言語解析を必要とする素性（言語素性）に分けた。表層素性には、原形化された単語の uni-gram, bi-gram, tri-gram が含まれている。また、言語素性は、品詞 tri-gram, 文や節の長さ、構文構造、句間の意味関係、他動性、時制などが含まれている。句間の意味関係には、例えば、“Verb-Subject-Noun（ある名詞 Noun はある動詞 Verb の主体 Subject を表す）” などがある。

上記のそれぞれの素性情報の有効性を検証したところ、表層素性に加えて、言語素性を用いることで分類精度が向上することを確認した。また、ほぼすべての素性情報の組合せにおいて、対数尤度比 (Dunning 1993) の高い上位 n 個の素性のみを選択的に利用する素性削減法を適用することによって、さらに高い精度を得た。

Gamon は、肯定極性でも否定極性でもない（以降、中立と呼ぶ）文書の存在について言及しているが、問題設定としては従来と同じく 2 値の分類問題に設定した。2 値分類ではなく、中立クラスを加えた 3 値の評価文書分類に初めて取り組んだのは (Koppel and Schler 2005) である。Koppel らは、まず、予備調査として、2 値分類器を学習し、それをそのまま、肯定 / 否定 / 中立の 3 値分類に援用する方法では低い精度しか達成できないことを示した。そして、stacking (Wolpert 1992) に基づく 3 値の評価文書分類を行う手法を新たに提案している。

3.2.3 より細かい分類粒度へ

これまでは、評価文書分類として、肯定極性が否定極性かの 2 値の分類問題を扱うことが多かった。この点から、中立クラスを加えた (Koppel and Schler 2005) の 3 値分類を扱った事例は、これまでになく、かつ、現実のデータに即した視点を評価文書分類に導入したと言える。

さらに、分類クラスの 2 値からの自然な拡張として、評価文書をより細かい分類粒度へ分類する課題を扱った研究がある (Pang and Lee 2005; 岡野原 辻井 2005)。これらの研究はいずれも中立クラスを明示的に取り込むことを目的としているわけではないが、細かな粒度を扱うことによって、結果的に中立クラスが反映されていると考えられる。

(Pang and Lee 2005)

Pang et al. (Pang and Lee 2005) は、2 値よりも粒度の細かい映画レビューの評点 (rating) に注目し、各映画レビューをそれに付与された評点毎に分類する（以下、評価文書評点分類）こ

とを試みている。

多値分類への素直な対応として、分類ではなく回帰の考え方を導入することが考えられる。しかし、映画レビューの評点が順序尺度であることを考慮すると、単純な回帰では適切に処理できない可能性がある。この点については、(Koppel and Schler 2005)でも議論されている。この問題を扱うために、Pangらは、metric labeling法 (Kleinberg and Tardos 1999)を適用することによって、既存の分類器によって得られた評点分類結果を補正することを提案した。

高い類似度をもつ評価文書群は、同じ評点を持つと考えられる。metric labeling法は、この考えを自然に明示的に取り込むことができ、評価文書間の類似度が高い場合に、それらの評価文書がもつ評点間の差を小さくするように学習が進む。metric labeling法では、評価文書間の類似度関数を利用するが、Pangらは、肯定極性文比率 (positive-sentence percentage; PSP) という尺度に基づく類似度関数を提案している。これは、評価文書中の全 (意見) 文における肯定極性文の比率を示す (すなわち、「肯定極性文の数」÷「評価文書中の全 (意見) 文の数」で求められた値)。

Pangらの調査から、評価文書の肯定極性文比率は、その評点と高い相関をもつことがわかっている。

2つの評価文書 r_1 と r_2 に対して、以下の類似度関数を提案している；

$$\text{sim}(r_1, r_2) = \cos(\overrightarrow{\text{PSP}(r_1)}, \overrightarrow{\text{PSP}(r_2)}) . \quad (4)$$

$\overrightarrow{\text{PSP}(r_i)}$ は、文書 r_i に対する $(\text{PSP}(r_i), 1 - \text{PSP}(r_i))$ という2次元ベクトルを示す。

評価実験では、サポートベクトルマシン分類器と One-vs-Rest 法 (Rifkin and Klautau 2004) の組合せ、Support Vector Regression (Smola and Scholkopf 1998)、および、それぞれに metric labeling法を適用したものを比較し、metric labeling法の有効性を示した。

Pangらは、評点の粒度に関する予備調査として、人間が区別できる評点差の違いを調査し、3値と4値の評点粒度からなるデータセットを作成し、評価実験に用いた。これは、4人の著者が記述した文書を各著者ごとにまとめた4セットの評価文書集合からなる。(Pang and Lee 2004)の知見を基に、あらかじめ事実文を取り除いている。Pangらの実験は、同一著者のデータにしか適用されていないため、複数の著者によって生成された評価文書群に対して評価文書評点分類を行うには、著者間の評点値がもつ意味を正規化する手法が別に必要になる。

(岡野原・辻井 2005)

岡野原ら (岡野原・辻井 2005) は、日本語の評価文書を対象にして、評価文書評点分類に取り組んだ。彼らの実験では、サポートベクトルマシン分類器と pair-wise 法 (Kresel 1999) の組合せと、Support Vector Regression が比較されている。

評価文書評点分類では、分類クラスの粒度が2値分類よりも細かいために、従来の評価表

現に加え，極性強調子の情報をうまく扱うことが重要である．岡野原らは，素性として，単語 uni-gram，bi-gram，tri-gram の3種類の素性を試した．結果として，単語 uni-gram と比較して，単語 bi-gram，tri-gram を用いた方が精度が高かった．「very good」のような評価表現と極性強調子の並びをうまく捉えたことが精度向上により影響を与えていると考えられる．

3.2.4 まとめ

評価文書分類を扱った論文を紹介した．評価文書分類の手法は，評価表現の比率に基づく手法と，教師あり機械学習に基づく手法に分けることができる．また，肯定か否定かを分類する評価文書分類の拡張として，より細かな分類粒度を扱う評価文書評点分類に関する研究を紹介した．

評価文書分類は，既存のトピックに基づく文書分類とは分類に有効な情報が必ずしも同じではない．(Pang et al. 2002) による実験結果によると，評価文書分類は，トピックに基づく文書分類よりも難しい課題であると考えられ，今後の進展が望まれる．

既に述べたように，これまでは，肯定極性か否定極性かのいずれかのみを扱う2値分類として扱われることが多かった．しかし，現実には，肯定極性か否定極性のいずれにも該当しない文書が存在する状況が多く，(Koppel and Schler 2005) のように，中立クラスを加えた3値の評価文書分類に対する考察についても今後の成果が期待される．中立クラスの扱いについては，6.1 節で再び取り上げる．

現在の評価文書分類では，多くの場合，ある評価対象についての評価情報が含まれている文書群が既に収集されているという前提の基で研究が進められる．これは，2 節で述べた分類のうち，「意見の収集，集約が目的となっているテキスト」が主な研究題材となっていることに一因がある．しかし，一般には，ある文書が与えられた時，その文書に記述されている評価対象が何であるかについても判定する必要がある．今後は，評価対象の分類を考慮した評価文書分類についても検討していく必要がある．

後の 3.4 節では，文書内に含まれる評価情報の要素に注目した諸研究を紹介する．ここでは，評価情報要素のひとつとして，評価対象の存在が意識される．

3.3 評価情報を含む文の抽出に関する諸研究

前節では，文書レベルで評価極性を判定する手法を紹介した．本節では，文書に含まれている文の評価極性を判定し，抽出する手法を紹介する．

評価分析を実施する際，前節で紹介した評価文書分類を行うことによって，評価情報，評価極性に関する文書全体の傾向を把握することができる．しかし，評価対象がどのように肯定的

あるいは否定的な評価を受けているかを知るには、評価文書の内容に目を通す必要がある。この時、各文書内に含まれている評価極性付きの文を抽出することができれば、評価対象がどのように評価を受けているかを知る上で見通しがよくなる。また、ひとつの文書に複数の評価が混じっているような場合には、文書レベルではなく、文レベルで評価極性を判定した方がよい。

評価文抽出の基本的な考え方は評価文書分類と同様であり、評価表現の比率に基づく手法や機械学習に基づく手法がそれぞれ提案されている。

Yu et al. (Yu and Hatzivassiloglou 2003) は、評価表現の比率に基づく手法によって評価文分類を実現した。分類手続きは 3.2.1 節で紹介した (Turney 2002) や (Taboada and Grieve 2004) の手法と基本的に同じであるが、Yu らは、分類クラスとして肯定、否定に中立を加えた 3 クラスを考えるため、分類手続きの 3 ステップ目に拡張を施した。幾つかの精度評価実験の中で評価表現の品詞に関する考察を行っており、実験の結果、品詞が形容詞、副詞、動詞となる評価表現を合わせて利用した場合が精度が高かったと報告している。一方、この組合せに品詞が名詞である評価表現を加えると精度が低下する結果を得たと報告している。

Gamon et al. (Gamon and Aue 2005) は、評価表現の比率に基づく手法と機械学習に基づく手法をそれぞれ提案している。評価表現の比率に基づく手法では、(Yu and Hatzivassiloglou 2003) の手法を拡張した。評価極性値を計算する際に、文内では「同じ極性の単語が共起しやすい」という仮定に加え、「逆極性の単語は共起しにくい」という仮定を陽に取り入れたり、一度求めた評価極性値に基づいて種表現を漸次的に増やす手法を提案している。また、機械学習に基づく手法では、Nigam らが提案した半教師あり学習の手法 (Nigam, McCallum, Thrun, and Mitchell 2000) によって、評価極性情報が付与されていない大量の生データの情報を評価文分類に利用した。

評価文分類で採用される分類手法は評価文書分類と共通するものが多いが、評価文分類に利用される素性情報も評価文書分類と重なる部分が多い。評価文書分類の場合と同様に、評価文分類でも単語、特に評価表現が分類の際の主要な情報となる。また、文内の語の系列情報 (笈島、嶋田、遠藤 2005) や構文情報 (工藤 松本 2003) を利用したアプローチも報告されている。評価文分類では、評価文書分類のように文書内の位置情報などは利用されない。その一方で、評価文分類に固有な情報として、例えば、文の文型パターン (村野 佐藤 2003; 峠 山本 2004) などが利用される。

また、評価文書分類の場合と同様に、評価文分類においても、評価対象が何であるかについては考慮されることがほとんどない。数少ない評価対象を考慮した評価文分類の研究として、(Hurst and Nigam 2004; Nigam and Hurst 2004) がある。Hurst と Nigam は、評価対象を含む文を抽出する過程と文の評価極性を判定する過程を独立に並行して行う。そして、「ある文に、ある対象についての記述と、評価極性を示す記述が含まれている場合、その評価極性は同一文内の

対象に関する極性である」という仮定に基いて、2つの過程の結果を併合することで対象に関する評価極性を判定した。

3.4 評価情報の要素組の抽出に関する諸研究

評価情報は「良い」や「悪い」などの評価表現の他に、評価者、評価対象などの幾つかの要素によって構成される。これらの評価情報の要素組のことを以下、評価組と呼ぶ。本節では、文書から評価組を抽出し、その評価極性を判定する研究を紹介する。評価組抽出の研究背景にある動機付けは、評価文抽出と同様であり、評価文書全体の評価ではなく、評価文書中の個々の評価情報に関する評価に注目する。

この課題は3.1節で紹介した評価表現辞書の構築に似ている。しかし、評価組の抽出では、評価表現が評価している評価対象や評価をしている評価者などを合わせて特定、抽出する点が評価表現辞書の構築と異なる。

例えば、下の例文(2a)からは、「美味しい」という評価表現と共に、その評価を受けているのは「りんご」であり、両者を合わせて、「りんごが肯定的な評価を受けている」ことを特定したい。また、例文(2b)であれば、評価を受けているのは「帽子」であり、評価をしているのは「次郎」であることを特定したい。

このように実際に評価表現が現れている文脈の情報を考慮することで、注目している評価情報の評価極性が変化する場合もある。例文(2c)と(2d)は同じ「眠気をさそう」という表現を含んでいるが、それぞれの対象(「ベッド」と「講義」)を考慮し、<ベッド, 眠気をさそう>は肯定極性、<講義, 眠気をさそう>は否定極性であることを判定することが求められる。

さらに、例文(2e)は、肯定極性をもつ評価表現「美味しい」を含んでいるが、文自体は質問を表しているのであって、りんごを評価しているわけではない。そのため、(2e)から得られる要素の組に対しては、評価ではないことを判定することが求められる。

- (2) a. りんごが美味しい。
b. 次郎は太郎からもらった帽子を気に入っている。
c. このベッドは眠気をさそう。
d. この講義は眠気をさそう。
e. このりんご、美味しかった？

評価組の抽出は、次のような副課題に分解して考えることができる。

- 要素抽出：評価表現、評価対象などを文書から特定する(3.4.1節)。
- 関係抽出：各要素を結びつける。例えば、ある評価表現がどの評価対象について評価しているかを特定する(3.4.2節)。
- 動的評価極性：関係抽出の結果得られた評価組の評価極性を判定する(3.4.3節)。

それぞれの副課題について順に紹介する。

3.4.1 要素抽出

要素抽出は、評価情報を構成する要素を抽出する課題である。現在のところ、評価情報の要素としては、評価表現、評価対象、評価対象の属性、評価者などが主に取り上げられている。

評価表現の抽出には 3.1 節で構築した辞書が利用される。また、多くの研究では評価対象はあらかじめ与えられていると想定する。そこで以下では、評価対象の属性と評価者の抽出に関する研究を紹介する。

3.4.1.1 評価対象の属性

評価対象の属性とは、評価対象の仕様（性質や特徴など）や評価対象の一部分などのことを指す。デジタルカメラについての属性の例を (3) に示す。下線部が属性表現である。

- (3) a. 値段 が少し高い。
- b. 満足のいく 画質 です。
- c. レンズ が明るい。

実際の属性抽出処理は、あらかじめ属性表現の集合を属性辞書として獲得しておき、評価情報獲得時には辞書照合によって属性抽出が実現される。そこで、属性辞書を（半）自動で構築することがここでの課題となる。

属性の中には、値段やサイズなどのように評価対象間で共通の属性も存在するが、一般には、評価対象ごとに属性は異なる。そのため、属性辞書は評価対象ごとに用意する。ある対象に関する属性辞書を構築する場合、その対象について記述された文書群を用意し、そこから属性辞書のエントリを見つけ出し、辞書登録をおこなう。

以下ではまず、ブートストラップに基づく小林ら (小林他 2005) の手法を紹介する。小林らの手法は、評価対象の属性と同時に評価表現も合わせて抽出する。

(小林他 2005)

小林ら (小林他 2005) は、文書に含まれる意見は <対象, 属性, 評価> の 3 つの要素からなると捉え、この 3 つ組要素のうち、属性表現と評価表現を効率よく収集する手法を提案した。次の (4) は 3 つ組の例である。(4a) が意見が含まれるテキストであり、(4b) が (4a) から得られる 3 つ組である。

- (4) a. 商品 A は、ボディはコンパクトですが、安定感は抜群ですね！
ただ足回りは固いので、人を乗せる時は気を使います。

- b. <商品_A ボディ, コンパクト>
 - <商品_A 安定感, 抜群>
 - <商品_A 足回り, 固い>

小林らの手法は、ブートストラップに基づいており、対象、属性、評価に関する共起パターンを介して、評価表現と属性表現を相互に獲得する。共起パターンの例を(5a)に示す。また、(5a)の共起パターンに照合する意見テキストを(5b)に示す。

(5) a. <対象>の<属性>は<評価>

b. 商品_Aの インテリア はきれいですね。

共起パターン中の<対象>、<属性>、<評価>のそれぞれのスロットは、新規表現スロット(下線で示す)と既知表現スロット(下線部以外)に分かれている。各パターンにはひとつの新規表現スロットが含まれる。上記の例の場合、「商品_A」に関して、「きれい」が評価表現であることが既知である状態で、「インテリア」が「商品_A」の新しい属性候補であるとして収集される。

小林らは8タイプの共起パターンを用意しているが、タイプによって、<属性>が新規表現スロットになる場合と<評価>が新規表現スロットになる場合を分けており、文書集合に対して、タイプの異なる共起パターンを繰り返し適用することで、属性表現と評価表現を相互に獲得している。

(5a)のように、小林らが作成した共起パターンは汎用的なパターンである。そこで、パターン照合の後、ノイズとなる属性候補、評価表現候補を排除するために幾つかのフィルタリングを行う。フィルタリングに利用される情報には、スロットに埋まる語句の品詞に関するものや、獲得された属性候補と評価表現候補の統計量(頻度、対数尤度比)などがある。

なお、小林らの手法は、収集できた評価表現における評価極性は決定しない。

その他にも、統計量に基づく属性抽出手法が幾つか提案されている。これらは、小林らのパターンによる候補獲得とフィルタリングという一連の処理のうち、パターン照合部に重点を置かず、フィルタリング部のみを行っていることに対応する。Yi et al. (Yi and Niblack 2005)は、限定名詞句に注目し、評価対象の文書に偏って出現する限定名詞句を対数尤度比に基づいて抽出した。Hu et al. (Hu and Liu 2004a, 2004b)は、評価文書に対して、相関ルールマイニングの手法を用いて、評価文書集合中に頻出し、かつ評価表現から近い位置に出現する名詞あるいは名詞句を属性表現として抽出した。Liu et al. (Liu, Hu, and Cheng 2005)は、(Hu and Liu 2004a, 2004b)と同様に相関ルールマイニングの手法を用いて属性を抽出した。他の研究では属性表現の抽出対象を名詞に制限することが多いが、Liuらは名詞以外の表現も抽出している。Morinaga et al. (Morinaga et al. 2002)は、本稿でいう属性に近い概念として、対象を特徴付ける語をマイニングする手法を提案した。

3.4.1.2 評価者 (Holder, Source)

評価者 (Holder, Source) とは, 評価対象を実際に評価している人物や組織のことである。まず, 次の文を例に取りながら, 評価情報の要素として評価者を考慮することの必要性について検討する。

(6) りんご 1 箱あたりの値段が上がった。

この文の評価極性を考える。りんごの消費者にとって, りんごの値上がりは好ましいとは言えない。つまり, 例文 (6) をりんごの消費者から見た場合, この例文の評価極性は否定極性であると考えることができる。一方, りんごの生産者にとっては, りんごの値上がりは好ましい可能性があり, 例文 (6) をりんごの生産者から見た場合は, この例文の評価極性は肯定極性となる可能性がある。この例から, まったく同じ表現でも誰の視点から見るかによって評価極性が異なることが確認できる。つまり, より厳密なテキスト評価分析を実施するには, 評価表現や評価対象に加えて, 評価をおこなう者, つまり, 評価者を特定する必要があるであろう。

人物名辞書や組織名辞書があらかじめ用意できる環境であれば, 評価者候補の特定は単純な辞書照合として容易に実現できる。評価者候補の特定には, ほかに, 既存の固有表現抽出 (named entity recognition; NER) (MUC6 1995; MUC7 1997; Sekine and Isahara 1999) に関する技術や意味役割同定 (semantic role labeling; SRL) (CoNLL-ShareTask 2004, 2005) に関する技術が利用できるだろう。

いずれの場合において, 特定された評価者候補が実際に評価組の要素となるか否かは続く関係抽出処理に任せることになる。

テキスト評価分析に関する研究全体の中でも, 評価者の視点を考慮した研究は現時点ではそれほど多くはない (Kim and Hovy 2004; Bethard, Yu, Thornton, Hatzivassiloglou, and Jurafsky 2004; 中山, 江口, 神門 2005)。この理由のひとつとしては, 対象データが評価文書に偏っているという現状がある。評価文書の場合は, ほぼ,

評価者 = 書き手

の関係が成り立つ。このことから, 現時点では評価者を特定する処理が省かれていると言える。今後, 評価文書ではない, 一般の Web 文書や新聞記事などに含まれる意見をテキスト評価分析で扱う場合, あるいは, 評価文書を対象にしてより精緻なテキスト評価分析を実現する場合には, 評価者の視点をより積極的に考慮していく必要がある。

3.4.2 関係抽出

関係抽出は, 評価情報の各要素を関係づける課題である。先に述べたように, 評価組抽出は, 要素抽出, 関係抽出, 動的極性判定の 3 つの副課題に分解できるが, 現在のところ, 要素抽出,

動的極性判定に比べて、関係抽出に焦点を当てた研究は乏しい。評価組抽出を扱う研究の多くは、要素間の近接情報や構文情報などに基づいた素朴な手法によって関係抽出を実現しているが、当然のことながら、十分な抽出精度が得られているわけではない。

関係抽出に積極的に焦点を当てた研究としては、(小林, 飯田, 乾, 松本 2005; 飯田他 2005) がある。事前調査から、意見を含む日本語文書では、<属性, 評価>間に係り受け関係が成立しない事例が少なからず存在することを指摘し、機械学習に基づく<属性, 評価>関係を特定する手法を提案している。彼らのグループでは、<属性, 評価>対を特定する課題が照応解析における<先行詞, 照応詞>対を特定する課題に類似していることに着目し、照応解析モデルを関係抽出に応用した。

近年では、固有表現間の関係を抽出する課題 (National Institute of Standards and Technology 2000) に対して、機械学習に基づく手法が盛んに研究されており (Zelenko, Aone, and Richardella 2003; Culotta and Sorensen 2004)、そこで得られた知見が本課題に利用できる可能性がある。

3.4.3 動的極性判定

関係抽出の結果得られた評価組の評価極性を判定する研究を紹介する。これまでに見たように、要素抽出と関係抽出は、評価組抽出の副課題としてその他の副課題とは独立に研究されるケースが多い。一方、動的極性判定を実施するには、要素抽出と関係抽出の出力結果が必要である。要素抽出と関係抽出の出力結果を手手で用意することによって、それらの結果が与えられた状態で動的極性判定のみを課題とすることもできるが、現在の動的極性判定に関する研究では、要素抽出と関係抽出も同時に合わせて扱われることが多い。

以下では、パターンに基づく Nasukawa et al. (Nasukawa and Yi 2003) の手法と、機械学習に基づく鈴木ら (鈴木, 高村, 奥村 2004) の手法を紹介する。

(Nasukawa and Yi 2003)

Nasukawa et al. (Nasukawa and Yi 2003) は、評価情報の要素として、<対象, 評価>の2つ組を考え、構文、意味関係パターン付きの評価表現辞書を用いて2つ組に対する動的評価極性を行った。

彼らは、辞書知識として、3.1節で紹介した評価極性付きの単語のエントリに加え、構文・意味関係を考慮したパターン付きのエントリを用意した。単語「admire」と「provide」のエントリ例を(7)に示す。

- (7) a. good VB admire obj
- b. transfer VB provide obj sub

(7a)は、「admire」の目的語の位置ある要素が肯定極性になることを表しており、(7b)は、

「provide」の目的語の位置ある要素の極性が判明していれば、主語の位置にある要素は目的語の位置ある要素の極性と等しくなる（評価極性が transfer する）ことを表している。

このような、語句間における評価極性の同一（相違）関係が記述された辞書知識を組み合わせることで、評価表現と、それと同一文内の主語や目的語の位置に表れる評価対象の組を同定し、それらの評価極性を判定する。

(7) に示したような構文・意味関係を考慮したパタン付きのエントリは、評価表現辞書を素直に文脈情報を取り込むように拡張したものと考えられる。

Kanayama et al. (Kanayama, Nasukawa, and Watanabe 2004) は、要素抽出から評価極性判定までを、文書から評価組への翻訳であると捉えた。そして、transfer-based な機械翻訳の機構を援用し、既存の翻訳パタンを、上記 (7) と類似した極性判定のための構文・意味関係を考慮したパタンに置き換えることによって、低コストで評価情報の要素組を抽出する枠組みを提示した。

(鈴木他 2004)

鈴木ら (鈴木他 2004) は、評価情報の要素として、〈対象, 属性, 評価〉の3つ組を考え、評価情報とその周辺情報のブートストラップによって、3つ組の動的極性判定を行った。評価極性としては、肯定 / 否定 / 中立の3値を扱っている。

先の (Nasukawa and Yi 2003) の手法では、要素抽出、関係抽出と動的極性判定が並行して進行していたが、鈴木らの手法では、〈対象, 属性, 評価〉の3つ組抽出と、3つ組の極性判定が逐次的に行われる。まず、評価表現辞書を用いて文書内の評価表現を同定する。そして、評価表現と係り受け関係にある名詞を評価対象、属性として、〈対象, 属性, 評価〉の3つ組を同定する。ただし、経験から評価対象は具体名詞、属性は抽象名詞という制約を課す。具体名詞、抽象名詞の判定は NTT 日本語語彙大系 (Ikehara, Miyazaki, Shirai, Yokoo, Nakaiwa, Ogura, Ooyama, and Hayashi 1997) の情報に基づく。

次に、得られた3つ組を評価極性分類器に入力することで、評価極性を判定する。彼らの分類器は半教師あり機械学習によって構築される。彼らは、まず、ベースとなる分類器として、ナイーブベイズ分類器を考え、〈対象, 属性, 評価〉組の評価極性を分類することを考える。そして、さらに教師なしデータの情報を取り込むために、EM アルゴリズム (Dempster, Laird, and Rubin 1977) をナイーブベイズ分類器と併用する手法 (Nigam et al. 2000) を用いた。また、教師なしデータを取り入れる別の手法として、同一著者による (鈴木 2005) では、先の手法から獲得された確率モデルからフィッシャーカーネル (Jaakkola and Haussler 1998) を作成し、そのカーネル関数をサポートベクトルマシン分類器で利用する手法を検討している。

3.4.4 語の組合せと評価極性

3.1 節で述べた評価表現辞書の構築では、「良い - 肯定」や「悪い - 否定」のように、単独で評価極性が特定できる単語を扱うことが多い³。しかし、単語の中には、それ単独では評価極性をもたないが、幾つかの単語が組み合わさることによって初めて、その組合せに評価極性が生じる場合がある (Baron and Hirst 2004)。ここでは、この現象を極性発現と呼ぶ。例えば、英語の “par for the course” という句は、個々の単語は否定極性を持たないが句全体としては否定極性をもつ (Channell 2000)。

極性発現に着目した研究には、Baron et al. (Baron and Hirst 2004) と高村ら (高村, 乾, 奥村 2005b) の研究がある。Baron らは、Xtract (Smadja 1994) を用いてコーパスから共起語を抽出し、共起語が出現していた文脈の評価極性に従って共起語の評価極性を決定した。この考え方は、評価表現辞書の構築で用いられた単語の評価極性の判定手法 (3.1.2 節) に非常に近い。一方で、高村らは、極性発現が生じる語句の内部構造を直接モデリングし、「ノートパソコンが軽い - 肯定」のように語句の評価極性が判定できる確率モデルを提案した。高村らの提案モデルは極性発現と同時に極性変化子にも対応しており、「リスクが低い - 肯定」のような例も適切に扱える。長江らは極性発現を考慮した評価表現辞書を作成している (長江, 望月, 白井, 島津 2002)。

4 テキスト評価分析の応用研究

4.1 評価分析システム

これまでに、既に幾つかの評価分析システムが構築されている (Li and Yamanishi 2001; Morinaga et al. 2002; Dini and Mazzini 2002; Dave, Lawrence, and Pennock 2003; Sano 2004; 佐野 2004; 立石, 福島, 小林, 高橋, 藤田, 乾, 松本 2004; 奥村, 南野, 藤木, 鈴木 2004; 藤村, 豊田, 喜連川 2004; Yi and Niblack 2005; Liu et al. 2005)。評価分析システムの多くは、基本機能として、文書集合から文書、文、あるいは、語句などの単位について、肯定 / 否定の評価極性を判定する。この機能を実現するために、前節で紹介した各要素技術が評価分析システムに組み込まれる。

評価分析システムの利用者は、大きく次の 2 つに分けられる。

- 対象の提供者である企業やサイト運営者、
- 対象を利用する立場の一般の人々。

前者は、マーケティング等で、対象の利用者としての個人の評価を、今後の企業活動に活かすことが典型的な用途となる。後者は、他の個人の評価を、自分の意志決定 (たとえば、何かを購

³ (那須川・金山 2004) のように、複数の単語から構成される評価表現を考慮した事例もあるが、現状では複数語からなる評価表現を扱う研究は少ない。

入しようとして、いくつかの候補の中からの選択が必要な場合)の参考にしようというのが典型的な用途となる。

言うまでもなく、大量の抽出された評価情報は、そのまま提示されるのではなく、類似するものはまとめられ、また、いくつかの観点で分類され、ユーザにとって負担のない情報量で、提示されることが望ましい。そのため、評価分析システムでは、評価情報を抽出するという基本機能に加えて、分析データの集約機能や、分析結果の可視化機能を備えている場合が多い。可視化にはグラフ表示が採用されることが多く、例えば、(Yi and Niblack 2005)では棒グラフ、(立石他 2004)ではレーダーチャートが表示できる⁴。この可視化機能によって、単一の対象、あるいは複数の対象間で、属性ごとにその対象の評価値を比較するといった作業が容易となる。

分析結果の評価情報を可視化するのではなく、テキストとして提示する研究も存在する。Beineke et al. (Beineke et al. 2004), Roman et al. (Roman and Piwek 2004)はいずれも、抽出した評価情報集合をテキストの形で集約して提示する試みを行なっている。文書要約(奥村 難波 2005)の観点からみた場合、これらの研究は、評価情報が含まれている文の重要度を高く定めて、複数文書を対象とした重要文抽出を行って要約を作成していることに相当する。この観点から見れば、評価文抽出は評価文書のテキストを要約する手法の中の特殊なケースとして捉えることもできる。

複数文書要約では、異なるソース(情報源)(たとえば、新聞社)からのテキストで、内容が重複する場合、重複する内容は重要であるとして、重複する内容を選択する手法が見られるが、評価文書集合の要約でも、同様に、評価情報が重複することは、共通の評価が複数存在することを意味しており、重要な情報としてその頻度情報が出力のひとつとして利用される。また、一般の複数文書要約では重複部分を同定することは困難な課題であるが、評価文書の場合は、基本的には、評価対象、評価対象の属性、評価極性の3つの要素のみから一致(重複)判定が可能である。

4.2 その他の応用領域

前節で紹介したテキスト評価分析のための各要素技術は、評価分析システムの他にも、計算機を介したコミュニケーション(computer mediated communication; CMC)(Boucouvalas 2002; Liu, Lieberman, and Selker 2003)、質疑応答システム(multi-perspective question-answering; MPQA)(Cardie, Wiebe, Wilson, and Litman 2003; Stoyanov, Cardie, Litman, and Wiebe 2004)、株価予測(Koppel and Shtrimberg 2004; Das and Chen 2001)など、幅広い応用領域でその有効性が示されている。

⁴ 対象についての属性が非常に多い場合、可視化手法によってはすべての属性をまとめて閲覧することが困難となる。その場合には、抽出時とは別に表示時において、属性の取捨選択や階層化などの処理が要求される。

5 テキスト評価分析に関連するその他の話題

本節では、テキスト評価分析に関連するその他の話題を紹介する。

5.1 主観性に関する諸研究

多くの文書は、ひとつの文書の中で、客観的な事実と主観的な意見が混じり合って記述される。本稿でとりあげた評価は、意見の中の下位分類に位置すると考えられ、本稿で述べた評価情報を扱う諸研究と意見情報を扱う諸研究は非常に関連が深いと考えられる。

これまで、Wiebe らの研究グループが中心となって、主観的な意見、主観性 (subjectivity) に焦点を当てた研究が進められている (Wiebe, Wilson, Bruce, Bell, and Martin 2004)。主観性に関する研究では、主に、文書から主観的な意見が記述された文 (以下、主観文) を判定し、抽出する課題が扱われる。具体的な事例研究としては、Wiebe et al. (Wiebe, Bruce, and O'hara 1999), Riloff et al. (Riloff, Wiebe, and Wilson 2003), Pang et al. (Pang and Lee 2004) などがある。

Hatzivassiloglou et al. (Hatzivassiloglou and Wiebe 2000) は、一部の形容詞の情報が主観文判定に有効であると報告しており、Weibe (Wiebe 2000) や Vegnaduzzo (Vegnaduzzo 2004) によって、文の主観性に強い影響を与える形容詞を獲得する手法が提案されている。また、形容詞以外の情報として、文の主観性に強い影響を与える名詞 (Riloff and Wiebe 2003) や単語の n-gram (Wiebe et al. 2001) を自動獲得する手法も提案されている。

主観文抽出の具体的な応用先としては、情報抽出や文書要約などが挙げられる。情報抽出では、実世界で実際に起きた事実の客観的な記述を抽出することに焦点が当たっており、主観文抽出を前処理として適用することで、抽出範囲を適切に絞り込むことができる。また、近年では、個人の意見が Web 掲示板や Weblog などを通じて発信されている。そして、この個人の意見を要約することが求められている。この時、意見文抽出を実施することによって、要約対象となる意見を自動的に抽出することができる。

5.2 その他の題材

書き言葉で構成された文や文書以外にも、チャットや実対話を題材とした研究もある。Maeireizo et al. (Maeireizo, Litman, and Hwa 2004), Chambers et al. (Chambers, Tetreault, and Allen 2004) は、実対話データを扱っている。また、Wu et al. (Wu, Khan, Fisher, Shuler, and Pottenger 2002), Holzman et al. (Holzman and Pottenger 2003) は、チャット対話中の発話に焦点を当てた。これらの研究では、本稿で取り上げた評価情報だけでなく、広く感情情報を意識する

傾向にあり，実対話やチャット内の発話の感情を推定することが中心的な課題となる．この際，言語的な特徴だけでなく，音素や韻律などの音声的な特徴も積極的に利用される．

6 課題

本節では，テキスト評価分析において，現在までのところあまり議論がされておらず，今後，進展が望まれる話題を整理する．

6.1 「中立」の取り扱い

既に述べたように，これまでの評価分析の要素技術研究では，評価極性として肯定極性と否定極性のみを考慮し，2値分類問題として定式化されることが多い．しかし現実には，肯定極性と否定極性のいずれにも該当しない文や文書が存在する状況も多くある．そのため，今後は，(Koppel and Schler 2005) や (Yu and Hatzivassiloglou 2003)，あるいは(鈴木他 2004)のように，中立クラスを加えた3値での評価分析に対する考察が求められる．

ただし，中立クラスを扱う際には，次の点に注意しなければならない．すなわち，現在でも中立クラスを考慮した研究は存在するが，研究者間でその意味するところは必ずしも一致しているわけではないという点である．ある事例が肯定極性でも否定極性でもない状況には幾つかの可能性がある．しかし現状では，それらすべてがまとめて中立と呼ばれている．

ある事例が評価を含むか，また，評価極性があるか，という点から整理した場合，各事例は図2の木構造のいずれかの葉ノードに割り振られる．本稿で紹介した各論文が主に扱っていた肯定極性と否定極性は，右隅の四角で囲った部分に位置する．それ以外の点線で囲まれた3つの葉ノードは，肯定極性か否定極性のいずれかの極性をもつわけではないという共通の特徴があり，現状では，これらのすべてあるいはいずれかのノードに該当する事例が「中立」と見なされている．しかしながら，図の階層分類からも明らかなように，点線で囲まれた3つの葉ノードはそれぞれ，評価を含まない，評価を含むが肯定極性でも否定極性でもない，肯定極性と否定極性の両方を含む，と言ったそれぞれに異なった特徴がある．今後，評価分析に関する諸研究を進展させていくには，「中立」という概念を再整理する必要があるだろう．

6.2 評価の分類軸

現在，評価分析の分類軸としては，肯定極性と否定極性からなる評価極性というひとつの軸を扱うのが主流であるが，今後は，目的に応じて分類軸は細分類化していく必要がある．

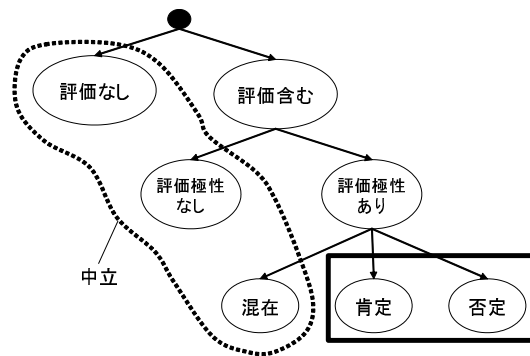


図 2: 「中立」の曖昧性

評価極性は、感情や情緒 (乾, 徳久, 徳久, 岡田 2000; 目良, 市村, 相澤, 山下 2002) と関連が深く、今後の展開として、これらの領域との融合研究も興味深い。近年では、自由記述アンケートから書き手の意図を抽出する研究 (大塚 2004) や、意図の中でも特に要望 (金山 那須川 2005) や賛否 (Galley, McKeown, Hirschberg, and Shriberg 2004) に注目した研究も行われている。評価分析の一つの方向性として、これらの研究の今後の動向にも注意を向けたい。

新たな分類軸の方向性について検討することも重要であるが、現状の評価極性自体にも目を向ける必要がある。現状の評価極性は、研究者によって意味する所が異なる部分があり、テキスト評価分析に利用される評価の軸を再考察する必要がある。例えば、次の例文 (8) を考えてみよう。本稿で紹介した諸研究が扱う評価極性を見る限り、(8a) が肯定極性をもつ評価情報を含んでいるとすることに異論をもつ研究者は少ないと考えられる。しかし、(8a) と同様に肯定極性をもつと考えられる (8b) も同じように評価情報を含んでいるとするか否かは研究者によって立場を異にする。(8b) はある人物の嗜好の表出であって評価は表されていないとする立場も存在する。

- (8) a. このりんごは味がよい。
- b. りんごが好きです。

Martin の提案している Appraisal system (Martin 2000, 2003) は、評価や意見、態度などの主観的な側面に関する言語能力を解明、説明する枠組みである。Taboada et al. (Taboada and Grieve 2004) は、実際に Appraisal system での考え方に従って、複数の異なる評価の分類軸を扱っており、評価極性の再整理という点から見て、非常に興味深い研究である。

6.3 明示的に表されない評価

3節で述べた評価文書分類，評価文抽出，評価組抽出の各技術は，現在のところ，評価表現辞書のエントリとなるような明示的な評価表現に頼っている部分が多い．このため，現状では，評価表現辞書に登録されている語句がまったく出現しない，つまり，明示的に評価が表されない文書や文を取り扱うことが困難である．今後は，明示的に表されない評価を取り扱う方法についての技術開発も進める必要があると考えられる．

評価が明示的か明示的でないかの境界を明確にすることは難しいが，例えば，3.4.4節で紹介した極性発現はどちらかと言えば明示的でない評価を扱っている一例と言える．

また，評価は意見の中の下位分類として位置付けられることが多いが，明示的に表されない評価は，意見の記述というよりは事実の記述となっていることがしばしばある（例えば，(Nigam and Hurst 2004)は，このことを説明するために「opinion」の対極として「evaluative factual」という用語を導入している）．例えば，次の例文(9)は，明示的に表されない評価を含む文の例である．今後の評価分析に関する技術開発が進む方向次第では，評価分析が既存の意見分析の域には収まらず，独自の新たな研究領域を開拓しつつ進展していく可能性もある．

- (9) a. 買ってすぐに電源が入らなくなった。
- b. 空港でパスポートがないことに気が付いた。
- c. おもわず息子の頭を撫でていた。

6.4 基礎言語解析技術

テキスト評価分析で扱うテキストデータ（評価文書）には，非専門家によって記述されたWeb上の掲示板への書き込みやWeblogなどが含まれる．これらのテキストデータは，従来から言語処理の対象とされてきた新聞記事ほど形式的に記述されていないため，従来からある形態素解析等の基礎言語解析器をそのまま利用するだけでは，高い解析精度が得られない．また，評価文書には，表記の多様性や，局所的なコミュニティー特有の言い回し，略語などの現象が多く見受けられる．今後，これらの諸現象に柔軟に対応できる，より頑健な基礎言語解析技術の開発が望まれる．

6.5 評価文書の収集

現在のテキスト評価分析に関する要素技術の諸研究では，多くの場合，ある評価対象についての評価情報が含まれている文書群が既に収集されているという前提のもとで研究がされている．しかし，明らかに，注目したい評価対象のすべてについて，この前提を置くことは適切で

はなく、現実には、評価対象についての評価情報が含まれている文書群を獲得する方法、評価文書の収集方法を確立しなければならない。特に、2節で示したテキスト評価分析の題材となるテキストデータの分類のうち、潜在的に意見を含むテキスト（Web 掲示板，Weblog，チャット）を処理対象とする場合には、この問題が顕在化するだろう。

7 おわりに

本論文では、近年盛んに研究活動が行われているテキスト評価分析に関する研究について、基盤となる研究から最近の研究動向までをまとめた。紹介した一連の研究領域は、いずれも成熟しているわけではなく、現在、急激に進展している状況にある。その中において、本論文がテキスト評価分析に関する現状あるいは今後の方向性を見極めるのに役立てれば幸いである。

謝辞

本論文は筆者を含む有志による集い「Affect Analysis 勉強会」の活動から生まれた。勉強会に参加し、議論に加わって頂いたすべての方に感謝する。

参考文献

- Bai, X., Padmanand, R., and Airoidi, E. (2004). “Sentiment Extraction from Unstructured Text using Tabu Search-Enhanced Markov Blanket.” In *Proceedings of the International Workshop on Mining for and from the Semantic Web (MSW-2004)*.
- Baron, F. and Hirst, G. (2004). “Collocations as Cues to Semantic Orientation.” In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Beineke, P., Hastie, T., and Vaithyanathan, S. (2004). “The Sentimental Factor: Improving Review Classification via Human-Provided Information.” In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004)*.
- Berger, A. L., Pietra, V. J. D., and Pietra, S. A. D. (1996). “A maximum entropy approach to natural language processing.” *Computational Linguistics*, **22** (1), 39–71.
- Bethard, S., Yu, H., Thornton, A., Hatzivassiloglou, V., and Jurafsky, D. (2004). “Automatic Extraction of Opinion Propositions and their Holders.” In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Boucouvalas, A. C. (2002). “Real Time Text-to-Emotion Engine for Expressive Internet Communications.” In *Proceedings of International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP-2002)*.

- Cardie, C., Wiebe, J., Wilson, T., and Litman, D. J. (2003). "Combining Low-Level and Summary Representations of Opinions for Multi-Perspective Question Answering." In *Proceedings of the New Directions in Question Answering*, pp. 20–27.
- Chambers, N., Tetreault, J., and Allen, J. (2004). "Approaches for Automatically Tagging Affect." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Chandler, D. (1987). *Introduction to Modern Statistical Mechanics*. Oxford University Press.
- Channell, J. (2000). *Corpus-based Analysis of Evaluative Lexis*, chap. 3 in EVALUATION IN TEXT: Authorial Stance and the Construction of Discourse, Edited by Susan Hunston, University of Birmingham, and Geoff Thompson, pp. 38–55. Oxford University Press.
- Church, K. W. and Hanks, P. (1989). "Word association norms, mutual information, and Lexicography." In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp. 76–83. Association for Computational Linguistics.
- CoNLL-ShardTask (2004). "The 9th. Conference on Computational Natural Language Learning. Shared Task: Semantic Role Labeling."
- CoNLL-ShardTask (2005). "The 10th. Conference on Computational Natural Language Learning. Shared Task: Semantic Role Labeling."
- Culotta, A. and Sorensen, J. (2004). "Dependency Tree Kernels for Relation Extraction." In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL2004)*.
- Das, S. R. and Chen, M. Y. (2001). "Yahoo! for Amazon: Opinion Extraction from Small Talk on the Web." In *Proceedings of the 8th Asia Pacific Finance Association Annual Conference*.
- Dave, K., Lawrence, S., and Pennock, D. M. (2003). "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews." In *Proceedings of the 12th International World Wide Web Conference (WWW-2003)*.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society Series B*, **39** (1), 1–38.
- Dini, L. and Mazzini, G. (2002). *Opinion classification through information extraction*, pp. 299–310. in A. Zanasi, C. A. Brebbia, N. F. F. Ebecken and P. Melli (eds), *Data Mining III*, WIT Press.
- Dunning, T. (1993). "Accurate methods for the statistics of surprise and coincidence." *Computational Linguistics*, **19**, 61–74.
- Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. The MIT Press.
- Galley, M., McKeown, K., Hirschberg, J., and Shriberg, E. (2004). "Identifying Agreement and Disagreement in Conversational Speech: Use of Bayesian Networks to Model Pragmatic Dependencies." In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004)*.

- Gamon, M. (2004). "Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis." In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-2004)*.
- Gamon, M. and Aue, A. (2005). "Automatic Identification of Sentiment Vocabulary: Exploiting Low Association with Known Sentiment Terms." In *Proceedings of the ACL Workshop on Feature Engineering for Machine Learning in Natural Language Processing*.
- Hatzivassiloglou, V. and McKeown, K. R. (1997). "Predicting the Semantic Orientation of Adjectives." In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL-1997)*.
- Hatzivassiloglou, V. and Wiebe, J. M. (2000). "Effect of Adjective Orientation and Gradability on Sentence Subjectivity." In *Proceedings of the 18th International Conference on Computational Linguistics (COLING-2000)*, pp. 299–305.
- Holzman, L. E. and Pottenger, W. M. (2003). "Classification of Emotions in Internet Chat: An Application of Machine Learning Using Speech Phonemes." Tech. rep., Lehigh univ (LU-CSE-03-002).
- Hu, M. and Liu, B. (2004a). "Mining and Summarizing Customer Reviews." In *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining (KDD-2004)*, pp. 168–177.
- Hu, M. and Liu, B. (2004b). "Mining Opinion Features in Customer Reviews." In *Proceedings of 19th National Conference on Artificial Intelligence (AAAI-2004)*.
- Hurst, M. and Nigam, K. (2004). "Retrieving Topical Sentiments from Online Document Collections." In *Proceedings of the 11th Document Recognition and Retrieval*.
- Ikehara, S., Miyazaki, M., Shirai, S., Yokoo, A., Nakaiwa, H., Ogura, K., Ooyama, Y., and Hayashi, Y. (1997). *Goi-Taikai - A Japanese Lexicon*. Iwanami Shoten.
- Inoue, J. and Carlucci, D. M. (2001). "Image restoration using the q-ising spin glass." *Physical Review*, **64** (036121-1-036121-18).
- Jaakkola, T. and Haussler, D. (1998). "Exploiting generativ models in discriminative classifiers." In *Proc. of the Advances in Neural Information Processing System 2*, pp. 487–493.
- Kamps, J., Marx, M., Mokken, R. J., and de Rijke, M. (2004). "Using WordNet to Measure Semantic Orientations of Adjectives." In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC-2004)*.
- Kanayama, H., Nasukawa, T., and Watanabe, H. (2004). "Deeper Sentiment Analysis Using Machine Translation Technology." In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-2004)*.
- Kennedy, A. and Inkpen, D. (2005). "Sentiment Classification of Movie and Product Reviews using

- Contextual Valence Shifters.” In *Workshop on the Analysis of Informal and Formal Information Exchange during Negotiations (FINEXIN-2005)*.
- Kim, S.-M. and Hovy, E. (2004). “Determining the Sentiment of Opinions.” In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-2004)*.
- Kleinberg, J. and Tardos, E. (1999). “Approximation Algorithms for Classification Problems with Pairwise Relationships: Metric Labeling and Markov Random Fields.” In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*.
- Koppel, M. and Schler, J. (2005). “The Importance of Neutral Examples for Learning Sentiment.” In *Workshop on the Analysis of Informal and Formal Information Exchange during Negotiations (FINEXIN-2005)*.
- Koppel, M. and Shtrimberg, I. (2004). “Good News or Bad News? Let the Market Decide.” In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Kresel, U. (1999). “Pairwise classification and support vector machines.” *Advances in kernel methods: support vector learning*, 255–268.
- Landauer, T. K. and Dumais, S. T. (1997). “A solution to Plato’s problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge.” *Psychological Review*, **104**, 211–240.
- Li, H. and Yamanishi, K. (2001). “Mining from open answers in questionnaire data.” In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 443–449.
- Liu, B., Hu, M., and Cheng, J. (2005). “Opinion Observer: Analyzing and Comparing Opinions on the Web.” In *Proceedings of the 14th International World Wide Web Conference (WWW-2005)*.
- Liu, H., Lieberman, H., and Selker, T. (2003). “A Model of Textual Affect Sensing using Real-World Knowledge.” In *Proceedings of the 2003 International Conference on Intelligent User Interfaces (IUI-2003)*.
- Mackay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.
- Maireizo, B., Litman, D., and Hwa, R. (2004). “Co-training for Predicting Emotions with Spoken Dialogue Data.” In *In ACL-04. Companion Volume to the Proceedings of the Conference. Proceedings of the Student Research Workshop, Interactive Posters/Demonstrations and Tutorial Abstracts*, pp. 203–206.
- Martin, J. (2000). *Beyond Exchange: Appraisal systems in English*, pp. 142–175. In Hunston, S. and Thompson, G. eds., *Evaluation in Text* (Oxford University).
- Martin, J. R. (2003). “Introduction, special issue on Appraisal.” *Text*, **23** (2), 171–181.

- Matsumoto, S., Takamura, H., and Okumura, M. (2005). "Sentiment Classification using Word Sub-Sequences and Dependency Sub-Trees." In *Proceedings of the 9th Pacific-Asia International Conference on Knowledge Discovery and Data Mining (PAKDD-2005)*.
- Mitchell, T. (1997). *Machine Learning*. McGraw-Hill.
- Morinaga, S., Yamanishi, K., Tateishi, K., and Fukushima, T. (2002). "Mining Product Reputations on the Web." In *Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2002)*.
- MUC6 (1995). "The 6th. Message Understanding Conference."
- MUC7 (1997). "The 7th. Message Understanding Conference."
- Mullen, T. and Collier, N. (2004). "Sentiment analysis using support vector machines with diverse information sources." In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004)*.
- Nasukawa, T. and Yi, J. (2003). "Sentiment Analysis: Capturing Favorability Using Natural Language Processing." In *Proceedings of the 2nd International Conference on Knowledge Capture (K-CAP 2003)*.
- National Institute of Standards and Technology (2000). "Automatic Content Extraction." <http://www.nist.gov/speech/tests/ace/index.htm>.
- Nigam, K. and Hurst, M. (2004). "Towards a Robust Metric of Opinion." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Nigam, K., McCallum, A., Thrun, S., and Mitchell, T. (2000). "Text Classification from Labeled and Unlabeled Documents using EM." *Machine Learning*, **39** (2/3), 103–134.
- Pang, B. and Lee, L. (2004). "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts." In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004)*.
- Pang, B. and Lee, L. (2005). "Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales." In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL-2005)*.
- Pang, B., Lee, L., and Vaithyanathan, S. (2002). "Thumbs up? Sentiment Classification using Machine Learning Techniques." In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2002)*, pp. 76–86.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.
- Polanyi, L. and Zaenen, A. (2004). "Contextual Valence Shifters." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.

- Qu, Y., Shanahan, J., and Wiebe, J. (2004). "Exploring Attitude and Affect in Text: Theories and Applications." Tech. rep. SS-04-07, Papers from 2004 AAAI Spring Symposium.
- Rifkin, R. and Klautau, A. (2004). "In Defense of One-Vs-All Classification." *Journal of Machine Learning Research*, **5**, 101–141.
- Riloff, E. and Wiebe, J. (2003). "Learning Extraction Patterns for Subjective Expressions." Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-2003).
- Riloff, E., Wiebe, J., and Wilson, T. (2003). "Learning Subjective Nouns using Extraction Pattern Bootstrapping." In *Proceedings of the 7th Conference on Computational Natural Language Learning (CoNLL-2003)*, pp. 25–32.
- Roman, N. T. and Piwek, P. (2004). "Politeness and Summarization: an Exploratory Study." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Salvetti, F., Lewis, S., and Reichenbach, C. (2004). "Impact of Lexical Filtering on Overall Opinion Polarity Identification." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Sano, M. (2004). "An Affect-Based Text Mining System for Qualitative Analysis of Japanese Free Text." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Sebastiani, F. (2002). "Machine learning in automated text categorization." *ACM Computing Surveys*, **34** (1), 1–47.
- Seerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990). "Indexing by latent semantic analysis." *Journal of the American Society for Information Science*, **41** (6), 391–407.
- Sekine, S. and Isahara, H. (1999). "IREX project overview." In *Proceedings of the IREX Workshop*.
- Smadja, F. Z. (1994). "Retrieving Collocations from Text: Xtract." *Computational Linguistics*, **19** (1), 143–177.
- Smola, A. and Scholkopf, B. (1998). "A tutorial on support vector regression."
- Stone, P. J., Dunphy, D. C., Smith, M. S., and Ogilvie, D. M. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. MIT Press, Cambridge.
- Stoyanov, V., Cardie, C., Litman, D., and Wiebe, J. (2004). "Evaluating an Opinion Annotation Scheme Using a New Multi-Perspective Question and Answer Corpus." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Strapparava, C. and Valitutti, A. (2004). "WordNet-Affect: an Affective Extension of Word Net." In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC-*

2004).

- Taboada, M. and Grieve, J. (2004). "Analyzing Appraisal Automatically." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Takamura, H., Inui, T., and Okumura, M. (2005). "Extracting Semantic Orientation of Words using Spin Model." In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL-2005)*.
- Turney, P. and Littman, M. L. (2003). "Measuring Praise and Criticism: Inference of Semantic Orientation from Association." *ACM Transactions on Information Systems (TOIS)*, **21** (4).
- Turney, P. D. (2002). "Thumbs up? thumbs down? Semantic Orientation Applied to Unsupervised Classification of Reviews." In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-2002)*, pp. 417–424.
- Turney, P. D. and Littman, M. L. (2002). "Unsupervised Learning of Semantic Orientation from a Hundred-Billion-Word Corpus." Tech. rep., Technical Report NRC Technical Report ERB-1094, Institute for Information Technology, National Research Council Canada.
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer.
- Vegnaduzzo, S. (2004). "Acquisition of Subjective Adjectives with Limited Resources." In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- Wiebe, J., Wilson, T., Bruce, R., Bell, M., and Martin, M. (2004). "Learning subjective language." *Computational Linguistics*, **30** (3).
- Wiebe, J., Wilson, T., and Bell, M. (2001). "Identifying Collocations for Recognizing Opinions." In *Proceedings of the ACL/EACL Workshop on Collocation*.
- Wiebe, J. M. (2000). "Learning Subjectives Adjectives from Corpora." In *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000)*.
- Wiebe, J. M., Bruce, R. F., and O'hara, T. P. (1999). "Development and Use of a Gold-Standard Data Set for Subjectivity Classifications." In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL-1999)*, pp. 246–253.
- Wolpert, D. H. (1992). "Stacked Generalization." *Neural Networks*, **5**, 241–259.
- Wu, T., Khan, F. M., Fisher, T. A., Shuler, L. A., and Pottenger, W. M. (2002). "Posting Act Tagging Using Transformation-Based Learning." In *Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM-2002)*.
- Yi, J. and Niblack, W. (2005). "Sentiment Mining in WebFountain." In *Proceedings of the 21st International Conference on Data Engineering (ICDE-2005)*.
- Yu, H. and Hatzivassiloglou, V. (2003). "Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences." In *Proceedings of the Con-*

ference on Empirical Methods in Natural Language Processing (EMNLP-2003).

- Zelenko, D., Aone, C., and Richardella, A. (2003). "Kernel methods for relation extraction." *Journal of Machine Learning Research*, **3**, 1083–1106.
- 大塚 (乾) 裕子 (2004). 自由記述アンケート回答の意図抽出および自動分類に関する研究-要求意図を中心に-. 博士論文, 神戸大学大学院自然科学研究科.
- 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一 (2005). "意見抽出のための評価表現の収集." *自然言語処理*, **12** (2), 203–222.
- 小林のぞみ, 乾孝司, 乾健太郎 (2001). "語釈文を利用した「p/n 辞書」の作成." 人工知能学会言語・音声理解と対話研究会 (SLUD-33), pp. 45–50.
- 小林のぞみ, 飯田龍, 乾健太郎, 松本裕治 (2005). "照応解析手法を利用した属性-評価値対および意見性情報の抽出." 言語処理学会第 11 回年次大会.
- 箆島郁子, 嶋田和孝, 遠藤勉 (2005). "系列パターンを利用した評価表現の分類." 言語処理学会第 11 回年次大会.
- 奥村学 難波英嗣 (2005). テキスト自動要約. オーム社.
- 奥村学, 南野朋之, 藤木稔明, 鈴木泰裕 (2004). "blog ページの自動収集と監視に基づくテキストマイニング." 人工知能学会セマンティックウェブとオントロジー研究会 (SIG-SWO-A401-01).
- 中山記男, 江口浩二, 神門典子 (2004). "感情表現の抽出手法に関する提案." 情報処理学会自然言語処理研究会 (NL-164-3), pp. 13–18.
- 中山記男, 江口浩二, 神門典子 (2005). "感情表現のモデル." 言語処理学会第 11 回年次大会.
- 乾健太郎, 徳久雅人, 徳久良子, 岡田直之 (2000). "感情の生起とその反応." *日本ファジィ学会誌*, **12** (6), 741–751.
- 立石健二, 福島俊一, 小林のぞみ, 高橋哲朗, 藤田篤, 乾健太郎, 松本裕治 (2004). "Web 文書集合からの意見情報抽出と着眼点に基づく要約生成." 情報処理学会自然言語処理研究会 (NL-163-1), pp. 1–9.
- 乾孝司, 乾健太郎, 松本裕治 (2004). "出来事の望ましさ判定を目的とした語彙知識獲得." 言語処理学会第 10 回年次大会, pp. 91–94.
- 江崎晃司, 松井藤五郎, 大和田勇人 (2005). "Weblog 上の評判情報における形容詞の出現位置を考慮した賛否分類." 情報処理学会第 67 回全国大会.
- 藤村滋, 豊田正史, 喜連川優 (2004). "Web からの評判および評価表現抽出に関する一考察." 夏のデータベースワークショップ (DBWS2004).
- 佐野真 (2004). "感性品質評価語辞書を利用したテキストマイニング." 情報処理学会論文誌: データベース, **45** (SIG7(TOD22)).
- 村野誠治 佐藤理史 (2003). "文型パターンを用いた主観的評価文の自動抽出." 言語処理学会第 9 回年次大会, pp. 67–70.

- 峠泰成 山本和英 (2004). “手がかり語自動取得による Web 掲示板からの評価文抽出.” 言語処理学会第 10 回年次大会, pp. 107–110.
- 鈴木泰裕 (2005). “Web データを利用した評価表現辞書の自動作成.” Master’s thesis, 東京工業大学大学院総合理工学研究科.
- 鈴木泰裕, 高村大也, 奥村学 (2004). “Weblog を対象とした評価表現抽出.” 人工知能学会セマンティックウェブとオントロジー研究会 (SW-ONT-A401-02).
- 岡野原大輔 辻井潤一 (2005). “評価文に対する二極指標の自動付与.” 言語処理学会第 11 回年次大会.
- 高村大也, 乾孝司, 奥村学 (2005a). “スピンモデルによる単語の感情極性判定.” 情報処理学会自然言語処理研究会 (NL-166-11).
- 高村大也, 乾孝司, 奥村学 (2005b). “極性反転に対応した評価表現モデル.” 情報処理学会自然言語処理研究会 (NL-168-22), pp. 13–18.
- 工藤拓 松本裕治 (2003). “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 (NL-158-9), pp. 55–62.
- 那須川哲哉 金山博 (2004). “文脈一貫性を利用した極性付評価表現の語彙獲得.” 情報処理学会自然言語処理研究会 (NL-162-16), pp. 109–116.
- 金山博 那須川哲哉 (2005). “要望表現の抽出と整理.” 言語処理学会第 11 回年次大会.
- 長江朋, 望月源, 白井清昭, 島津明 (2002). “製品コンセプトと製品評価文章の関係の分析.” 言語処理学会第 8 回年次大会, pp. 583–586.
- 飯田龍, 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一 (2005). “意見抽出を目的とした機械学習による属性-評価値対同定.” 情報処理学会自然言語処理研究会 (NL-165-4).
- 目良和也, 市村匠, 相澤輝昭, 山下利之 (2002). “語の好感度に基づく自然言語発話からの情緒生起手法.” 人工知能学会論文誌, **17** (3), 186–195.

略歴

乾 孝司: 1976 年生. 1999 年九州工業大学情報工学部卒業, 2001 年九州工業大学大学院情報工学研究科修士課程修了, 2004 年奈良先端科学技術大学院大学情報科学研究科博士課程修了. 同年, 東京工業大学 21 世紀 COE ポスドク研究員, 2005 年日本学術振興会特別研究員, 2006 年東京工業大学統合研究院助手, 現在に至る. 博士 (工学), 主に自然言語処理の研究に従事. 情報処理学会, 言語処理学会, ACL 各会員.

奥村 学: 1962 年生. 1984 年東京工業大学工学部情報工学科卒業. 1989 年同大学院博士課程修了. 同年, 東京工業大学工学部情報工学科助手. 1992 年北陸先端科学技術大学院大学情報科学研究科助教授, 2000 年東京工業大学精密工

学研究所助教授，現在に至る．工学博士．自然言語処理，知的情報提示技術，
語学学習支援，テキストマイニングに関する研究に従事．情報処理学会，人
工知能学会，AAAI，言語処理学会，ACL，認知科学会，計量国語学会各会員．
oku@pi.titech.ac.jp, <http://oku-gw.pi.titech.ac.jp/~oku/>.

(2005年10月7日受付)

(2006年1月2日再受付)

(2006年1月23日採録)