

2011年度 卒業論文

構成性に基づく評価極性知識獲得

2012年3月26日

情報知能システム総合学科

(学籍番号: A8TB2163)

鍋島 啓太

東北大学工学部

概要

事態の評価極性を判別することは言語処理において重要である。特に、単語の評価極性は、文への評価極性付与において重要な手がかりとなる場合が多い。しかし、周りの単語の組み合わせによって全体の評価極性が変化する場合がある。例えば、「風邪の予防」、「不幸の誕生日」という表現は、いずれも「ネガティブ」の「ポジティブ」という組み合わせだが、全体を見ると「風邪の予防」はポジティブ、「屈辱の勝利」はネガティブである。このように、単語の構成性によって評価極性が変化する場合がある。

本研究では、前述の例のような複数語表現のうち、さまざまな表現が存在する「名詞の名詞」(以下AのB)への評価極性付与を行う。まず、名詞への評価極性付与手法として、東山ら手法[1]を再実装し、「AのB」に対する評価極性付与とその評価を行う。次に、大量の「AのB」に対して評価極性を付与、その結果を分析することで、構成性を考える上で重要となる名詞句の主辞である「B」の性質を分析する。

目次

第1章	序論	1
第2章	関連研究	3
2.1	単語への評価極性付与	3
2.1.1	語彙ネットワーク	3
2.1.2	共起情報	4
2.2	複数語への評価極性付与	5
第3章	評価極性の定義	7
3.1	用言の評価極性の定義	7
3.2	名詞の評価極性の定義	7
3.3	「AのB」の評価極性の定義	7
第4章	複合語への評価極性付与	8
4.1	名詞の評価極性付与	8
4.1.1	基本的なアイデア	9
4.1.2	共起の抽出	9
4.1.3	共起行列の作成	9
4.1.4	分類器	10
4.2	「AのB」の評価極性付与	10
第5章	実験	12
5.1	実験設定	12
5.2	評価方法	12
5.3	実験結果	13
5.3.1	名詞との比較	13
5.3.2	低頻度との比較	13
第6章	被修飾名詞[B]の分析	15
6.1	全体の評価極性に対し、Bの極性が影響しない	16
6.2	全体の評価極性に対し、Bの極性が支配的である	16
6.3	BがAの評価極性を反転する	17
第7章	まとめ	18

第1章 序論

近年、個人が Web 上に情報を発信する機会の増加に伴い、Web 上から意見を分析する技術に、産業界を含め、さまざまな場所で注目が集まっている。例えば、大量の意見の中から肯定的な意見か否定的な意見かを自動分類する技術は、商品などの調査の高速化を行うためには必要不可欠である。特に新商品や新作映画が消費者にとって好評か不評かといった情報をウェブ上から集めるのは、製作者や販売者にとってマーケティングをする上で非常に重要である。ここで商品が好評か不評かといった情報は、テキスト中に含まれる事態の評価極性を抽出することにより自動分類可能である。ここで事態の評価極性とは、事態が望ましい(ポジティブ)か望ましくない(ネガティブ)かそのどちらでもない(ニュートラル)かの3値からなる。

これまでにテキストへの評価極性付与の研究は数多くなされてきた。その中でも単語の評価極性は、以下の例のように文への評価極性付与において重要な手がかりとなる場合が多い。

1. 東北大学の入試試験に 合格した ポジティブ
2. 部屋に ゴミ ネガティブが増えた

しかし複数語の極性は単純な極性の和では決まらず、周りの単語の組み合わせによって全体の評価極性が変化する場合がある。例として以下の2つをあげる。

3. 「ゴミ ネガティブの 削減 ニュートラル」
4. 「洗濯物 ニュートラルの 乾燥 ネガティブ」

上の例3について、「ゴミ」単体の評価極性はネガティブだが、そこに「削減」という極性を反転させる単語が組み合わさることで前の極性が反転し、全体の評価極性がポジティブとなっている。例4について、「洗濯物」単体の評価極性はニュートラルであり、「乾燥」という単語も単体ではネガティブだが、「洗濯物の乾燥」と2つの単語が組み合わさることで、全体の評価極性がポジティブとなっている。このように、構成語のタイプや組み合わせやによって全体の評価極性が変化する場合がある。

本研究では数ある複数語表現のうち、名詞句である「名詞の名詞」(以下 A の B) への評価極性付与を行う。「A の B」という表現は日本語の中では頻繁に出現する。A と B の組み合わせによって、イベントを表すものや修飾の関係にあるものなど、さまざまなタイプが存在し [2]、他の複数語表現への評価極性付与を考える上で基盤となると考えている。このような理由から本研究では「A の B」を評価極性付与の対象とした。

「A の B」へ評価極性付与の方法として、東山ら [1] の名詞への評価極性付与の手法を「A の B」へ適応することで評価極性付与を行い、その性能の評価を行う。次に、大量の「A の B」に対して上記の方法により評価極性を付与、その結果を分析することで、構成性を考える上で重要となる名詞句の主辞「B」の性質を分析する。

2章では評価極性に関わる関連研究についての調査を行い、3章で評価極性の定義について述べる。4章では「AのB」への評価極性付与を行う提案手法について説明し、5章でその手法の評価実験を行う。6章では4章で提案した手法を大規模に行うことにより、「AのB」のうち名詞句の主辞である被修飾名詞(B)の分析を行った結果を報告する。7章で全体のまとめと今後の課題について述べる。

第2章 関連研究

事態の評価極性を自動判別する技術についての研究は今まで様々なものが研究されてきた。単語への評価極性付与を対象としたものや、複数語や句単位、述語項構造単位の評価極性付与などさまざまである。本章ではそれぞれの代表的なものについて説明する。

2.1 単語への評価極性付与

単語への評価極性を付与する研究はこれまでに様々なものが存在する。主な分類として、語彙ネットワークを用いた手法 [3, 4, 5, 6, 7, 8] や、共起情報を利用した手法 [9, 10, 11, 1] がある。それぞれについて説明する。

2.1.1 語彙ネットワーク

語彙ネットワークを用いた評価極性辞書の構築には、まず WordNet などのを用いてそこに含まれる語彙をノード、同義語、類義語、反義語関係にあるものをエッジで結ぶことで語彙ネットワークを構築し、評価極性をもつ少量のシードを与えることで評価極性を伝播させてネットワーク全体のノードの評価極性付与を行う。

Kamps ら [4] は、WordNet の類義関係 (synonymy) にある形容詞をリンクで結んで語彙ネットワークを構築し、以下のスコア関数により形容詞 w がポジティブであるかネガティブであるかを判定する。

$$SO - score(w) = \frac{d(w, bad) - d(w, good)}{d(good, bad)} \quad (2.1)$$

ここで $d(t_a, t_b)$ は2つの形容詞の最短経路の距離を示す。つまりこの式は形容詞 w がネットワーク上で「good」、「bad」どちらに近いかを表しており、「good」に近い、つまり式の値が正の値であればポジティブの極性、逆に「bad」に近い、つまり式の値が負の値であればネガティブの極性を示す。

Hu ら [5] も、WordNet を用いて形容詞の評価極性を判定する手法を提案している。彼らは Kamps らが提案した類義関係に加え、反義関係 (antonymy) の関係も利用した。彼らは評価表現が既知の形容詞 30 語を用意し、順々に WordNet の類義関係、反義関係を辿っていく。類義関係を辿れば、辿り元と辿り先は同じ評価極性を割り当て、逆に反義関係を辿ると、辿り元と辿り先は違う評価極性を割り当てていく。この操作を辿り先が WordNet 内になくなるまで繰り返すことで形容詞に評価極性付与を行った。

次に、共起情報と語彙ネットワークを用いた手法を紹介する。Hatzivassiloglou ら [3] は、「and」や「but」といった接続表現に着目した。「simple and well-received」のように and で接続されてい

る単語の極性は同じであり、「simplistic and well-received」のように but で接続されている単語の極性は反対であるという考え方に基づく手法である。コーパス中から「形容詞+接続詞+形容詞」という品詞列ををすべて抽出し、形容詞をノード、形容詞と形容詞をエッジで結ぶことにより語彙ネットワークを構築する。次にノードを結んでいる接続詞の種類によって、ノードである2つの形容詞が一致、不一致であるかを付与していく。例えば「and」であれば2つの形容詞の評価極性が一致、「but」であれば不一致などである。そして、クラスタリングを行い2種類に分類し、最終的に一方をポジティブ、もう一方をネガティブとすることで語彙ネットワーク全体の形容詞の評価極性を決定する。この手法の都合上、評価対象は形容詞のみである。

Takamura ら [6] は、各電子がスピンと呼ばれる方向をもつスピン系のモデルを語彙ネットワークに用いた。彼らは単語の極性(ポジティブまたはネガティブ)を電子のスピン(上向きまたは下向き)の方向と見なし、それぞれの単語をネットワークのノードとし、WordNet の語釈文、シソーラスでの類義語ペア、上位語下位語ペア、「healthy and delicious」のようにコーパス中の and などの接続詞で連結されて出現する形容詞ペアを正の重みで連結する。このリンクの集合を SL とする。同様に反義語ペア、「but」のように but などの接続詞で連結されて出現する形容詞ペアを負の重みで連結する。このリンクの集合を DL とする。すると、エッジの重みは以下の式で表される。

$$w_{ij} = \begin{cases} -\frac{1}{\sqrt{d(i)d(j)}} & (l_{ij} \in SL) \\ \frac{1}{\sqrt{d(i)d(j)}} & (l_{ij} \in DL) \\ 0 & otherwise \end{cases} \quad (2.2)$$

ここで l_{ij} は単語 ij 間のリンクを表し、 $d(i)$ は単語 i がリンクしている数を表す。彼らは平均場近似を利用することでスピン系の状態を近似的に求め、それぞれの単語の評価極性の判定を行った。

最近の研究では、Ahmed ら [7] が、語彙ネットワークにマルコフランダムウォークモデルを適応した手法がある。このモデルの利点として、他のネットワークを利用した手法に比べ正確かつ迅速に任意の単語の評価極性を求めることができる。

2.1.2 共起情報

共起情報を利用した研究では、まず初めにシードとなる代表的なポジティブとなる単語(例えば good,excellent) とネガティブな単語(例えば bad,poor) をいくつか用意しておいて、それらと評価極性を求めたい単語の文章中での共起頻度の比を比べることで、評価極性を決定する。

Turney ら [9] は、評価極性を付与したい語句と評価表現のコーパス中の共起を測ることで評価極性を決定した。彼らは以下の SO-PMI(Semantic Orientation Pointwise Mutual Information) を用いてポジティブな単語との共起の強さとネガティブな単語との共起の強さの差分を求めることで、評価極性を決定した。

$$SO - PMI(w) = PMI(w, excellent) - PMI(w, poor) \quad (2.3)$$

ここで PMI とは自己相互情報量のことであり、単語 word1 と単語 word2 の自己相互情報量は以

下の式で与えられる。

$$PMI(word_1, word_2) = \log \frac{P(word_1, word_2)}{P(word_1)P(word_2)} \quad (2.4)$$

単語が共起関係にあるかどうかは、単語の前後 10 単語以内に存在するかどうかで決定する。単語が「excellent」と共起しやすいほど SO-PMI の値が正に大きく、「poor」と共起しやすいほど SO-PMI の値は負に大きい。彼らの手法では、WordNet やシソーラスなどの語彙資源を必要としないが、信頼のおける大規模なコーパスが必要になる。

Kaji ら [10] は語彙統語パターンを用いた手法を提案している。彼らは「良い事は～こと」や、「～は悪いことだ」など、評価表現の出現しやすいパターンをいくつか用意しておいて、そのパターンを利用して評価文コーパスを自動構築する。次にコーパス中から抽出した候補表現(名詞+格助詞+述語)の極性を人手でラベル付した文との SO-PMI を計算することで、極性値の高い表現を順次獲得できる。

東山ら [1] は述語の選択選好性に着目した名詞評価極性の獲得の手法を提案している。彼らは「心掛ける」のヲ格には「安全運転」や「心配り」といったポジティブな行為を表す単語が出現しやすく、反対に「防ぐ」のヲ格には「危険」や「事故」などといったネガティブな出来事を表す単語が共起しやすいという性質を利用した。名詞の素性として、名詞と共起しやすい「格助詞+述語」を用意し、分類器学習を行うことで評価極性の決定を行った。この手法は Turney らの手法に比べより精緻な文脈ベクトルを与えることができるが、評価対象は名詞のみである。

同じく述語の共起パターンを利用した研究に、De ら [12] の研究がある。彼らは評価極性付与ではなく、トラブル表現の収集にこれを利用した。トラブル表現とは、災害や病障壁や規制など、人間の活動に負荷を与えたり、マイナス効果をもたらしたりする事物に関する表現のことである。

2.2 複数語への評価極性付与

複数語や句単位, または述語とその項からなる述語項構造に対しての評価極性の判定を行う研究など、もさまざまなものが存在する。

共起情報などの周辺情報を利用した研究としては、先ほども述べた Turney らの研究があげられる。この手法は、既存の辞書や WordNet のエントリ情報などを利用しないため、複数語への評価極性付与も行うことができる。

酒井ら [13] は、評価表現の主要部分は「属性、対象、評価表現」といった小林ら [14] の研究と同様に、この 3 つ組を評価対象とし、「綺麗だ」と「汚い」などの評価表現と反対の意味を持つ反義語の評価極性は異なるという特徴を利用し、その共起情報を比較することで判定を行った。

しかしこれらの手法の場合、ニュートラルの判別を行うことができない。なぜならばニュートラルとの共起を図ろうにも、ニュートラルとして特徴的な語は存在しないためである。

構成語の属性を考慮した研究もいくつか存在する。Willson ら [15] は、句単位での評価極性の付与を行い、先に評価極性の有無を分類してから評価極性の判定を行うことで精度が向上することを示した。

高村ら [16] は名詞と形容詞からなる表現に対して評価極性を行うにあたり、複数語の特徴を捉えるためにモデルに隠れ変数を導入した。彼らも述べていることだが、この手法では訓練データ

に出現しない未知語に対しての評価極性の付与は行えない。この問題に対処するには Web データなどの大規模データを利用する必要があると思われる。

Lu ら [17] は、評価極性の文脈依存問題に対処するために、評価極性付与に有効であることが知られている複数の手がかりを用い、最適化問題を解くことでドメインごとの辞書を構築する手法を提案した。文脈依存問題とは、「電池のバッテリーが長い」、「トイレの待ち時間が長い」のように、同じ「長い」という単語であっても前者はポジティブ、後者はネガティブと、文脈によって評価極性が変化する問題である。彼らは既存の評価極性辞書、同義語反義語辞書、and や but などの接続表現、レーティングの付いたドメインごとのレビュー文章を手がかりとし、ドメインを表す単語が含まれる句に対して評価極性付与を行った。ドメイン依存の問題だけでなく、同じドメインでも場面によって評価極性が変わる場合でもきちんと判別できることを示した。

第3章 評価極性の定義

評価極性の定義は既存研究である東山らの定義 [1] と同様のものを用いた。「A の B」の評価極性の定義は、「A の B」という名詞句を一つの名詞と見ること、名詞に対する評価極性の定義と同様のものと定義する。名詞の評価極性を定義するため、まずは用言の評価極性の定義を行う。

3.1 用言の評価極性の定義

動詞、形容詞などの用言の評価極性については、その事態が望ましいか否かで評価極性を決定する。望ましいものをポジティブ、望ましくないものをネガティブ、さらにそのどちらにも分類できないものをニュートラルとする。例えば、「合格する」はポジティブなイベントであり、「汚い」はネガティブな方向への評価である。

3.2 名詞の評価極性の定義

名詞の評価極性は、以下の3つの表現をを補い述語にした上で、評価極性を決定する。

1. (～する) を付与して動詞化
2. (～だ) を付与して形容詞化
3. (～が増える) や (～がある) を付与して動詞化

1. は「信頼」などのサ変名詞、2. は「綺麗」などの形容詞語幹の名詞に対しての評価極性付与の基準であり、それぞれ動詞化、形容詞化を行い用言の評価極性の定義の元、評価極性を決める。3. について、例えば「お金」という名詞について、それ自体はニュートラルだが、「お金がある」や「お金が増える」のように「お金」の後に plus 演算子を付与した場合、全体の極性がポジティブであるならば、「お金」はポジティブな名詞と判断する。逆に「お金がない」や「お金が減る」のように、「お金」の後に minus 演算子を付与した場合、全体の極性がネガティブであるならば、「お金」はポジティブな名詞と判断する。

3.3 「A の B」の評価極性の定義

「A の B」の評価極性の定義は、被修飾名詞 (B) だけでなく、修飾名詞 (A) も考慮し、「A の B」という名詞句全体に対して前節の名詞の評価極性の定義を適応する。例えば、B にくる名詞が「誕生日」というポジティブな単語であったとして、「友達の誕生日」であればポジティブだが、「一人ぼっちの誕生日」のように修飾されればネガティブとなる。

第4章 複合語への評価極性付与

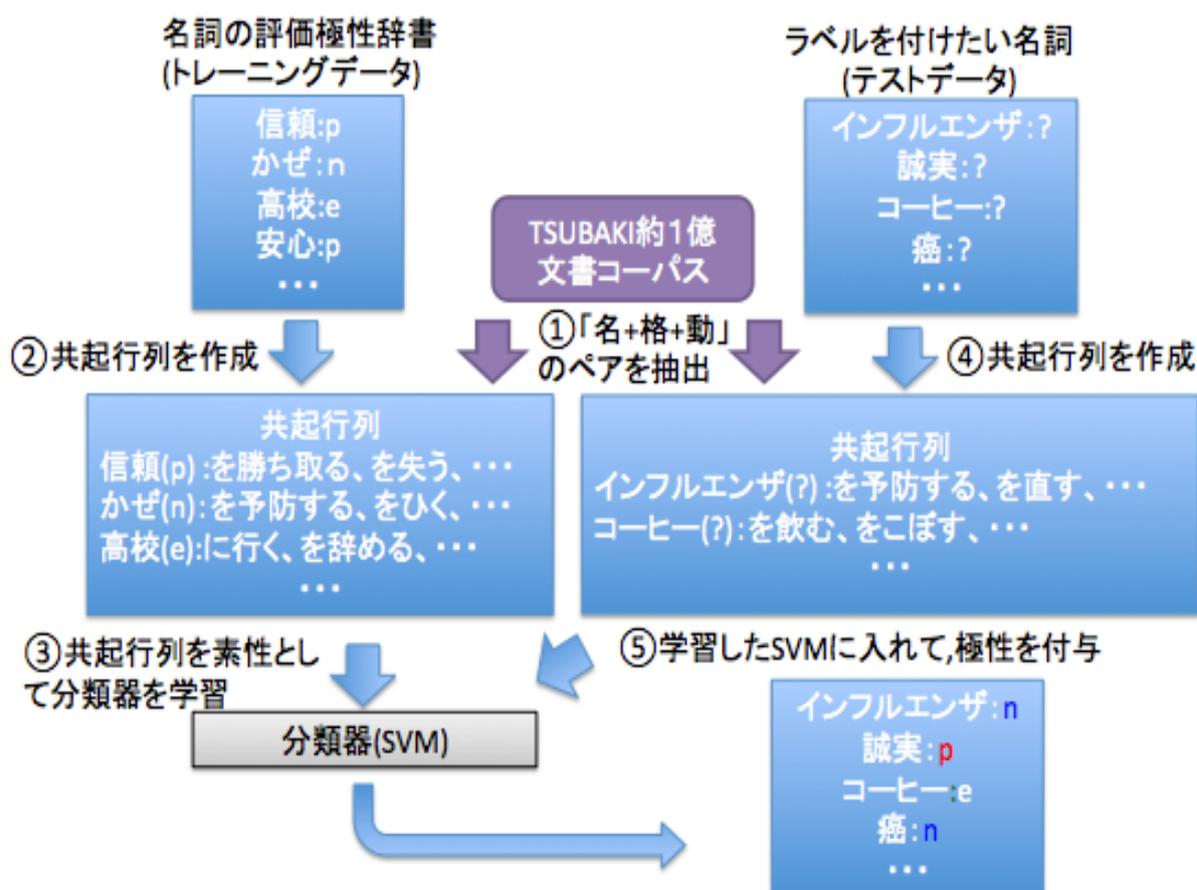


図 4.1: 名詞の評価極性付与

「A の B」へ評価極性を付与するために、本研究では東山ら [1] の名詞に対する評価極性付与の既存手法を拡張することにより、「A の B」への評価極性付与を行った。まず初めに名詞への評価極性付与の方法について説明し、次に「A の B」への適応方法を示す。

4.1 名詞の評価極性付与

「A の B」に対して評価極性を付与するために、東山らの述語の選択選好性に着目した名詞の評価極性付与手法 [1] を利用することにした。この手法を選んだ理由としては、名詞への評価極性を高い精度で行えるということ、さらに辞書にない単語に対しても評価極性を付与できるということがあげられる。

4.1.1 基本的なアイデア

この手法の基本的なアイデアである述語の選択選好性について、これは「防ぐ」のヲ格には「危険」や「事故」などといったネガティブな単語が共起しやすいという性質を利用する。手法の概要図は図 4.1 になる。手順としては、単語と共起しやすい「格助詞+動詞」をコーパスから抽出、共起行列を作成する。次に人手でポジティブ、ネガティブ、ニュートラルのラベルを付与した名詞を訓練事例としてその共起行列とともに分類器で学習、分類を行う。学習後、極性の付与を行いたい名詞とその共起行列を学習された分類器で分類することで、評価極性の付与を行う。

以下でそれぞれの部分を具体的に説明する。

4.1.2 共起の抽出

コーパスから「名詞+格助詞+動詞」の3つ組の共起を網羅的に抽出する。抽出する上で、以下に示すいくつかの制約を設けた。

1. 格助詞をガ格、ヲ格、デ格に限定
2. 動詞とその直前の共起のみを抽出

2.について、これは「名詞」と「格助詞+動詞」の共起の距離に近いほど、意味的なつながりが強いためこの制約を設定した。また動詞について、「走る」「走った」「走ります」などのように複数の表現をまとめるために、いくつかのストップワードを設け、動詞は原形に統一した。例外として、「走らない」などの否定表現や、「走ることができる」などの可能表現は分類の素性として有効であると思われるので、「走る(否)」や「走る(可)」のように「走る」とは別の素性として区別した。

4.1.3 共起行列の作成

名詞と、それと共起する述語パターンをそれぞれ行と列とする共起行列を作成する。抽出してきた共起から共起行列を作成する上で、すべてのペアを共起行列に加えるのではなく、以下の4つの制約を設けた。

3. 名詞と「格助詞+動詞」の自己相互情報量が閾値(0)以下のものを削除
4. 頻度が閾値(100万)以上の「格助詞+動詞」を削除
5. 頻度が閾値(50)以下の名詞または「格助詞+動詞」を削除
6. 素性数が閾値(2)以下の訓練事例を削除

3.について、自己相互情報量(PMI)は以下の式で表される。

$$PMI(word_1, word_2) = \log \frac{P(word_1, word_2)}{P(word_1)P(word_2)} \quad (4.1)$$

このスコアが低いと、「名詞」と「格助詞+動詞」の間の相互依存が弱く、素性としては有用でないため削除する。

4. について、頻度が高い「格助詞+動詞」(以下述語パターン)には、「を行う」や「が話す」など、一般的な述語パターンが多く、分類に有効な素性でないので削除する。

5. について、これは特殊な名詞、述語パターンを削除するために設けた。

6. について、素性数が少ないとうまく学習できないので共起行列から該当する名詞を除去した。

4.1.4 分類器

作成した共起行列を素性として、名詞の評価極性を分類する分類器の学習を行う。分類器には線形カーネルのサポートベクターマシン (SVM) を用いた。分類するのはポジティブ、ネガティブ、ニュートラルの3値であるが、ニュートラルを表す強い特徴はないと思われるので、訓練事例をポジティブとその他、ネガティブとその他の2つにわけ、両方のモデルでその他に分類されたものをニュートラルとした。

4.2 「AのB」の評価極性付与

名詞の評価極性付与の手法について説明してきたが、ここで「AのB」へ適応する方法について説明する。「AのB」への評価極性付与の手法の概要図を図4.2に示す。図に示したとおり、分類器の学習には先程と同様評価極性が付いた名詞の共起行列を用いる。モデル学習後、名詞の代わりに今度は「AのB+格助詞+動詞」のペアをコーパスから抽出、ラベルを付与したい「AのB」と抽出した述語パターンとの共起行列を作成、学習した分類器に与えることで「AのB」へ極性の付与を行う。

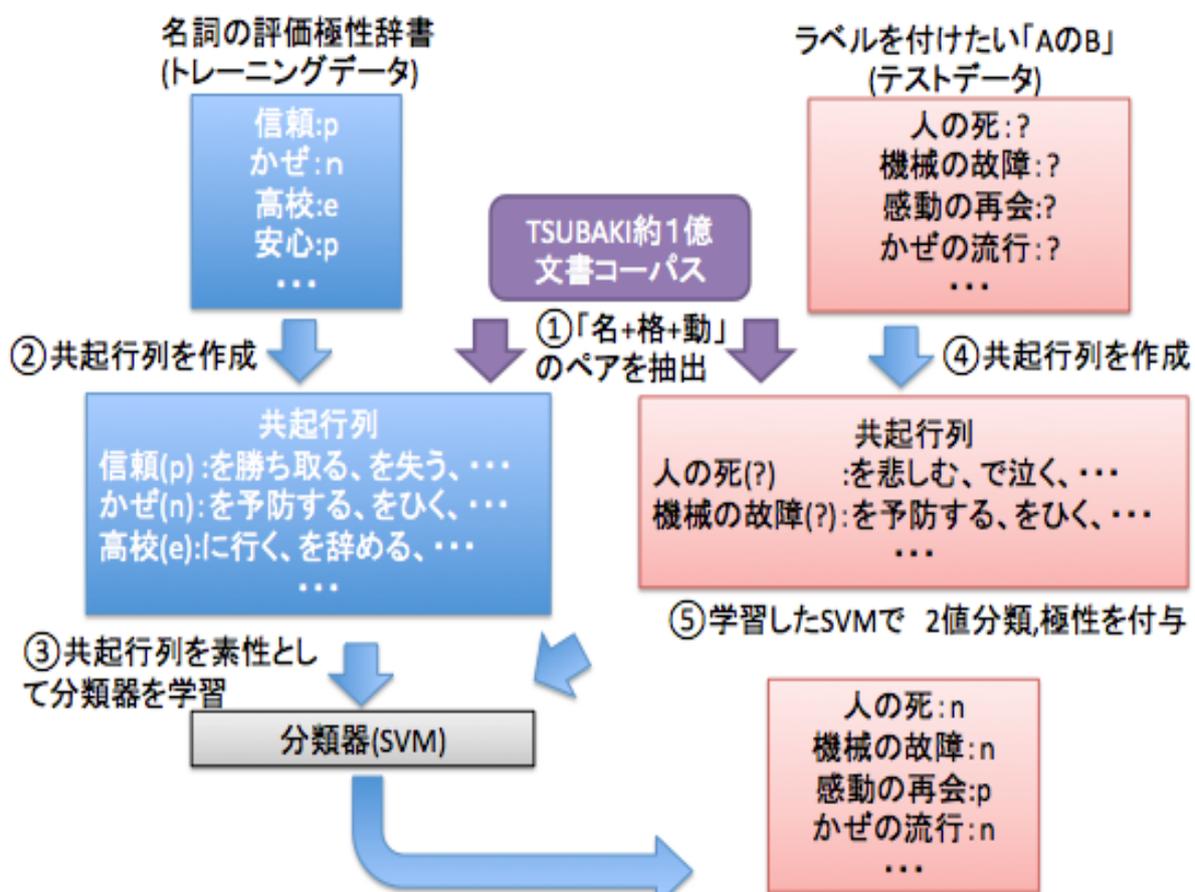


図 4.2: A の B への評価極性付与

第5章 実験

5.1 実験設定

分類器の学習データとして、日本語評価極性辞書(名詞編)と高度言語融合フォーラムから公開されている意見抽出ツール用モデルに付属されている評価表現辞書を利用した。ここからポジティブな単語 6330 語、ネガティブな単語 15636 語、ニュートラルな単語 3459 語をトレーニングデータとして用いた。コーパスは検索エンジン TSUBAKI[18] の約 1 億 1 千万文書を利用した。正解データは 2 種類を用意した。1 つは「A の B」の単語頻度上位 10 万単語からランダムに抽出、人手でラベルを付与したポジティブな単語 247 個、ネガティブ単語 86 個、ニュートラルな単語 492 個である(以下高頻度データ)。もう一つは、「A の B」の単語頻度上位 10 万から 200 万単語の間からランダムに抽出、人手でラベルの付与を行い、ポジティブな単語 224 個、ネガティブな単語 92 個、ニュートラルな単語 621 個である。(以下低頻度データ)。

この手法を名詞へ適応した場合との性能を比較するために、名詞に評価極性を付与した場合との性能を比較した。名詞の正解データは高頻度データと同様に名詞頻度上位 10 万単語からランダムに抽出、人手でラベルの付与を行い、ポジティブな単語 117 個、ネガティブな単語 174 個、ニュートラルな単語 838 個を評価事例として用いた。

また実験結果を比較するため、ベースラインとしてはランダムに評価極性を付与するものを用い、スコアを比較することで評価を行った。

5.2 評価方法

実験の評価は、Precision(精度)、Recall(再現率)、F1 を用いて評価した。評価極性付与におけるそれぞれの値は以下のとおりである。

$$Precision = \frac{\text{ポジティブを正しく出力できた数}}{\text{システムがポジティブと出力した数}} \quad (5.1)$$

$$Recall = \frac{\text{ポジティブを正しく出力できた数}}{\text{ポジティブが正解である事例数}} \quad (5.2)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5.3)$$

ネガティブ、全体(ポジティブとネガティブを合わせたもの)の場合は、それぞれ上記の式のポジティブの部分置き換えたものである。

表 5.1: 「A の B」への評価極性付与のスコアが高い上位 10 件

ポジティブな「A の B」	スコア	ネガティブな「A の B」	スコア
天使の羽	3.16	不慮の事故	4.91
神の力	2.81	新型インフルエンザの感染拡大	4.75
日本のお土産	2.60	今回の危機	4.17
音楽の力	2.38	自分のエゴ	3.81
今年の秋	2.34	全身の痛み	3.64
旅の魅力	2.19	価値観の相違	3.57
京都の紅葉	2.15	システムの変更	3.43
左手の薬指	2.14	顔のシミ	2.95
花の香	2.11	歪みの原因	2.94
手書きの手紙	2.06	横領の疑い	2.40

5.3 実験結果

5.3.1 名詞との比較

提案手法を名詞へ適応した場合と、「A の B」へ適応した場合とでの精度の差を比較した。実験結果を表 5.2 に示す。この表より、「A の B」へ極性を付与した場合より名詞へ極性を付与した場合のほうが精度は若干高かった。この結果より、名詞の評価極性付与の既存手法が「A の B」へ十分適応可能だということが分かる。

表 5.1 に、高頻度データにおいて実際に高いスコア (分離平面からの距離が大きい) でポジティブ、ネガティブと判別された「A の B」トップ 10 を表示する。上位の単語に対しての精度の良さを読み取ることができる。

名詞と比べ「A の B」のほうが精度が下がった理由として、「A の B」の定義のゆれの問題がある。名詞への評価極性の場合、名詞そのものの普遍的な評価が多いが、「A の B」の場合、「ソニーの製品」、「国産の肉」のように、名詞句の好悪を判定するものが多くなり、評価にブレがあったためと思われる。

また名詞と「A の B」のどちらに評価極性付与を行った場合でも、Recall の値が高く、Precision の値が低かった。これは学習に用いたポジティブ、ネガティブの訓練事例がニュートラルの訓練事例と比べ非常に少なかったのに対し、実際のテストデータではポジティブ、ネガティブと比べニュートラルの事例が多かったためと思われる。Shoushan ら [19] も述べていることだが、評価極性において訓練事例に偏りがある場合とない場合では精度に差が出てくる。比率を合わせるために単純に訓練事例を減らすと、全体の性能が下がるため、今後改善する必要がある。

5.3.2 低頻度との比較

実験結果を表 5.3 に示す。この表から、高頻度データ、低頻度データともにベースラインと比較して提案手法の性能が上回っていることが分かる。しかし高頻度データにくらべ、低頻度データの精度は低くなった。これは低頻度は高頻度に比べ共起する述語の数が少なく、有用な素性が得られにくいからである。

高頻度データに対しては、ある程度の精度で評価極性付与を行うことができたが、このような高頻度データに含まれる「A の B」は、コーパスから抽出した「A の B」全体の約 0.2% にすぎず、

表 5.2: 名詞と「A の B」への評価極性付与の比較

	全体 (PR F1)	ポジティブ (PR F1)	ネガティブ (PR F1)
A の B への評価極性付与 (高)	.375 .621 .468	.415 .591 .487	.305 .709 .427
ベースライン (A の B)	.201 .327 .249	.290 .329 .306	.109 .337 .165
名詞への評価極性付与	.386 .804 .522	.315 .768 .447	.530 .851 .653
ベースライン (名詞)	.146 .363 .209	.164 .362 .225	.129 .366 .190

表 5.3: 「A の B」への評価極性付与における高頻度と低頻度

	全体 (PR F1)	ポジティブ (PR F1)	ネガティブ (PR F1)
提案手法 (高)	.375 .621 .468	.415 .591 .487	.305 .709 .427
ベースライン (高)	.201 .327 .249	.290 .329 .306	.109 .337 .165
提案手法 (低)	.241 .516 .328	.289 .473 .359	.183 .516 .285
ベースライン (低)	.163 .322 .217	.220 .316 .260	.102 .337 .157

低頻度の問題に対処する必要がある。次節では、提案手法によって大量の「A の B」に評価極性を付与することで、名詞句のヘッドである B の性質に着目し、その分析を行う。

第6章 被修飾名詞[B]の分析

全体の単語の極性

Aの単語の極性		p	n	e	Aの合計
	P	C _{pp}	C _{pn}	C _{pe}	C _p *
	n	C _{np}	C _{nn}	C _{ne}	C _n *
	e	C _{ep}	C _{en}	C _{ee}	C _e *
	全体の合計	C*p	C*n	C*e	C**

図 6.1: 評価極性の分布

Bの性質を分析するために、提案手法を用いて「AのB」の頻度上位200万件に対し評価極性を付与した。名詞句の主辞である「B」の単語に着目し、Aにポジティブな単語、ネガティブな単語がついたときに、全体の評価極性がどのように変化するか観点から「B」の分類を行った。

評価極性の変化を図6.1に示す。この図は、Aの単語の評価極性がどのようなときに、全体の評価極性がどうなったかの分布を表している。例えば、C_{pp}というのは、Aの単語の評価極性がポジティブな単語のとき、全体の評価極性がポジティブになった単語の数を表している。この分布を用いることで、Bの単語の分類を行った

分析の結果、Bを以下の3パターンに分類することができた。

1. 全体の評価極性に対し、Bが影響しない
2. 全体の評価極性に対し、Bが支配的である
3. BがAの評価極性を反転する

以下ではそれぞれどのようなものを説明し、そして実際に獲得できたものを紹介する。

6.1 全体の評価極性に対し、Bの極性が影響しない

このタイプについて、これは被修飾名詞Bの評価極性がポジティブ、ネガティブにかかわらず、全体の評価極性が修飾名詞Aの評価極性を受け継ぐものである。例えば「才能の芽」のようにAにポジティブな単語がくると全体がポジティブになり、「犯罪の芽」のようにAにネガティブな単語がくれば全体がネガティブとなるような名詞Bのことである。

このタイプに分類される名詞は、図6.1のC_{pp}やC_{nn}などの数が多いと思われるので、以下の式で特徴な単語を獲得する。

$$\frac{C_{pp}}{C_{p*}} + \frac{C_{nn}}{C_{n*}} \quad (6.1)$$

C_{pp},C_{nn}それぞれ3回以上出現するものでフィルタリング、かつ上の式でランキングを行い、その上位15件を表6.1に示す。獲得できた単語について、「増加」や「拡大」など、単独では極性が見つからないものや、「夫」「天気」など、元々の評価極性がニュートラルの一般的な名詞などが上位に現れた。

表 6.1: 「B」が影響しない

拡大 成長 増大 広がり 増加 高まり お母さん 芽 上昇 多く 天気 ことば 企業 恋 夫
--

6.2 全体の評価極性に対し、Bの極性が支配的である

このタイプについて、これは修飾名詞Aの評価極性がポジティブ、ネガティブにかかわらず、被修飾名詞Bの評価極性が全体に対して支配的なものである。全体が常にポジティブな例として、「震災の復興」、「町の復興」や、逆に全体的に全体が常にネガティブな例として「人材の流出」、「化学物質の流出」などがあげられる。

このタイプに分類される名詞は、「B」がポジティブな名詞の場合は図6.1のC^{*p}が多く、ネガティブな場合はC^{*n}などの数が多いと思われるので、以下の式で特徴な単語を獲得する。

$$\frac{C^{*p}}{C^{**}} \text{ or } \frac{C^{*n}}{C^{**}} \quad (6.2)$$

C^{*p},C^{*n}それぞれ5回以上出現するものでフィルタリング、かつ上の式でランキングを行い、その上位15件を表6.3に示す。

獲得できた単語について、ポジティブの方では「安定」「回復」など、Aがネガティブなときでも全体がポジティブになる単語が得られたが、「役目」「誕生日」など、この分類にはふさわしくないものも得られた。これは提案手法の評価極性付与の誤りにより、上位にランクインしてしまったためである。

ネガティブの方もポジティブと同様に、「劣化」や「流失」などAにかかわらず全体がポジティブになる単語が得られた。

表 6.2: 「B」が支配的

B が支配的 (ポジティブ)	誕生日 役目 お手伝い 役割 雪辱 ご連絡 成功 異名 復興 回復 実現 チャンス 登場 浄化 安定
B が支配的 (ネガティブ)	食い違い 劣化 ミスマッチ 不足 地震 流出 渋滞 遅延 借り 荒廃 流失 損傷 制止 トラブル 拡散

6.3 B が A の評価極性を反転する

このタイプについて、これは修飾名詞 A の評価極性がポジティブ/ネガティブの場合、評価極性を反転させ全体の極性をネガティブ/ポジティブにするものである。

評価極性を反転させる単語の例として、「やる気の低下」、「死亡率の低下」や「可能性の消滅」、「癌の消滅」などがあげられる。

このタイプに分類される名詞は、「B」がポジティブな名詞の場合は図 6.1 の C_{pn} や C_{np} が多いと思われるので、以下の式で特徴な単語を獲得する。

$$\frac{C_{pn}}{C_{p*}} + \frac{C_{np}}{C_{n*}} \quad (6.3)$$

C_{pn}, C_{np} それぞれ 3 回以上出現するものでフィルタリング、かつ上の式でランキングを行い、その上位 15 件を表 6.3 に示す。

獲得できた単語について、「低下」「縮小」「中止」などこのタイプに該当する単語を得られたが、他の分類と比べ、「彼氏」「向こう側」「掲示板」など評価極性反転とは関係ない単語も多く得られた。これはこの分類に属する名詞の総数が他に比べ少ないためだと思われる。

表 6.3: A の極性を反転

引き下げ 洞窟 低下 縮小 減少 衝動 彼氏 報 向こう側 掲示板 渦 溝 効能 ないこと 中止
--

第7章 まとめ

本研究では、名詞の評価極性付与の既存手法を拡張することにより、「AのB」という複合語に対して評価極性を付与し、ベースラインを上回る精度を実現した。次に、この手法を大規模に行い、名詞句の主辞である「B」に着目して分析を行うことで、低頻度な「AのB」に対しても「AのB」の評価極性を予測できる可能性を検証した。

今後の課題としては、「AのB」について、それぞれの名詞の性質を利用することでモデルの改善を行いたい。また、今回は「AのB」のみに焦点をあてて評価極性付与を行ったが、今後は名詞にかぎらず、動詞や形容詞など、他の評価表現への評価極性付与や、述語項構造や句単位の評価極性付与を行っていく。

謝辞

本研究を進めるにあたり、ご指導、御助言を頂きました主指導教員である乾健太郎教授に深く感謝いたします。

本研究内容に関して、ご指導、御助言を頂きました岡崎直観准教授に深く感謝いたします。

本研究内容に関して、いろいろと御助言を頂きました渡邊陽太郎助教授に深く感謝いたします。

本研究内容に関して、数多くの御助言、相談にのっていただきました水野淳太さんに深く感謝いたします。

研究生生活や学生生活を暖かく見守って下さいました八巻智子秘書に心から感謝いたします。

最後になりましたが、研究生生活の様々な場面でお世話になりました研究室の皆様有難うございました。

参考文献

- [1] 東山昌彦. 述語の選択選好性に着目した名詞評価極性の獲得. 2008.
- [2] 植村将人. 生成語彙論に基づく名詞句「aのb」の意味解釈. 2005.
- [3] V. Hatzivassiloglou and K.R. McKeown. Predicting the semantic orientation of adjectives. In *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*, pp. 174–181. Association for Computational Linguistics, 1997.
- [4] J. Kamps, MJ Marx, R.J. Mokken, and M. De Rijke. Using wordnet to measure semantic orientations of adjectives. 2004.
- [5] M. Hu and B. Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177. ACM, 2004.
- [6] H. Takamura, T. Inui, and M. Okumura. Extracting semantic orientations of words using spin model. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pp. 133–140. Association for Computational Linguistics, 2005.
- [7] A. Hassan and D. Radev. Identifying text polarity using random walks. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pp. 395–403. Association for Computational Linguistics, 2010.
- [8] L. Velikovich, S. Blair-Goldensohn, K. Hannan, and R. McDonald. The viability of web-derived polarity lexicons. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 777–785. Association for Computational Linguistics, 2010.
- [9] P.D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pp. 417–424. Association for Computational Linguistics, 2002.
- [10] N. Kaji and M. Kitsuregawa. Building lexicon for sentiment analysis from massive collection of html documents. In *Proceedings of the joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, pp. 1075–1083, 2007.
- [11] H. Kanayama and T. Nasukawa. Fully automatic lexicon expansion for domain-oriented sentiment analysis. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pp. 355–363. Association for Computational Linguistics, 2006.
- [12] S. De Saeger, K. Torisawa, and J. Kazama. Looking for trouble. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pp. 185–192. Association for Computational Linguistics, 2008.

- [13] 酒井義和, 荒木健治. 反対語を利用した文脈依存評価表現の感情極性判定 (自然言語処理). 電子情報通信学会論文誌. D, 情報・システム, Vol. 93, No. 9, pp. 1778–1789, 2010-09-01.
- [14] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一. 意見抽出のための評価表現の収集. 自然言語処理 = Journal of natural language processing, Vol. 12, No. 3, pp. 203–222, 2005-07-10.
- [15] T. Wilson, J. Wiebe, and P. Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 347–354. Association for Computational Linguistics, 2005.
- [16] 高村大也, 乾孝司, 奥村学. 極性反転に対応した評価表現モデル. 情報処理学会研究報告, 2005.
- [17] Y. Lu, M. Castellanos, U. Dayal, and C.X. Zhai. Automatic construction of a context-aware sentiment lexicon: an optimization approach. In *Proceedings of the 20th international conference on World wide web*, pp. 347–356. ACM, 2011.
- [18] Keiji Shinzato, Tomohide Shibata, Daisuke Kawahara, Chikara Hashimoto, and Sadao Kurohashi. TSUBAKI: An open search engine infrastructure for developing new information access methodology. In *Proc. the 3rd International Joint Conference on Natural Language Processing (IJCNLP2008)*, pp. 189–196, 2008.
- [19] S. Li, Z. Wang, G. Zhou, and S.Y.M. Lee. Semi-supervised learning for imbalanced sentiment classification. *Proceedings of IJCAI-2011*, 2011.