

# 複数語から成る評価表現のモデル化

高村 大也† 乾 孝司‡ 奥村 学†  
† 東京工業大学 精密工学研究所 ‡ 学術振興会  
{takamura,oku}@pi.titech.ac.jp, tinui@lr.pi.titech.ac.jp

## 1 序論

テキストにおける感情情報処理技術が、産業界を含む多くの場所で注目を集めている。そのような技術は、レビューの解析による新製品のサーベイ、アンケート処理など様々な応用の場を持つ。大抵の応用においては大量のデータを処理するので、感情情報処理の自動化は、包括的な調査のためには必要不可欠である。

テキストの感情情報処理における最も基礎的な技術は、単語の感情極性の獲得であるといえる。ここで感情極性とは、ポジティブ（望ましい）か或はネガティブ（望ましくない）かを表す。例えば、“美しい”はポジティブだが、“汚い”はネガティブである。このタスクについては、いくつかの研究があり良い結果が出ている。次に解くべき問題の一つとして、複数語から成る表現の感情極性をいかにして扱うかという問題が挙げられる。これまでは、単語の感情極性をそのまま複数語表現に適用した研究はあったが、複数語の特性を考慮に入れたモデルは提案されていない。本稿の目的は、複数語から成る評価表現のモデル及びそれに基づいた分類手法を提案することである。複数語評価表現のモデル化は、文や文書など、より大きな言語単位でのモデル化への足掛かりになると期待される。

複数語表現の感情極性は、その構成語の極性の単純な和ではない。例えば、“ノートパソコンが軽い”という表現はポジティブであるが、“軽い”も“ノートパソコン”もそれら自体はポジティブではない。また、極性を反転させる作用を持つ単語も存在する。例えば、“リスクが低い”においては、“リスク”のネガティブ極性が“低い”によって反転させられている。このようなある種の非線形な演算がモデルに取り入れられる必要がある。本稿では、複数語表現の特性を有する最も単純なパターンである、「名詞+形容詞」を取り扱う。

複数語から成る評価表現の特性を捉えるため、隠れ変数モデルを導入する。“感染率”と“死亡率”のように、感情極性という観点から類似している単語（“低い”と結び付いてポジティブになる単語）が同じクラスターに属するようにモデルが構築される。我々のモデルは、単語と感情極性の共起データのみを利用しているという点において、言語非依存であるといえる。さらに我々は、訓練データに出現しなかった単語も取り扱えるように、単語の感情極性のためのスピンモデル (Takamura et al., 2005) を拡張して、本タスクに適用する。

## 2 関連研究

本研究の関連研究として、一般的な単語対の分類問題という側面と、感情極性分類という側面の両方から見ていくことにする。単語対の分類に関しては、Tori-sawa (2001) は、名詞と動詞の対に対する適切な格の付与に、Fujita ら (2004) は言い替え文における誤った格割り当ての抽出に、Probabilistic Latent Semantic Indexing (PLSI) モデル (Hofmann, 2001) を利用した。

単語の感情極性分類についてはいくつかの研究があり、良い成果が出ている (Hatzivassiloglou and McKeown, 1997; Turney and Littman, 2003; Kamps et al., 2004; Takamura et al., 2005; 小林ら, 2001)。しかし、複数語からなる評価表現の特徴を捉えた確率モデルは

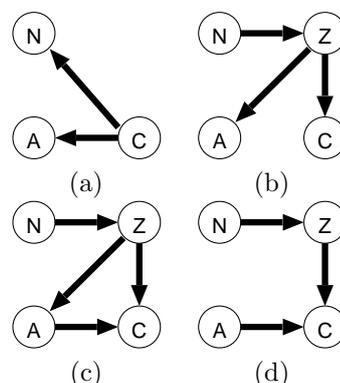


図 1: モデルのグラフィカル表現

未だに提案されていない。

鈴木ら (2006), Turney (2002), Baron ら (2004) は、文脈情報を利用し、複数語評価表現を抽出することを試みた。これらはコーパスに基づいた、いわば表現とそれが出現する外部文脈との関係を利用した手法である。一方、我々の手法は複数語表現の感情極性の内部構造を明らかにしようとするものである。

Inui (2004) は、複数語表現の感情極性分類において、各単語に *plus/minus* のいずれかの値をとる属性を考え、その属性値と構成語の感情極性に基づいた極性決定規則を提案した。例えば、[negative+minus=positive] という規則は“リスク (negative)+低い (minus)”がポジティブであると決定する。Wilson ら (2005) の手法も同様である。本稿で提案する手法は、非常に一般的な視点からは、Inui や Wilson らのアイデアを確率モデルを用いて自動化した手法であると捉えることができる。

**3 複数語評価表現のための隠れ変数モデル**  
隠れ変数モデルを用いて、そのような隠れた意味クラスターを捉え、複数語表現の高精度な分類を実現する。

隠れ変数を導入することにより、直観的には例えば、“リスク”や“感染率”などのようにその程度や量が減少することによりポジティブになるような名詞が同じクラスターに属するといった効果が期待できる。一方、減少を意味するような形容詞あるいは動詞が同じクラスターに属するといった効果も期待できる。

図 1 は、本タスクに関連する様々なモデルのグラフィカル表現である。N, A, Z, C は各々、名詞、形容詞、クラスター、感情極性に対応する。図 1-(a) は、感情極性変数が与えられたときに名詞と形容詞が独立になる、いわばナイーブベイズモデルである。図 1-(b) は、PLSI に感情極性に対応する観測変数を加えたもので、ここでは *3PLSI* モデルと呼ぶ。図 1-(c) は三角形モデル、図 1-(d) は U 字型モデルと呼ぶ。三角形モデルと U 字型モデルにおいては、 $P(c|az)$  を通して、形容詞が直接的に感情極性変数に影響を与える。

### 3.1 モデル推定

三角形モデルを例にとって、推定方法を説明する。D を、名詞  $n$ 、形容詞  $a$ 、その極性  $c$  のタプルの集合で

あるとする：  $D = \{(n_1, a_1, c_1), \dots, (n_{|D|}, a_{|D|}, c_{|D|})\}$ .  
ここで  $c \in \{-1, 0, 1\}$  とし、 $-1$  はネガティブに、 $0$  はニュートラルに、 $1$  はポジティブに対応するものとする。我々の目的は未知のペア  $x$  と  $y$  に対して、その極性  $c$  を予測することである。

図 1-(c) より、ここで、 $n, a, c, z$  の同時確率は、

$$P(nacz) = P(z|n)P(a|z)P(c|az)P(n). \quad (1)$$

と表せると仮定する。ただし、紙面の節約のため、確率変数間のカンマを省いて表記する。

モデル推定には、Expectation-Maximization (EM) アルゴリズム (Dempster et al., 1977) を用いる。Q 関数 (隠れ変数の事後確率に関する対数尤度の期待値) は

$$Q(\theta) = \sum_{nac} f_{nac} \sum_z \bar{P}(z|nac) \log P(nacz|\theta), \quad (2)$$

である。ここで、 $\theta$  はパラメータの集合を表し、 $f_{nac}$  はタプル  $\langle n, a, c \rangle$  のデータ中での頻度を表す。  $\bar{P}$  は、古いパラメータを用いて計算された確率値である。これより、以下の更新式を得る：

$$\bar{P}(z|nac) = \frac{P(z|n)P(a|z)P(c|az)}{\sum_z P(z|n)P(a|z)P(c|az)}, \quad (3)$$

$$P(z|n) = \frac{\sum_{ac} f_{nac} \bar{P}(z|nac)}{\sum_{ac} f_{nac}}, \quad (4)$$

$$P(y|z) = \frac{\sum_{nc} f_{nac} \bar{P}(z|nac)}{\sum_{nac} f_{nac} \bar{P}(z|nac)}, \quad (5)$$

$$P(c|az) = \frac{\sum_n f_{nac} \bar{P}(z|nac)}{\sum_{nc} f_{nac} \bar{P}(z|nac)}. \quad (6)$$

式 (3) は E ステップに、式 (4), (5), (6) は M ステップに対応する。この二つのステップは Q 関数の変化が十分に小さくなるまで交代しつつ繰り返される。

極性が未知の単語ペア  $n, a$  に対し、確率値

$$P(c|na) = \frac{\sum_z P(z|n)P(a|z)P(c|az)}{\sum_{cz} P(z|n)P(a|z)P(c|az)}. \quad (7)$$

を計算し、この値が最大になるような  $c$  を、求める極性の予測値として出力する。

このモデルにおいて  $P(a|z)$  を一定とすることにより、U 字型モデルが得られる。

### 3.2 比較のための他のモデル

図 1-(b) の 3PLSI モデルも比較対照として試す。

隠れ変数モデルに加え、次のような単純な確率モデルを用いたベースライン分類器を用意しておく：

$$P(c|na) \propto P(n|c)P(a|c)P(c). \quad (8)$$

これは、素性が 2 つのナイーブベイズ分類器 (NB) (Mitchell, 1997) と等価であり、図 1-(a) に対応する。パラメータは、ラプラス推定により求める。

結局我々は、NB, 3PLSI モデル、三角形モデル、U 字型モデルの 4 つのモデルを用意したことになる。

### 3.3 ハイパーパラメータの推定

実際の EM の計算では、通常の EM アルゴリズムでなく、tempered EM アルゴリズム (Hofmann, 2001) を用いる。これにより、計算途中の隠れ変数の事後確率値を過信することによるモデル推定の失敗を回避しやすくなる (Hofmann, 2004)。通常の EM アルゴリズムの E ステップに僅かな変更を加えるだけで、tempered EM アルゴリズムが実現できる。例えば、三角形モデルの場合は、

$$\bar{P}(z|nac) = \frac{(P(c|az)P(z|n))^\beta}{\sum_z (P(c|az)P(z|n))^\beta}, \quad (9)$$

となる。ここで、 $\beta$  はハイパーパラメータで、正の値をとる。この値が小さいほど、計算途中の隠れ変数の事後確率値を信用しないことになる。他のモデルに関しても同様に tempered EM アルゴリズムが導出できる。

また、隠れ変数の状態数を  $M$  で表す。結局我々は、 $\beta$  と  $M$  の二つのハイパーパラメータを決定する必要がある。これらの値の決定は、ヘルドアウト法を用いて行う。すなわち、与えられた訓練データの 90% を一時的な訓練データとして学習を行い、残りの 10% を一時的なテストデータとして評価を行う。これを様々な  $\beta$  と  $M$  の組について行い、最も正解率が高いハイパーパラメータの組を選ぶ。選ばれた値を用いて改めて訓練データ全体で学習を行い、確率モデルを求める。

## 4 Potts モデルによる極性推定

上記隠れ変数モデルは、未知の単語対について感情極性を予測するが、単語そのものが訓練データに出現しない場合は適用不可能である。そこで、Potts スピンモデルを導入し、これに対処する。

Takamura ら (2005) はイジングスピンモデルを利用して、単語の感情極性抽出を行った。辞書の見出し語と語釈文に出現する単語は極性が等しくなりやすいと考えられる。まず、そのような単語同士を連結し、単語をノードとする語彙ネットワークを構築した。スピン系の計算手法を用い、隣り合ったノードがなるべく同じ極性を持つように極性を割り当て、高精度の抽出を実現した。我々は、形容詞ごとに一つのスピン系を構築し、それと結び付いた場合の名詞の極性を分類することにより、複数語表現の感情極性を予測する。これにより、訓練データに含まれていない名詞に対しても、辞書に含まれていれば、形容詞と対になったときの感情極性が予測できる。

しかし、イジングスピンは、ポジティブかネガティブの二値分類しかできない。我々は Potts モデルを導入して、ニュートラルも加えた三値分類を可能にする。

ネットワークの構築方法は、Takamura ら (2005) に従う。すなわち、ある単語が別の単語の語釈文に出現したらその単語を連結し、“ない” などの否定語の後に出現した場合は、重みを負とする。彼らはシソーラスやコーパスも利用しているが、今回は辞書のみを利用する。リンクの重み  $W = (w_{ij})$  の決定方法については、Takamura らの論文 (2005) を参照して頂きたい。

ネットワーク全体の状態を表す、エネルギー関数  $H(\beta_p, c)$  を導入する：

$$H(\beta_p, c) = -\beta_p \sum_{ij} w_{ij} \delta(c_i, c_j) + \alpha \sum_{i \in L} -\delta(c_i, a_i) \quad (10)$$

ここで、 $L$  は初期単語集合、 $a_i$  は初期単語  $i$  の感情極性 ( $1, 0$ , または  $-1$ )、 $\alpha$  はラベル付きデータに対する重み (正の定数) を表す。また、 $\delta$  は、二つの引数が等しいときに 1 を、そうでないときに 0 を返す。つまり、 $c_i$  ( $i \in L$ ) が  $a_i$  と異なる場合、ペナルティが与えられる。この  $H(\beta_p, c)$  を用いると、系全体の確率は、

$$P(c) = \exp\{-H(\beta_p, c)\} / Z \quad (11)$$

と表される。しかし、これを厳密に扱うのは計算量的に非常に困難であるので、近似関数  $\rho(c) = \prod_i \rho_i(c_i)$  を導入し、平均場近似を用いる。Nishimori (2001) の手法に従うことにより、各単語  $i$  に対して、感情極性の確率  $\rho_i(c)$  が求められるので、この値が大きい極性を予測するものとする。これらの計算が、各形容詞に対して行われることに注意して頂きたい。 $\beta_p$  の値は、Takamura

表 1: 予測された  $\beta$  及び  $M$  を用いたときの分類正解率

	標準		極性不定	
	正解率	$\beta, M$	正解率	$\beta, M$
NB	73.40	-	65.93	-
3PLSI	67.02	.73, 91.7	60.51	.80, 87.4
三角形	81.39	.60, 174.0	77.95	.60, 191.0
U字型	81.94	.64, 60.0	75.86	.65, 48.3

ら (2005) と同様に、臨界点を用いる。ただし、Potts モデルは多値なので、近似確率の空間平均がいずれかの感情極性クラスに偏り始めた点を臨界点とした。

## 5 実験

### 5.1 実験設定

まず、事例集合について述べる。新聞記事 (1995) から、主語となる名詞とその述語となる形容詞の対を抽出し、各対にポジティブ、ニュートラル、ネガティブのいずれかの感情極性タグを付けた。その結果、12066 事例から成るラベル付き事例集合が得られた。12066 事例中、異なる対は 7416 事例ある。また、そのうち 4459 事例がネガティブ、4252 事例がニュートラル、3355 事例がポジティブである。名詞の種類数は 4770 であり、形容詞の種類数は 384 である。

評価には 10 分割の交差検定を用い、その正解率を算出した。ただし、訓練データとテストデータに同じ単語対が出現しないように分割した。

また、訓練データに出現しなかった形容詞を含む単語対は評価に使用しなかった。名詞が訓練データに出現している 7581 事例を、隠れ変数モデルの評価に用いた。名詞が訓練データに出現していない 4304 事例のうち、名詞が語彙ネットワークのノードとなっているような 2130 事例を、Potts モデルの評価に用いた。

モデルの有効性を明らかにするため、“高い”、“低い”、“大きい”、“小さい”など、極性が一定でないと思われる 17 語の形容詞を含む対を元の事例集合から抜き出し、新しい事例集合を作成した。この新しい事例集合を極性不定形容詞事例集合、元の事例集合を標準事例集合と呼ぶ。極性不定形容詞事例集合は 4787 の異なる対を含み、標準事例集合の部分集合となっている。極性不定形容詞事例集合は、評価データとしてのみ使用した。訓練には、常に標準事例集合を用いた。

$\beta$  の値としては、0.1, 0.2, ..., 1.0 を試した。また、 $M$  の値としては、10, 30, 50, 70, 100, 200, 300, 500 を試した。適切なハイパーパラメータの値を予測する場合は、これらの値の組の中から、3.3 節で述べたヘルドアウト法を用いて最も高い正解率を出す組を選んだ。

### 5.2 実験結果 (隠れ変数モデル)

表 1 に、ヘルドアウト法で決定した  $\beta$  と  $M$  を用いたときの 4 手法の分類正解率を示す。この表からわかるように、三角形モデルと U 字型モデルは他と比較して良い性能を示している。この結果は、隠れ変数を通して複数語表現の感情極性の内部構造を捉えることに成功したことを示唆している。

3PLSI モデルはうまく働かなかった。Hofmann (2004) は、協調フィルタリングには 3PLSI モデルは制限が強すぎる (モデルの自由度が低過ぎる) としており、本タスクにおいても同様のことがいえることが示された。

次に、ハイパーパラメータの値の影響を見る。図 2, 3, 及び 4 は、それぞれ 3PLSI モデル、三角形モデル、U 字型モデルの交差検定された正解率の  $\beta$  に対する変化を、いくつかの  $M$  についてプロットしたものである。つまり、ここではハイパーパラメータの予測は行われていない。図からわかるように、分類性能は  $\beta$  の値に大

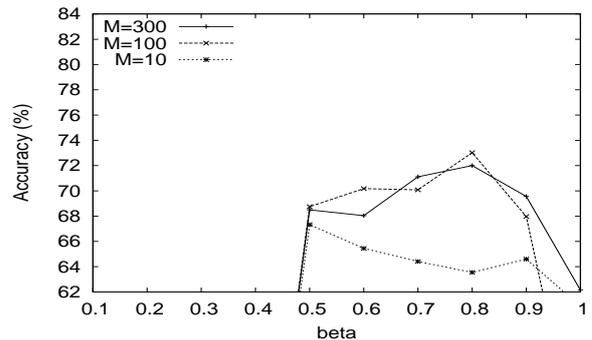


図 2: 3PLSI モデル, 標準事例集合

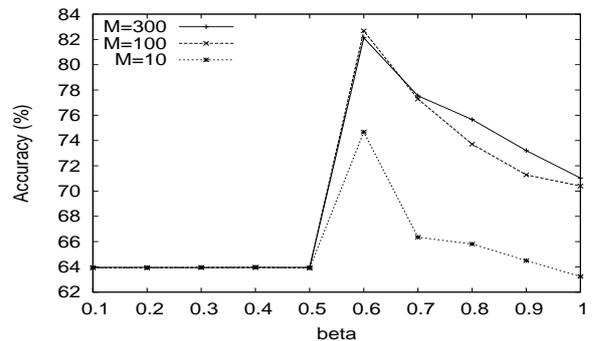


図 3: 三角形モデル, 標準事例集合

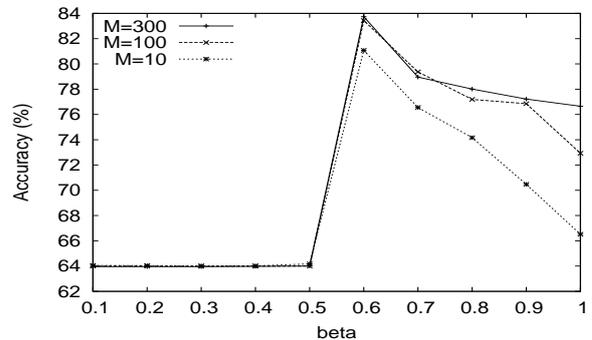


図 4: U 字型モデル, 標準事例集合

きく影響を受けている。つまり、より高精度なハイパーパラメータ予測手法が使用できれば、分類性能はさらに上がるものと思われる。大きめの  $M$  の値 ( $M = 100, M = 300$ ) の方が、小さめの  $M$  の値より良い結果を出している。しかし、これは分類性能と学習時間とのトレードオフであり、 $M$  が大きくなれば学習に多大なコストがかかる。そのような観点から、U 字型モデルは比較的小さな  $M$  でも良い分類性能を示しており、実際の応用に有用であると考えられる。

さらに、全体的なエラーの傾向を見るために、予測されたハイパーパラメータを用いたときの U 字型モデルでの分類結果の分割表を表 2 に示す。この表からわかるように、エラーのほとんどはニュートラル絡みであり、ポジティブとネガティブを間違えて予測した例は全体の 2.26% に過ぎない。つまり、提案モデルは、極性を逆に予測してしまうような大きな間違いをすることは非常に少ないことがわかる。

極性を逆に予測してしまったような少数の例を観察

表 2: U 字型モデルによる分類結果の分割表

		U 字型			合計
		pos	neu	neg	
解答	pos	1856	281	69	2206
	neu	202	2021	394	2617
	neg	102	321	2335	2758
	合計	2160	2623	2798	7581

表 3: Potts モデルによる分類正解率

	正解率
訓練データ未出現全体	74.4
Potts モデル使用のみ	62.9

し、簡単にエラー分析をしてみる。

“食品 + 高い”のように、実際には“食品の価格が高い”ことを意味しているが、“価格”の部分が解釈されていないと思われるものが多く見られた。このような事例に対しては、例えば対象と属性を前処理で正確に特定するような枠組が必要である。

その他、“手詰まり感 + 色濃い”のように、低頻度語に対して判定を誤る例があった。

### 5.3 獲得されたクラスタの例 (隠れ変数モデル)

定性的に結果を見るために、U 字型モデルで得られたいくつかのクラスタ  $z$  に対して、名詞  $x$  を  $P(z|x)$  の値の降順でソートし、上位 50 語に入っている名詞  $x$  のうちデータ中で 3 回以上出現しているようなものを示す。条件は、 $\beta = 0.6$ ,  $M = 60$  である。

- 1 トラブル, 反対意見, 病気, 苦情, 心配, 既往症, 再発
- 2 リスク, 死亡率, 感染率, 発症率
- 3 絆, 意見, 愛着, 意味合い, 憧れ, 意志
- 4 得票, 応募, 話題, 支持者
- 5 弊害, 悪化, ショック, 衝撃, 負担
- 6 悪化, 差別, 負荷, 弊害
- 7 比重, 影響度, 数字, ウェイト, 帰属意識, 波, 呼び声

人間の直観に合ったクラスタが得られている。例えばクラスタ 2 は、“高い”と対になってネガティブに、“低い”と対になってポジティブになる。実際、

$$P(\text{negative}|\text{high, cluster 2}) = 0.995,$$

$$P(\text{positive}|\text{low, cluster 2}) = 0.973.$$

である。単純な共起情報に基づいたクラスタリングでは、クラスタ 2 に“成功率”のような極性が逆になるものが含まれてしまう。極性クラス  $c$  を導入した結果、感情極性判定に合ったクラスタが獲得できた。

### 5.4 実験結果 (Potts モデル)

各形容詞について Potts モデルを利用し、訓練事例を初期単語集合 (4 節の  $L$  に相当) として、近似確率分布を計算した。初期単語集合が一種類の感情極性しか持たない場合は、スピン計算を行わず、その形容詞に対しては常にその極性を予測することにする。結果は表 3 に示す。訓練データに出現しない名詞に関しても、表 3 からわかるように 75% 近い正解率を得られることがわかった。

## 6 結論

複数語から成る評価表現のモデル及びそれに基づいた分類手法を提案した。複数語から成る評価表現の特質を考慮し、モデルに隠れ変数を導入した。実験により、提案した隠れ変数モデルは複数語から成る評価表現分類において高い性能を持つことを示した。

また、Potts モデルにより、訓練データに出現しない単語対についても極性分類が可能であることを示した。両手法の組合せを今後の課題としたい。

## 参考文献

- F. Baron and G. Hirst. 2004. Collocations as cues to semantic orientation. In *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*.
- A. P. Dempster, N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B*, 39(1):1–38.
- A. Fujita, K. Inui, and Y. Matsumoto. 2004. Detection of incorrect case assignments in automatically generated paraphrases of Japanese sentences. *The 1st IJCNLP*, pp. 14–21.
- V. Hatzivassiloglou and K. R. McKeown. 1997. Predicting the semantic orientation of adjectives. *The 35th EACL*, pp. 174–181.
- T. Hofmann. 2001. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42:177–196.
- T. Hofmann. 2004. Latent semantic models for collaborative filtering. *ACM Transactions on Information Systems*, 22:89–115.
- T. Inui. 2004. *Acquiring Causal Knowledge from Text Using Connective Markers*. Ph.D. thesis, Graduate School of Information Science, Nara Institute of Science and Technology.
- J. Kamps, M. Marx, R. J. Mokken, and M. de Rijke. 2004. Using wordnet to measure semantic orientation of adjectives. *The 4th LREC, volume IV*, pp. 1115–1118.
- 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.
- Mainichi. 1995. 毎日新聞 CD-ROM version.
- T. Mitchell. 1997. *Machine Learning*. McGraw Hill.
- H. Nishimori. 2001. *Statistical Physics of Spin Glasses and Information Processing*. Oxford University Press.
- Y. Suzuki, H. Takamura, and M. Okumura. 2006. Application of semi-supervised learning to evaluative expression classification. *The 7th CICLing-06*, pp. 502–513.
- H. Takamura, T. Inui, and M. Okumura. 2005. Extracting semantic orientations of words using spin model. *The 43rd ACL*, pp. 133–140.
- K. Torisawa. 2001. An unsupervised method for canonicalization of Japanese postpositions. *NL-PRS 2001*, pp. 211–218.
- P. D. Turney and M. L. Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 21(4):315–346.
- P. D. Turney. 2002. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *The 40th ACL*, pp. 417–424.
- T. Wilson, J. Wiebe, and P. Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. *HLT/EMNLP*, pp. 347–354.