

## 概要

近年の情緒推定では，機械学習や統計情報を用いる手法が主流である．それらの手法では，情緒推定の手掛かりとして1文単位での表層的な情報を用いることが多い．情緒推定の性能を高めるためには，複数の文に着目することや，表層情報以外の素性を用いることが考えられる．そこで，本研究では，文数および素性の組合せ方を変えながら，機械学習 SVM(Support Vector Machine) を使用して情緒推定を行うことで，性能の違いを実験的に評価することを目的とする．

方法として，複数文の学習データを得るために対話行為および情緒が付与されたコーパスを用意し，そのコーパスより情緒推定 SVM に用いる各種素性を得る．この素性の有効性を調査し，組合せの種類ごとに実験することで性能向上を図る．

結果，複数文の情緒推定が性能向上に繋がることと，素性の組合せによる性能差を確認することができた．

# 目次

<b>第1章</b>	<b>はじめに</b>	<b>1</b>
1.1	複数の文からの情緒推定とは	1
1.2	本研究の目的	1
1.3	本論文の構成	2
<b>第2章</b>	<b>関連研究</b>	<b>3</b>
2.1	情緒推定の関連研究	3
2.1.1	感情の計算モデル	3
2.1.2	統計的機械学習を用いた研究	3
2.1.3	評価極性	4
2.2	本研究で使用するリソース	4
2.2.1	使用するコーパス	4
2.2.2	講談社類語大辞典	7
2.2.3	SVM について	8
<b>第3章</b>	<b>SVM を用いた情緒推定</b>	<b>9</b>
3.1	情緒推定の手順	9
3.2	単語 n-gram 素性	9
3.3	SO-Score のラベル素性	10
3.4	意味属性コード素性	10
3.5	対話行為素性	11
3.6	心的状態ラベル素性	11
<b>第4章</b>	<b>情緒推定の実装</b>	<b>12</b>
4.1	情緒推定の手順	12
4.2	情緒推定器の実装	13
4.3	各素性の実装	13

4.3.1	単語 n-gram 素性 . . . . .	13
4.3.2	SO-Score のラベル素性 . . . . .	13
4.3.3	意味属性コード素性 . . . . .	14
4.3.4	対話行為素性 . . . . .	14
4.3.5	心的状態ラベル素性 . . . . .	14
<b>第 5 章</b>	<b>実験</b>	<b>16</b>
5.1	参照文数の違いによる性能差 . . . . .	16
5.2	素性の種類ごとの性能差 . . . . .	17
5.3	単語 1~5-gram 素性での性能 . . . . .	18
5.4	2 種類の素性組みごとの性能 . . . . .	18
5.5	加算的に素性を用いる場合の性能差 . . . . .	19
5.6	比較実験 . . . . .	19
<b>第 6 章</b>	<b>おわりに</b>	<b>20</b>

# 表 目 次

2.1	対話行為タグの構成要素 . . . . .	6
2.2	情緒タグの件数と割合 . . . . .	7
4.1	対話行為認識精度 . . . . .	14
5.1	参照文数による性能の比較 . . . . .	16
5.2	素性のみ実験結果 . . . . .	17
5.3	単語 1~5-gram をあわせた素性による実験結果 . . . . .	18
5.4	2種類の素性組みによる実験結果 . . . . .	18
5.5	W1-5G+PKR4 への素性追加時の実験結果 . . . . .	19
5.6	コーパス別性能実験結果 . . . . .	19

# 第1章 はじめに

言語の意味理解の一つとして、文章から話者や登場人物の情緒を推定する技術の実現が期待されている。なぜなら、この技術はテキストマイニングへの応用に可能性があるからである。例えば、Web上のblogや掲示板、近年ではツイッターなどに蓄積されたテキストデータから情緒を推定することで、商品や社会事情に対する大衆の気持ちを機械的に知るといったことができるからである。

## 1.1 複数の文からの情緒推定とは

本研究では、テキスト対話からの情緒推定を想定する。これはWebにおける掲示板などの自動分析への応用が考えられるためである。

さて、まずは、情緒推定のために参照する文数を考えてみよう。1文から得られる情報は少ない。たとえば、「うん。」という文だけでは話し手の情緒が推定しがたい。2文からは明らかに情報が増える。たとえば、「遊びに行こう！」という相手からの要求に「うん。」と答えた場合は、この話し手に、《喜び》や《期待》が生じていると推定しても良いだろう。ゆえに相手との発話の対に注目する必要がある。一方、「うん。でも、外は寒いよ。」と、話し手が2文で答えたとすると、加えて、《嫌だ》という情緒も可能性がある。ゆえに話し手の複数の発話文にも注目する必要がある。

次に、文からの情報の種類を考えてみよう。たとえば、「遊び」には、陽気なイメージがある。「遊びに行こう！」-「うん。」では、対話理解を通じて心的状態が生じている。「でも」には逆接的な意味がある。

## 1.2 本研究の目的

近年の情緒推定では、機械学習や統計情報を用いる手法が主流である。それらの手法では、情緒推定の手掛かりとして1文単位での表層的な情報を用いることが多い。情緒推定の性能を高めるためには、複数の文に着目することや、表層情報以外の素性を用い

ることが考えられる。そこで、本研究では、文数および素性の組合せ方を変えながら、機械学習 SVM(Support Vector Machine) を使用して情緒推定を行うことで、性能の違いを実験的に評価することを目的とする。

### 1.3 本論文の構成

本論文の構成は以下の通りである。第2章で関連研究と、本研究で使用するリソースについて述べ、第3章で各素性について述べる。第4章で情緒推定の方法、情緒推定器・各素性の実装について述べ、第5章で実験の方法と各素性を用いた情緒推定の結果・考察を述べる。最後に、第6章で本研究をまとめる。

## 第2章 関連研究

本章は、まず、情緒推定についての関連研究と、本研究で使用するリソースについて紹介する。

### 2.1 情緒推定の関連研究

#### 2.1.1 感情の計算モデル

「人間の心は、ある種の計算装置とみなすことができる」という発想から、感情の生起過程を記号計算の過程とみなし、コンピューター上で実働する記号処理プログラムのように構築するモデルのこと。情緒主の感情生起過程の一つは感情の目標指向性であり、R.Pfeifer[1]は目標達成のために遂行する計画や下位目標の状態（達成、失敗、保留など）をネットワークで表現した。そして、このネットワーク上である条件をみたすような状態が発生した時に感情が生起する。[1]ではこれを「割込み」という機構を用いて感情誘機の契機とした。感情を生起させることが決定されると、次はどのような種類の感情を選択するかを決定する機構が必要である。この選択には、感情の生起した時点での状況を構築している要因が条件として用いられる。

このように、感情の計算モデルでは感情生起までには様々な条件や状態を参照する必要がある、1文のみの文章からこれらを特定するのは困難である。

#### 2.1.2 統計的機械学習を用いた研究

統計的機械学習法を用いた研究として、徳久良子らの研究[2]がある。これは、Webから人が感情を生起する要因を獲得し、感情生起要因コーパスを作成する。そして、このコーパスを学習データとして入力された発話の感情極性を推定、次にその感情極性と大量の感情生起要因を利用して情緒を推定するものである。しかし、これらは1文におけ

るデータを扱っており、深層的な情報が必要となる対話においては有効性が評価されていない。

### 2.1.3 評価極性

評価極性の研究として、[3]らのSO-Scoreというものがある。これは、ある単語  $t$  における評価極性で、その単語がPositiveかNegativeか、コーパス内の共起頻度を利用してどちらの傾向が強いが算出するというものだ。

$$SO-Score(t) = PMI(t, "Positive") - PMI(t, "Negative")$$

$$PMI(a, b) = \log_2(p(a, b)/p(a)p(b))$$

$p(a, b)$  はコーパス内で単語  $a$  と 情緒タグ  $b$  が同時に存在する確率。

$p(a)$  はコーパスで単語  $a$  を含む文が出現する確率。

## 2.2 本研究で使用するリソース

### 2.2.1 使用するコーパス

機械学習には、教師あり機械学習 SVM を用いる。そのため、本研究では、対話行為および情緒タグ付きコーパスを使用する [4], [5]。コーパスの対話文は、漫画ちびまる子ちゃんの1~10巻から収録する。対話行為タグ、情緒タグの種類と説明は後に記載する。漫画の1コマ内で同一人物の全ての発話文を1つの被付与単位（コマ人単位と呼ぶことにする）とし、1つ以上の情緒タグが付与されている。また、前コマ人と現コマ人との発話文のペアをコマ人ペアと呼ぶことにする。

コーパスの規模は、26,527文、17,078コマ人である。対話行為タグ数は26,527個、情緒タグ数は21,998個である。

#### 対話行為タグ

対話行為とは、話者が何かしらの意図を持って発話することであり、その意図することがどのような種類なのかを判断するために用いられるのが対話行為タグである。これは「行為タイプ」、「対象タイプ」、「時制」、「存在性」の4つの組で構成されている。この



組合せには制約があり，98通りとなる．徳久らによって改良が加えられ，極力文脈に依存しない判断でタグが付与できるようになった [6],[7]．次ページの表 2.1 にその構成要素の種類と説明を示す．

表 2.1 対話行為タグの構成要素

項目	要素	説明	
行為タイプ	質問	5W1Hの様に、仮説が述べられていない疑問文	
	確認	Yes/Noで答えられる様に、仮説が述べられている疑問文	
	伝達	事態を述べる文	
	要求	聞き手の行動を定める文	
	はい	「はい」等承諾、肯定を表す文	
	いいえ	「いいえ」等拒否、否定を表す文	
	その他	挨拶や相槌、擬音語等上記以外の文	
対象タイプ	生理	生理	身体の内面や表面での感覚
		欲求	生理や社会に起因して生じる願望
		情緒	喜び、悲しみ等の感覚
		プラン	人物の行動の1つ以上の系列 動物の動作も擬人化してプランとして扱う
		属性	物事に対する評価や性質
		その他	上記以外のもの
		時制	過去
	非過去		対象は過去のものではない
	存在性	否定	対象は想像上であり、否定する
		想像	対象は想像上であり、否定はしない
		実在	対象は実在した、あるいは実在する

以下に、対話文に付与するタグの例を示す。

このお茶新しい。：[伝達, 属性, 非過去, 実在]

弟子入りするよ。：[伝達, プラン, 非過去, 想像]

一緒に帰ろう。：[要求, プラン, 非過去, 想像]

## 情緒タグ

情緒タグは、《喜び》、《悲しみ》、《好ましい》、《嫌だ》、《驚き》、《期待》、《恐れ》、《怒り》、《なし》の9種類となっている。この分類は、本研究で扱う情緒と同様に [8] が参考とされている。タグ自体は、対話行為の理解はもちろん、タグ付与の前に1話分程度の練習を行った二人の熟練者によって付与されている。まず、1つの台詞につき2人の付与者がそれぞれ「一時タグ」としてタグを付与する。次に、2人の一時タグを比較して両者の協議により「正解タグ」として情緒タグが決定される。なお、文脈の影響を受けにくいよう設計されてあるが、完全に文脈に自由なわけではないため、タグ付与の仮定である程度前後文の参照を行っている。コーパス内に付与されている情緒タグの内訳は以下の通りである。

表 2.2 情緒タグの件数と割合

情緒名	件数	割合
《喜び》	4,895	23.4 %
《嫌だ》	3,511	16.8 %
《期待》	2,952	14.1 %
《驚き》	2,347	11.2 %
《恐れ》	2,156	10.3 %
《悲しみ》	1,727	8.3 %
《怒り》	1,522	7.3 %
《なし》	1,101	5.3 %
《好ましい》	680	3.3 %

### 2.2.2 講談社類語大辞典

講談社出版の類語辞典 [9]。約7万9000項目語が収録されており、それぞれが7桁のコードで表されている。桁数によって意味分類の粒度を分類することができ、先頭2桁で大分類、4桁で小分類、5桁で品詞レベル、7桁で単語レベルとなっている。これも詳しくは3章で説明するが、素性の1つとして用いる。

### 2.2.3 SVMについて

Support Vector Machine の略で，教師あり機械学習器の一つである．本研究では，これを情緒を推定するために用いるものとし，コーネル大学の Thorsten Joachims 氏が開発した SVM-Light を使用する．基本的に SVM はクラス数が 2 であるような場合のための分類器であるが，本研究では one-versus-rest 法を用いることで，この問題を解決した．

#### one-versus-rest 法

$n$  個のクラスに対し，あるクラスかそれ以外かを分類する二値分類器を，クラス数分構築する手法のこと．つまり， $n$  クラス分類問題では，入力ベクトル  $x$  の分類スコアが，あるクラスで正值ならば正例，それ以外ならば負例と判断する．

## 第3章 SVMを用いた情緒推定

本章では、情緒推定までの手順と、情緒推定SVMに用いる素性についての説明する。

### 3.1 情緒推定の手順

SVMを用いてコーパスから情緒を推定する。SVMはベクトル表現されたデータを分類することができる。そのため、文や文章を分類するためには、文や文章をベクトル化する必要がある。文や文章のベクトルは多次元であり、各次元には素性を対応させることで情緒推定を行う。素性に用いる文数は実験的に変更するものとし、素性範囲のことを参照文数と呼ぶ。ここで、SVMにおける複数クラスの多値分類では、one-versus-rest法を用いる。

### 3.2 単語 n-gram 素性

形態素解析の結果から、n個の単語を直結したものを単語 n-gram という。以降、実験条件で表す素性表記は、次のように略称を使う。略称：W1G,W2G,...W5G

【例】カンペキに逆らってんな、こいつ…。

W1G: ^ , W1G:カンペキ, W1G:に,..

W2G: ^カンペキ,..

例にある「^」は仮想記号である。これは、形態素解析後に文頭に「^」、文末に「\$」を付与することで、文章のはじまりと終端を明確にし、単語の意味を区別しやすくするためにつけられている。

### 3.3 SO-Scoreのラベル素性

SO-Scoreの値によって「Positive」か「Negative」にラベル化される。Turneyら[3]は形容詞との共起を参照したが、本研究では情緒タグとの共起を用いてSO-Scoreを算出する。なお、共起に用いる情緒は3分類（P/N/その他）にまとめている。略称をSOSとする。

【例】 こうなったら一緒に帰るぞっ。SOS:Positive  
ケチ。SOS:Negative

### 3.4 意味属性コード素性

講談社類語大辞典[9]の7桁のコードを用いる。これを意味属性コードと呼ぶことにする。1文には複数の語があり、それぞれに意味属性コードが付与されている。この文中のすべての意味属性コードを素性にした場合をPKとする。また、述語の意味を近似的に得るために、文末側の意味属性コードを参照することがある。そこで、文末の意味属性コードを使用した場合をPKRとする。そして、これらに参照する桁数をnとして後ろに加え、略称をKR<sub>n</sub>,PKR<sub>n</sub>とする。

【例】 ヘチマの観察はこうやって書いちゃおっと。  
KR2:16, KR4:1600, ..  
PKR2:22, PKR4:2201, ..

### 3.5 対話行為素性

話者の発話が意図する種類を判断するために用いられる。コーパスに付与されている対話行為タグを正解対話行為 (DA) とし、対話行為推定 SVM を用いてコーパスから 1 文ごとに対話行為を推定したものを、推定対話行為 (EDA) とする。対話行為推定 SVM に用いる素性は、単語 1-gram から 5-gram までとする。対話行為タグは組み合わせてできるので、各要素が文に対応するかどうかを対話行為推定 SVM で判定し、分類スコアの高いものを組み合わせて 1 つの対話行為を構成する。略称を DA, EDA とする。なお、推定対話行為 SVM の精度は以下の通りである。

【例】ちっ, まるつきり手え抜いて育てたのに…。

DA : {伝達, プラン, 過去, 実在}

EDA: {伝達, プラン, 過去, 実在}

### 3.6 心的状態ラベル素性

心的状態ラベルは、対話行為の認識結果から求める。これは、対話行為の効果、および、発話対の効果に基づく。たとえば、話し手の〈行為〉が「伝達」ならば〈対象〉となる心的状態が生じる。また、話し相手の〈行為〉が「要求」かつ話し手の〈行為〉が「はい」ならば要求された〈対象〉が心的状態に生じる。図 1 に具体例を示す。

相手：「遊びに行こう！」 [要求, プラン, 非過去, 想像]  
話し手：「うん。 [はい]  
でも、外は寒いよ。」 [伝達, 属性, 非過去, 実在]  
→話し手の信念：プラン：遊び：有り, 属性：寒い：有り

図 3.1 対話理解による心的状態の推定の例

ただし、「遊び」や「寒い」という字面は用いず、これらの意味属性コードおよび SO-Score ラベルを用いてラベル化する。

## 第4章 情緒推定の実装

本章では、実際に SVM を用いて情緒推定を行う際の具体的な手順、必要な項目を説明する。

### 4.1 情緒推定の手順

手順1 コーパスから必要なデータを抽出，各種素性を作成し，DataBase に収納する。

手順2 コーパスから参照文数ごとに ID 番号を付与して IDFiles に収納する。

手順3 DataBase から使用する素性に略称を付与して抽出し，IDFiles とあわせて再び DataBase に収納する。

手順4 手順3 で作成したものと正解クラスをあわせ，IDFiles を参照しながら学習文とテスト文を作成する。

手順5 手順4 の学習文とテスト文を用いて，SVM で情緒の推定を行う。



## 4.2 情緒推定器の実装

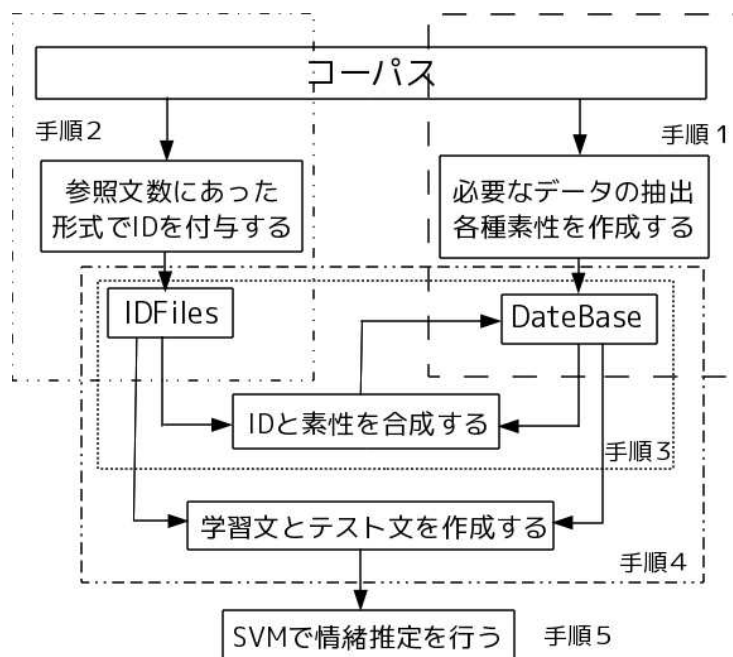


図 4.1 情緒推定器の実装図

## 4.3 各素性の実装

### 4.3.1 単語 n-gram 素性

手順1 コーパスのセリフ抽出し，mecab を使って形態素解析を行う。

手順2 手順1 で得られた結果から，単語 n 個を連結する。

### 4.3.2 SO-Score のラベル素性

手順1 DataBase から単語 n-gram のデータを参照し，それぞれの SO-Score を求める。

手順2 W5G から W1G までの SO-score の絶対値を比べ，その最大値のものが正ならば”Positive”，負ならば”Negative”とする。(単語 n-gram の最大値が4以下の文章の場合は，最大値の n-gram からはじめる)

### 4.3.3 意味属性コード素性

手順1 コーパスのセリフ抽出し，mecabを使って形態素解析を行う。

手順2.a KRnの場合，DataBaseの中から文中にある全ての意味属性コードの部分をも参照する桁数(n)だけ抜き出す。

手順2.b PKRnの場合，DataBaseの中から文中にある一番後ろのみ意味属性コードの部分をも参照する桁数(n)だけ抜き出す。

### 4.3.4 対話行為素性

手順1 DAについては，コーパスからそのまま抽出を行う。

手順2 EDAはコーパスのDAを学習データとし，SVMを用いて対話行為を推定する。

推定対話行為の認識精度は以下ようになった。

表 4.1 対話行為認識精度

対話行為確認	適合率	再現率	F 値 (出力数, 一致数)
	0.6248	0.6248	0.6248 (26,527, 16,575)

### 4.3.5 心的状態ラベル素性

手順1 コーパスから，話者と話し相手のEDAを求める。

手順2 心的状態ラベルを判断する条件に必要な項目を求める。

項目 a < 心的状態 > は対話行為における〈対象〉である。

項目 b.1 < 心的状態 > が「プラン」,「その他」の場合はKR4を求める。

項目 b.2 < 心的状態 > が上記以外の場合は単語 n-garam を抽出し，各 SO-Score を求める。

項目 b.3 n が最大の単語 n-gram から SO-Score の絶対値が一番大きいものを選択し，それが正ならば”Positive”，負ならば”Negative”とする。

項目 c : < 存在性 > は *if - then* ルールによって決まる。

手順 3 22 件の *if - then* から、話者の心的状態ラベルを作成する。

以下に心的状態ラベルに用いる具体的な項目と、22 件の *if - then* ルールを記載する。

心的状態ラベル : < 心的状態 > : < 状態クラス名 > : < 存在性 >

- < 心的状態 > : [生理 | 欲求 | 情緒 | プラン | 属性 | その他]
- < 状態クラス名 > : [ 意味属性コード | SO-Score ]
- < 存在性 > : [有り | なし | 不明]

```
if "C : [伝達, M, __, 実在] : S" then "C : [MS : M : S : 有り]"
if "C : [伝達, M, __, 想像] : S" then "C : [MS : M : S : 有り]", "C : [MS : M : S : なし]"
if "C : [伝達, M, __, 否定] : S" then "C : [MS : M : S : なし]"
if "C : [要求, M, __, __] : S" then "C : [MS : M : S : 有り]", "C : [MS : M : S : なし]"
if "P : [要求, M, __, 実在] : S, C : [はい]" then "C : [MS : M : S : 有り]"
if "P : [要求, M, __, 想像] : S, C : [はい]" then "C : [MS : M : S : 有り]"
if "P : [要求, M, __, 否定] : S, C : [はい]" then "C : [MS : M : S : なし]"
if "P : [要求, M, __, 実在] : S, C : [いいえ]" then "C : [MS : M : S : なし]"
if "P : [要求, M, __, 想像] : S, C : [いいえ]" then "C : [MS : M : S : なし]"
if "P : [要求, M, __, 否定] : S, C : [いいえ]" then "C : [MS : M : S : 有り]"
if "P : [確認, M, __, 実在] : S, C : [はい]" then "C : [MS : M : S : 有り]"
if "P : [確認, M, __, 想像] : S, C : [はい]" then "C : [MS : M : S : 有り]"
if "P : [確認, M, __, 否定] : S, C : [はい]" then "C : [MS : M : S : なし]"
if "P : [確認, M, __, 実在] : S, C : [いいえ]" then "C : [MS : M : S : なし]"
if "P : [確認, M, __, 想像] : S, C : [いいえ]" then "C : [MS : M : S : なし]"
if "P : [確認, M, __, 否定] : S, C : [いいえ]" then "C : [MS : M : S : 有り]"
if "P : [質問, M, __, 実在] : S, C : [はい]" then "C : [MS : M : S : 有り]"
if "P : [質問, M, __, 想像] : S, C : [はい]" then "C : [MS : M : S : 有り]"
if "P : [質問, M, __, 否定] : S, C : [はい]" then "C : [MS : M : S : なし]"
if "P : [質問, M, __, 実在] : S, C : [いいえ]" then "C : [MS : M : S : なし]"
if "P : [質問, M, __, 想像] : S, C : [いいえ]" then "C : [MS : M : S : なし]"
if "P : [質問, M, __, 否定] : S, C : [いいえ]" then "C : [MS : M : S : 有り]"
```

図 4.2 心的状態の *if - then* ルール

P は話者の話相手、C は話者、M は心的状態、S は状態クラス名となっている。  
心的状態ラベルの略称を MS とする。

## 第5章 実験

4分割クロスバリデーション法により情緒推定の動作を確認する。評価には適合率（一致率/出力数），再現率（一致率/正解数），F値（適合率と再現率の調和平均）を用いる。正解数は，すべて21,998個である。

### 5.1 参照文数の違いによる性能差

まず，参照文数の違いによりどの程度性能に差が出るのか調査する。情緒推定に用いる素性は，単語1～5-gramのみで実験を行った。

表 5.1 参照文数による性能の比較

参照文数	適合率	再現率	F値	(出力数, 一致数)
文単位情緒推定	0.4219	0.3302	0.3705	(26,841, 11,325)
コマン単位情緒推定	0.4785	0.3785	0.4227	(17,404, 8,327)
コマンペア単位情緒推定	0.4697	0.3686	0.4131	(17,265, 8,109)

1文単位での情緒推定よりも，コマン単位における複数文の情緒推定のほうが高い精度が得られた。対話であるコマンペア単位での情緒推定においても，コマン単位には及ばないものの，文単位情緒推定よりも高い精度となっている。この結果より，以降の実験において参照文数についてはコマン単位を用いるものとする。

## 5.2 素性の種類ごとの性能差

情緒推定に用いる，素性の違いによる性能差を調査する．

表 5.2 素性のみ実験結果

手法	適合率	再現率	F 値	(出力数, 一致数)
W1G	0.4444	0.3470	0.3897	(17,176, 7,633)
W2G	0.4592	0.3641	0.4061	(17,441, 8,009)
W3G	0.4476	0.3530	0.3947	(17,348, 7,765)
W4G	0.4092	0.3203	0.3593	(17,220, 7,046)
W5G	0.3415	0.2656	0.2988	(17,110, 5,843)
SOS	0.1295	0.1221	0.1257	(20,737, 2,686)
KR2	0.1829	0.1421	0.1599	(17,090, 3,126)
KR4	0.2933	0.2279	0.2565	(17,093, 5,014)
KR5	0.3050	0.2388	0.2679	(17,225, 5,253)
PKR2	0.2173	0.1687	0.1899	(17,081, 3,711)
PKR4	0.2833	0.2216	0.2486	(17,207, 4,874)
PKR5	0.2306	0.1824	0.2037	(17,401, 4,013)
DA	0.1532	0.1190	0.1339	(17,079, 2,617)
EDA	0.1817	0.1411	0.1589	(17,080, 3,104)
MS	0.2045	0.1639	0.1819	(17,632, 3,605)
SP	0.1967	0.1533	0.1723	(17,145, 3,372)

単語を n-gram を含まない素性の場合，性能が低い．情緒推定の内訳を見てみると，単語 n-gram 以外の素性の場合には情緒の内どれか 1 つに出力が極端に偏る傾向が確認できた．正解情緒タグの割合と無関係に「期待」や「悲しみ」等，完全にランダムで出力の偏りは半分近くにもおよぶ．一方，単語 n-gram の場合は「喜び」，「嫌だ」，「なし」の順に出力が多く表れ，正解情緒の割合とある程度似た傾向を示す結果となった．これは，ラベルやコードのみでは抽象度が大きすぎるため，正確な推定ができなかったと考える．

### 5.3 単語 1～5-gram 素性での性能

単語 n-gram の  $n = 1 \sim 5$  と、あわせた素性を用いる場合の性能を調査した。表 6.2 のいずれよりも、性能が高いという結果が得られた。

表 5.3 単語 1～5-gram をあわせた素性による実験結果

素性	適合率	再現率	F 値	(出力数, 一致数)
W1-5G	0.4785	0.3785	0.4227	(17,404, 8,327)

### 5.4 2種類の素性組みごとの性能

前節までで性能が最もよかった W1-5G をベースに、他の素性を組み合わせて性能差を調査する。

表 5.4 2種類の素性組みによる実験結果

素性	適合率	再現率	F 値	(出力数, 一致数)
W1-5G+SOS	0.4827	0.3819	0.4264	(17,403, 8,401)
W1-5G+KR2	0.4791	0.3794	0.4234	(17,418, 8,345)
W1-5G+KR4	0.4802	0.3809	0.4248	(17,449, 8,379)
W1-5G+KR5	0.4784	0.3794	0.4232	(17,445, 8,346)
W1-5G+PKR2	0.4813	0.3811	0.4254	(17,417, 8,383)
W1-5G+PKR4	0.4834	0.3829	0.4273	(17,424, 8,422)
W1-5G+PKR5	0.4820	0.3817	0.4260	(17,418, 8,396)
W1-5G+DA	0.4830	0.3825	0.4269	(17,422, 8,414)
W1-5G+EDA	0.4778	0.3784	0.4223	(17,421, 8,324)
W1-5G+MS	0.4818	0.3814	0.4257	(17,411, 8,389)

W1-5G+EDA 以外は性能が向上した。EDA を求める対話行為認識の性能は予備実験では F 値で 0.62 程度であった。その正確性に欠けることが EDA の不調の原因であると考えられる。これは、正解対話行為タグである DA との組合せの場合は性能が向上していることから分かる。一方、素性を単独で用いる際、PKR4 よりも KR5 のほうが性能はよかったが、W1-5G との組合せでは PKR4 のほうがよかった。これは、文末の述語が文全体の情緒に大きく関係していること、および、単語 n-gram よりも PKR4 のほうが高い抽象度であることが、再現率の向上した理由と考える。

## 5.5 加算的に素性を用いる場合の性能差

前節までで最も性能の良かった W1-5G+PKR4 をベースに加算的に素性を増やし、性能の調査を行う。

表 5.5 W1-5G+PKR4 への素性追加時の実験結果

ベースへの追加素性	適合率	再現率	F 値 (出力数, 一致数)
W1-5G+PKR4+SOS	0.4894	0.3876	0.4326 (17,423, 8,526)
W1-5G+PKR4+SOS+MS	0.4901	0.3886	0.4336 (17,444, 8,549)
W1-5G+PKR4+SOS+MS+EDA	0.4898	0.3884	0.4333 (17,445, 8,545)
W1-5G+PKR4+SOS+MS+DA	0.4919	0.3904	0.4353 (17,456, 8,587)

W1-5G+SOS+PKR4+MS+DA の性能が最もよかった。しかし、6.4 同様に EDA を加えると性能が低下した。DA を加えた場合は性能が向上しているのので、原因についても同様と考える。他のコーパスを用いる場合、対話行為素性については EDA を用いることしかできないので、対話行為素性を加えるかどうかは考慮が必要である。ただ、MS は EDA 素性にルールを適用して求めたものであるが、取捨選択が行われたことが効果をもたらしたと考える。取捨選択は、表 6.4 で KR<sub>n</sub> よりも PKR<sub>n</sub> のほうが高性能であったことから有効と考える。

## 5.6 比較実験

比較実験として、対話ではなく blog 文を用いて複数文における情緒推定の調査する。blog コーパスには情緒タグの付与された 1000 文、素性に使用するのは単語 1~5-gram, 参照する文数は 3 文とする。

表 5.6 コーパス別性能実験結果

コーパス	適合率	再現率	F 値 (出力数, 一致数)
漫画コーパス	0.3883	0.2988	0.3377 (631, 245)
blog コーパス	0.6262	0.5854	0.6051 (998, 625)

性能は漫画コーパスよりも高く見えるが、内約は全てが「なし」に偏っており、実質のところの blog コーパスでの情緒推定は困難であった。漫画と比べ、blog は事細かな物事の流れが書かれていないことや、話の場面が唐突に切り替わることがある為と思われる。

## 第6章 おわりに

1文のみならず，複数の文も考慮して得られる素性を用いた情緒推定をSVMにより実現した．素性については単語が最も性能に関与していたが，文章から得られる素性を適切に加えることによって，さらに性能を向上できることを確認した．今後の課題として，対話文以外でも運用可能である素性の考案，参照する文で得られる情報から状況の流れをさらに把握できる方法を考案する必要がある．



# 謝辞

本研究，本論文作成を勧めるにあたり，多大なる検討と数々の御指導を頂きました鳥取大学工学部知能情報工学科計算機 C 研究室の，村田真樹教授，村上仁一准教授に心からお礼を申し上げます。また，研究の終始に渡り様々な御指導を頂いた徳久雅人講師に大変感謝致します。

多忙な中，修士論文の副査を鳥取大学工学部知能情報工学科知能 B 研究室木村周平准教授に行って頂きましたことを，心よりお礼申し上げます。

また，その他様々な御協力をしてくださった，計算機 C 研究室のみなさまに，深く感謝致します。

## 参考文献

- [1] Pfeifer, R. (1982): Cognition and emotion: An information processing approach. CIP working paper, 436, Carnegie-Mellon University.
- [2] 徳久良子, 乾健太郎, 松本裕治: Web から獲得した感情生起要因コーパスに基づく感情推定, 情報処理学会論文誌, Vol.50, No.4, pp.1365-1374, 2009.
- [3] Turney,P.D.:Thumbs Up or Thumbs Down? Semantic Orientation Applied to Un-supervised Classification of Reviews, ACL, pp.417-424, 2002.
- [4] 徳久雅人, 村上仁一, 池原悟: 漫画における表情に着目した情緒タグ付きテキスト対話コーパスの構築, 自然言語処理, Vol.14, No.3, pp.192-217, 2007.
- [5] 徳久雅人, 前田浩佑, 村上仁一, 池原悟: 対話行為と情緒を解析するための文末表現パターンの作成, 信学技報, NLC2007-95, pp45-50, 2008.
- [6] 松浦大樹, 徳久雅人, 村上仁一, 池原悟: “心的状態の解析のための対話行ためタグ付きテキスト対話コーパスの試作,” 電子情報通信学会技術研究報告, 思考と言語,
- [7] 徳久雅人, 前田浩佑, 村上仁一, 池原悟: “心的状態を表す対話行ためタグ付きテキスト対話コーパスの構築,” 電子情報通信学会技術研究報告, 思考と言語, TL2007-45, pp.25-30, 2007.
- [8] Plutchik.R.: “Emotions and Life: Perspectives from Psychology, Biology, and Evolution,” American Psychological Association, 2002.
- [9] 柴田武, 山田進: 類語大辞典, 講談社, 2002.