

## 概要

人は日々感動を求め、感動によって動かされる生き物である。また、人間の英知は文という形で記録、保存され、後世に受け継がれていく。本研究では「感動」と「文」に重きを置き、感動を与える文か否かの自動判定に関する研究を行うことによって、感動を与える文の作成支援を行った。自動判定は教師あり機械学習である SVM や、パターンマッチングによって行う。また、感動を与える文で多く使われる単語を収集することで、感動を与える文の言語的特徴を明らかにする。

その結果、感動を与える文に多く出現する単語として、「人生」「人々」「幸福」「友情」「青春」「恋愛」などが得られた。これらを用いた文は感動的な文を作成する際に役立つものと考えられる。

また、機械学習により大規模ウェブデータから感動を与える文を適合率 0.4 で抽出できることを示した。

# 目次

第1章	はじめに	1
第2章	感動を与える文の判定方法	2
2.1	問題設定	2
2.2	実験環境	2
2.3	Support Vector Machine への適用	2
2.3.1	Support Vector Machine	2
2.3.2	素性	4
2.4	実験方法	5
2.4.1	人手による感動を与える文の収集	5
2.4.2	教師あり機械学習を利用した感動を与える文の収集	5
2.4.3	カッパ係数による正例と負例の評価	7
第3章	実験	8
3.1	感動を与える文の収集	8
3.1.1	人手による感動を与える文の収集	8
3.1.2	教師あり機械学習による感動を与える文の収集	10
3.1.3	カッパ係数による正例と負例の評価	11
3.2	収集した感動を与える文の分析	12
3.3	感動を与える文の自動抽出性能	14
第4章	関連研究	16
4.1	コンピュータによる小論文の自動採点システム Jess	16
4.2	機械学習を利用した研究	16
4.2.1	情報の重用度を定める要因の抽出・分析の自動推定	16
4.2.2	ユーザ個人の興味の影響を考慮した情報の重用度を定める要因の抽出・分析	17

4.2.3	小説テキストを対象としたジャンル推定と人物抽出 . . . . .	17
4.3	感情に関わる研究 . . . . .	17
4.3.1	語の意味情報を考慮した感情推定アルゴリズム . . . . .	17
4.3.2	登場人物の感情表現に着目した物語要約 . . . . .	17
4.3.3	web から獲得した感情生起要因コーパス . . . . .	18
4.3.4	リビング環境において "深い感動 "を喚起させる「場」の実現方法 の検討 . . . . .	18
4.3.5	音楽聴取における "感動 "の評価要因ー感動の種類と音楽の感情価 の関係 . . . . .	18
第5章	今後の課題	19
第6章	おわりに	20

# 表 目 次

2.1	素性例 . . . . .	4
2.2	カッパ係数評価指標 . . . . .	7
3.1	追加された正例と負例の個数 . . . . .	10
3.2	正例に出現する割合の高い単語の例 . . . . .	14
3.3	種々の手法の抽出性能 . . . . .	15
6.1	正例に出現する割合が1かつ頻度が5以上の単語 . . . . .	24
6.2	正例に出現する割合が0.8以上の単語 . . . . .	25
6.3	正例に出現する割合が0.8以上の単語 . . . . .	26
6.4	正例に出現する割合が0.8の単語 . . . . .	27

# 第1章 はじめに

人は日々感動を求め、感動によって動かされる生き物である。また、人間の英知は文という形で記録、保存され、後世に受け継がれていく。本研究では感動を与える文を作成する際に役立つ知見を収集することを目的とし、本研究の成果は名文作成支援システムの構築に役立つ。名文の作成支援は、感動を与える演説をする必要のある政治家の原稿作成の際にも有用である。石岡ら [1] は毎日新聞の社説およびコラムを学習し、採点を行う日本語小論文の自動採点システムを構築したが、感動に対する評価は行われていない。そこで本研究では「感動」と「文」に重きを置き、読み手の感情を推定し感動を与える文に関する研究を行う。具体的には以下のことを行う。

## 1. 感動を与える文の収集

まず、感動を与える文を人手で収集する。感動を与えない文も人手で収集する。これらを用いて教師あり機械学習により、ウェブ文書からさらに感動を与える文を収集する。

## 2. 収集した感動を与える文の分析

収集した感動を与える文を分析する。感動を与える文で多く使われる単語を収集することで、感動を与える文の言語的特徴を明らかにする。

読み手や聴き手となる人に感動を与えることで、良い印象を与える可能性が高まる。本研究は名文や感動を引き起こす文における計算言語学的特徴の追求を目指す。

本論文の構成は以下の通りである。第2章ではこれまでの関連する研究を説明する。第3章では本研究における、感動を与える文の判定方法と判定に利用する技術を説明する。第4章では、感動を与える文の自動判定とその評価を行う。第5章では、考察を行い効果的な利用法を考察する。第6章ではまとめを行う。

## 第2章 感動を与える文の判定方法

本章では，教師あり機械学習を利用した感動を与える文の判定の手法，また判定に利用する技術について説明を行う．

### 2.1 問題設定

本研究では，まず Google 検索等で感動を与える文と感動を与えない文を収集する．その後でこれらのデータを学習データとして教師あり機械学習を利用した自動判定を行う．この自動判定を用いてウェブコーパス [2] からさらに，感動を与える文と感動を与えない文を収集する．

### 2.2 実験環境

実験には，Support Vector Machine (以下 SVM) を実装している TinySVM[3]，形態素解析を行う ChaSen[4] を使用した．

### 2.3 Support Vector Machine への適用

#### 2.3.1 Support Vector Machine

サポートベクトルマシン法は，空間を超平面で分割することにより2つの分類からなるデータを分類する手法である．このとき，2つの分類が正例と負例からなるものとするとき，学習データにおける正例と負例のマージン（間隔）を大きくとるほど分類器の誤りが減少するという考えから，このマージンを最大にする超平面を求めそれを用いて分類を行う．一般的に上記の方法の他に「ソフトマージン」と呼ばれる学習データにおいてマージンの内部領域に少数の事例が含まれてもよいとする手法の拡張や，線形分離が不可能な問題に対応するために，超平面の線形の部分を非線型にする拡張（カーネル関数の導入）がなされたものが用いられる．この拡張された方法は，以下の識別関数を用い

て分類することと等価であり，その識別関数の出力値が正か負かによって二つの分類を判別することが可能である．

$$\begin{aligned}
 f(\mathbf{x}) &= \operatorname{sgn} \left( \sum_{i=1}^l \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right) \\
 b &= -\frac{\max_{i, y_i=-1} b_i + \min_{i, y_i=1} b_i}{2} \\
 b_i &= \sum_{j=1}^l \alpha_j y_j K(\mathbf{x}_j, \mathbf{x}_i)
 \end{aligned} \tag{2.1}$$

ただし， $\mathbf{x}$  は識別したい事例の文脈 (素性の集合) を， $\mathbf{x}_i$  と  $y_i (i = 1, \dots, l, y_i \in \{1, -1\})$  は学習データの文脈と分類先を意味し，関数  $\operatorname{sgn}$  は，

$$\begin{aligned}
 \operatorname{sgn}(x) &= 1 \quad (x \geq 0) \\
 &= -1 \quad (\text{otherwise})
 \end{aligned} \tag{2.2}$$

であり，また，各  $\alpha_i$  は式 (2.4) と式 (2.5) の制約のもと式 (2.3) の  $L(\alpha)$  を最大にする場合のものである．

$$L(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \tag{2.3}$$

$$0 \leq \alpha_i \leq C \quad (i = 1, \dots, l) \tag{2.4}$$

$$\sum_{i=1}^l \alpha_i y_i = 0 \tag{2.5}$$

また，関数  $K$  はカーネル関数と呼ばれ，様々なものが用いられるが本稿では以下の多項式のものを用いる．

$$K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^d \tag{2.6}$$

$C, d$  は実験的に設定される定数である．本論文ではすべての実験を通して  $C$  を 1 に  $d$  を 2 に固定した．ここで， $\alpha_i > 0$  となる  $\mathbf{x}_i$  は，サポートベクトルと呼ばれ，通常，式 (2.1) の和をとっている部分はこの事例のみを用いて計算される．つまり，実際の解析には学習データのうちサポートベクトルと呼ばれる事例のみしか用いられない．

## 2.3.2 素性

本節では、素性 (解析に用いる情報) について説明する。本研究では素性に、文章中の名詞、動詞、形容詞、形容動詞、連体詞、副詞、接続詞、感動詞の単語を用いる。各素性情報は、入力文章に対し ChaSen による形態素解析を行い、出力された各品詞を抽出し素性として付与する。表 2.1 に素性を付与した例を示す。表において「名詞:X」は X という名詞が文章中出现したことを意味する素性である。

表 2.1: 素性例

本文	素性
「毎日が未来」	名詞 名詞:毎日 助詞 助詞:が 名詞:未来
不幸な人は 希望をもて、 幸福な人は 用心せよ。	名詞 名詞:不幸 動詞 動詞:不幸 形容動 形容動詞:不幸 助動詞 助動詞:だ 名詞:人 助詞 助詞:は 名詞:希望 助詞:を 動詞:もつ 名詞:幸福 動詞:幸福 形容動詞:幸福 名詞:用心 動詞:する
このように あげられない写真も多く、 気持ちの整理もダメで、 どうしたらいいのか	連体詞 連体詞:この 名詞 名詞:よう 助動詞 助動詞:よう 助詞 助詞:に 副詞 副詞:に 動詞:あげる 動詞:られる 助動詞:ない 名詞:写真 助詞:も 形容詞 形容詞:多い 名詞:気持ち 助詞:の 名詞:整理 名詞:ダメ 動詞:ダメ 形容動詞 形容動詞:ダメ 助動詞:だ 副詞:どう 動詞:する 助動詞:た 形容詞:いい 名詞:の 助詞:か



## 2.4 実験方法

本研究で使用する学習データを作成するために、感動を与える文を収集する方法を示す。まず人手で感動を与える文と感動を与えない文を収集する。この方法を 2.4.1 節で述べる。

その後、これらのデータを学習データとして教師あり機械学習を行う。この結果からさらに感動を与える文と感動を与えない文を収集する。この方法を 2.4.2 節で述べる。

本論文では感動を与える文を正例、感動を与えない文を負例と呼ぶ。

### 2.4.1 人手による感動を与える文の収集

Google 検索にて「という言葉に感動した」という検索ワードで得た文を正例の候補、「という文」「という言葉」という検索ワードで得た文を負例の候補として取り出す。各候補文を人手で判定して正しい正例、負例を得る。

さらに、Google 検索で「名言集」を検索し、そこから正例の候補を、Yahoo のニュース記事等から負例の候補を集める。これらの候補文も人手で判定し正しい正例と負例を得る。

正例と負例の判断基準としては、「感動した」と書いてあるものと「名言」を正例として、感嘆符等が付いておらず、客観的事実を述べているものを負例とした。

### 2.4.2 教師あり機械学習を利用した感動を与える文の収集

2.4.1 節で得た、正例と負例を学習データとした教師あり機械学習を行う。ウェブコーパス [2] を教師あり機械学習で、感動する文かいないかを判定することで、ウェブコーパス [2] から感動する文を収集する。

具体的な手順は以下のとおりである。

1. 2.4.1 節の方法で得た正例と負例を学習データとする。
2. 学習データを用いて機械学習を行う。学習結果を利用して新しいウェブデータ 1 万文について正例、負例の判定を行う。正例とされた事例を人手でチェックして正しい正例と負例を新たに作成し学習データに追加する。
3. 2 を 10 回繰り返す。

機械学習結果の出力においての正例と負例の判定基準は以下の通りである．

何らかの感想を得ることの出来た文は正例，客観的事実しか書かれていないものは負例とする．

### 2.4.3 カッパ係数による正例と負例の評価

カッパ係数とは，ある現象を2人の観察者が観察した場合の結果がどの程度一致しているかを表す統計量である．

カッパ係数は以下の数式で求められる．

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (2.7)$$

$k$  はカッパ係数， $Pr(a)$  は見掛け上的一致率で， $Pr(e)$  は偶然的一致率を示す．

表 2.2: カッパ係数評価指標

値	評価
0.00-0.20	わずかに一致
0.21-0.40	一致
0.41-0.60	中程度的一致
0.61-0.80	かなりの一致
0.80-1.00	ほぼ完全な一致

## 第3章 実験

### 3.1 感動を与える文の収集

#### 3.1.1 人手による感動を与える文の収集

前章にて説明した方法を用いて、人手による感動を与える文の収集を行った。以下に収集した事例の例を示す。

1. 「という言葉に感動した」がつく例文
  - ・「毎日が未来」という言葉に感動した・「道は人と人との架け橋です」という言葉に感動した
  - ・「成功する確率が低いのは当然のことだが、確率が低いからこそその成功に価値がある」
2. 「名言集」から得た例文
  - ・山中の賊を破るのは易く、心中の賊を破るのは難し。
  - ・自分の全生命を女の愛というカードに賭けた男が、このカードが殺された時、がっくりとなって何事も手につかないほど安心してしまうようなら、そんな人間は男ではなく、ただのオスである
  - ・我々は翼が欲しいという欲望を持っている。にもかかわらず結局は空を飛ぶことはできない。要するに、我々は幸せなのだ。さもなければ、空気はやがて吸うに堪えなくなるに違いない。
3. 「という文」がつく例文
  - ・読み取り専用という文
  - ・このようなかたちでよろしいでしょうかという文
  - ・彼女は、スポーツをするのが好きですという文
4. 「という言葉」がつく例文
  - ・「被曝」という言葉

- ・「させていただく」という言葉
- ・「それはいい考えですね。」という言葉

正例と負例の例を以下に示す。

1. 正例の例

- ・残りの人生、世のため人のため働きます
- ・他人のためではなく、自分のために残された力でベストを尽くす
- ・がんばらないけど、諦めない

2. 負例の例

- ・フォルダにコピーする
- ・象は鼻が長い
- ・雪は白い。

この人手に基く収集により，正例 1,018 個，負例 406 個を得た。

### 3.1.2 教師あり機械学習による感動を与える文の収集

2.4.1 節の方法で得た正例 1,018 個，負例 406 個を学習データとし，2.4.2 節で述べた方法を用いて感動を与える文の収集を行う．

何らかの感想を得ることのできた文は正例，客観的事実しか書かれていないものは負例とする．正例と負例の例を以下に示す．

#### 1. 正例の例

- ・ 大学に合格した
- ・ 明日は映画の公開日だ
- ・ 私にはまだ父と母の親であるおじいちゃんとおばあちゃんが四人とも元気に過しています

#### 2. 負例の例

- ・ 大学がある
- ・ プラスチックは燃えないゴミだ
- ・ 図を挿入することができる

機械学習の判定を 10 回繰り返し，得られた正例と負例の個数を表 3.1 に示す．

表 3.1: 追加された正例と負例の個数

	正例	負例
1 回目	111	2108
2 回目	30	164
3 回目	15	137
4 回目	8	119
5 回目	8	111
6 回目	14	90
7 回目	19	96
8 回目	28	57
9 回目	23	65
10 回目	19	59
合計	275	3006

人手で収集したものとあわせて，正例が 1,293 個，負例が 3,412 個となった．

上記では 10 回しか繰り返さなかったが，さらに繰り返すことでより多くの正例と負例を収集できる．

### 3.1.3 カッパ係数による正例と負例の評価

カッパ係数を用いて，収集した正例と負例の一致率を算出した．カッパ係数の算出は 2.4.3 節で述べた方法を用いる．

正例と負例の評価は被験者 A(1 名) が行った．被験者 A(1 名) の評価した正例 20 個と負例 20 個を，被験者 A とは別の被験者 3 名により，評価させた．

被験者 3 名の多数決の結果と被験者 A の判定結果を比較すると，カッパ係数 0.58 を得た．中程度の一致であった．

## 3.2 収集した感動を与える文の分析

3.1.2 節で収集した 1,293 個の正例と 3,412 個の負例を用いて、感動を与える文の分析を行った。ここでは単語に基づく分析を行った。

正例と負例から単語を取り出し、各単語について正例または負例に出現する頻度、正例に出現する割合を求めた。正例に出現する割合が 0.8 より大きく、出現頻度が 5 以上である単語を、正例に出現しやすい単語として取り出した。取り出した単語は合計で 110 単語で、その一部を表 3.2 に示し、残りの単語は付録に付す。

「人生」「人々」「幸福」「友情」「青春」「恋愛」などの単語を得られた。それぞれの例文を以下に示す。

1. 「人生」を含む文の例
  - ・人生は道路のようなものだ。一番の近道は、たいてい一番悪い道だ。
  - ・介護職はおもしろい。死と向き合え、人生のことをいろいろと考えられる。おじいちゃんやおばあちゃん、障害を持った子も持っていない子も、誰にでも居場所を作ってあげていきたい
2. 「人々」を含む文の例
  - ・たいていの人々は、運命に過度の要求をすることによって、自ら不満の種をつくっている。
  - ・人々は悲しみを分かち合ってくれる友達さえいれば、悲しみを和らげられる。
3. 「幸福」を含む文の例
  - ・不幸な人は希望をもて。幸福な人は用心せよ。
  - ・幸福は小鳥のようにつかまえておくがいい。できるだけそっと、ゆるやかに。小鳥は自分が自由だと思い込んでさえいれば、喜んでお前の手の中にとどまっているだろう。
4. 「友情」を含む文の例
  - ・友情は人生の酒である。
  - ・友情のための最大の努力は、友人に我々の欠点を見せることではない。彼に彼の欠点を悟らせることだ。
5. 「青春」を含む文の例
  - ・青春は短い。宝石の如くにしてそれを惜しめ。



・青春とは、奇妙なものだ。外部は赤く輝いているが、内部ではなにも感じられないのだ。

6. 「恋愛」を含む文の例

- ・恋愛は人情の永久的な音楽であり、青年には希望を、老年には後光を与える。
- ・恋愛とは美しい少女に出会い、そしてその少女が鱈のように見えるのに気がつくまでの中間にある甘美な休憩時間です。

これらの単語を用いた文は感動を与える文になりやすいと考えられる。

表 3.2: 正例に出現する割合の高い単語の例

単語	正例に出現する割合	正例の頻度	負例の頻度
幸福	1.00	83	0
友情	1.00	29	0
青春	1.00	18	0
悲しみ	1.00	12	0
存在	1.00	10	0
...	...	...	...
我々	0.97	37	1
不幸	0.97	32	1
愛さ	0.96	23	1
恋愛	0.96	44	2
恋	0.95	122	7
孤独	0.94	32	2
この世	0.94	16	1
愛し	0.94	31	2
愛する	0.94	30	2
あらゆる	0.93	14	1
お前	0.93	13	1
瞬間	0.92	11	1
人生	0.91	145	14
未来	0.91	20	2
幸せ	0.91	20	2
喜び	0.91	10	1
女	0.91	115	12
運命	0.90	19	2
死ぬ	0.90	37	4
知る	0.90	9	1
...	...	...	...
人々	0.81	17	4
感動	0.80	8	2

### 3.3 感動を与える文の自動抽出性能

本研究の技術は、感動を与える文を自動抽出することに役立つ。本節では、感動を与える文を自動抽出する手法の性能を評価する。

表 3.3: 種々の手法の抽出性能

手法	適合率	再現率	F 値
ML 0 回目	0.06	0.25	0.10
ML 1 回目	0.26	0.08	0.12
ML 2 回目	0.29	0.07	0.11
ML 5 回目	0.31	0.05	0.09
ML 10 回目	0.40	0.05	0.09
ベースライン	0.07	1.00	0.12
パターン 1	0.11	0.08	0.09
パターン 2	1.00	0.002	0.003

評価結果を表 3.3 に示す。評価データはウェブコーパスの新たな 1 万文とし、各手法で正例とした事例からランダムに抽出した 100 個の事例を人手で評価し（ベースラインのみ 200 個の事例を人手で評価）、その結果から近似的に適合率、再現率、F 値を算出した。ベースラインは、すべてを正例と判断する手法であり、この手法で検出した正例の個数から、再現率の分母を推定している。

「ML  $x$  回目」は、2.4.2 節の機械学習に基づく方法で  $x$  回目の正例と負例の追加をした後の学習データを用いた場合である。パターン 1 は、3.2 節の分析において、正例に出現する割合が 0.8 以上であった単語を一つでも含む文をすべて正例として抜き出す方法である。パターン 2 は、「感動」という語を一つでも含む文をすべて正例として抜き出す方法である。

10 回正例と負例の追加をした後の機械学習では適合率が 0.40 が得られている。

## 第4章 関連研究

本章では関連研究について述べる。

4.1 節では第1章で紹介した、本研究と同様に文を評価する小論文の自動採点システム Jess の説明を行う。4.2 節では分析に機械学習を利用した研究についての紹介を行う。本研究では読者が感動するかどうか、すなわち読者の感情を推定を行っているのとらえることができる。4.3 節では感情に関わる研究についての紹介を行う。

### 4.1 コンピュータによる小論文の自動採点システム Jess

1章で紹介した石岡らの研究 [1] の小論文の自動採点システムの説明を行う。文章の形式的な側面、いわゆる文章作法を評価する「修辞」と、アイデアが理路整然と表現されていることを示す「論理構成」と、トピックに関連した語彙が用いられているかを示す「内容」の3つの観点から小論文を評価する。毎日新聞の社説およびコラムを学習し、これを模範とした場合に適切でないと判断される採点項目に対して減点することで採点を行う。

### 4.2 機械学習を利用した研究

#### 4.2.1 情報の重用度を定める要因の抽出・分析の自動推定

村田ら [5] は、情報の重用度を定める要因を明らかにし、その知見に基き情報の重要度を自動推定するシステムの構築を目指し、新聞記事やアンケートデータを用いた機械学習を利用した重要度に関する研究を行った。この研究により1つの記事のみを用意して重要かどうかを判別するよりも、2つの記事を用意してどちらがより重要であるかを判別する方が簡単であると報告している。

#### 4.2.2 ユーザ個人の興味の影響を考慮した情報の重用度を決める要因の抽出・分析

村田ら [6] は、情報の重要度を自動推定するシステムの構築を目指し、ユーザ毎に異なる情報の重要度について調査する研究を行った。どのような情報を重要と考えるかは個々の人により異なるものであり、ユーザごとの興味がアンケートより抽出しその結果を利用してユーザ毎に異なる情報の重要度についての調査を行った。この研究によりユーザ個人の興味情報がそのユーザの重要な記事の判断と相関があると報告している。

#### 4.2.3 小説テキストを対象としたジャンル推定と人物抽出

馬場ら [7] は小説テキストを「ジャンル」と「登場人物」により分類する手法を提案した。SVM を用いて登場人物とその特徴を抽出し、それらからジャンル推定を行う。ジャンルによって有効な特徴量が異なること、学習データとテストデータの語の頻度分布の差異で分類精度が変化すると報告している。

### 4.3 感情に関わる研究

#### 4.3.1 語の意味情報を考慮した感情推定アルゴリズム

松本ら [8] は、会話文から感情を推定するためには文に含まれる語句の感情的意味と文の表す事象の意味内容を読み取る必要があるとし、語の意味属性と感情生起事象文型パターンに基いた感情推定アルゴリズムを提案した。

#### 4.3.2 登場人物の感情表現に着目した物語要約

横野ら [9] は、登場人物の感情の変化が示されている部分は他の部分に比べて重要であると仮説を立て、これに基いた要約手法を提案した。登場人物の感情が描写されている部分を抽出するため、感情表現辞典に掲載されている感情表現を含んでいるかで判定する。動作文のみを抜き出した要約文よりも本手法によって生成された要約文の方が要約率が高く、重要な要素を絞り込むことができると報告している。

### 4.3.3 web から獲得した感情生起要因コーパス

徳久ら [10] は、話者の感情を適切に表現する応答を生成するために、発話の意味する感情を感情極性より粒度の細かい感情クラスで推定する必要があるとし、文の意味する感情を推定する手法を提案した。web 上のテキストから感情が生起する要因となる事態を獲得し、それを用いて感情を推定する。あらかじめ感情極性を推定することで感情推定精度が向上すること、複数の類似事例を用いることで感情推定精度が向上することを報告している。

### 4.3.4 リビング環境において ”深い感動 ”を喚起させる「場」の実現方法の検討

石川ら [11] は、リビング環境において ”深い感動 ”を喚起させることを目的とした「場」の実現方法を検討し、その結果を報告した。また、リビング環境において ”深い感動 ”を喚起させる「場」の一つの実現方法を提案した。「体感させること」「フレームレス」「音源に含まれる音が身体に直に伝わること」の重要性を報告している。

### 4.3.5 音楽聴取における ”感動 ”の評価要因—感動の種類と音楽の感情価の関係

大出ら [12] は音楽を聴取するという行為においても、喚起される感動の種類に違いがあることを実験的に示すことを目的とし、既往の研究において分類した感動語を評価語とした感動評価尺度を作成した。音楽を聴いてどういった良さを感じたのかを調べる目的においては感情評価尺度は有効であると報告している。

## 第5章 今後の課題

本研究の今後の課題として以下のことが挙げられる。

1. 構文パターン，修辭的表現など，単語以外での感動を与える文の分析
2. より多くの感動を与える文の収集

## 第6章 おわりに

本研究では感動を与える文の作成支援のために、感動を与える文の収集とそれらの分析を行った。分析の結果、感動を与える文に多く出現する単語として、「人生」「人々」「我々」「恋愛」「喜び」などが得られた。これらを用いた文は感動的なものになりやすいと考えられる。この知見は、感動を与える文を作成する際に役立つものと思われる。

本研究では感動を与える文を機械学習で抽出することも行った。現状で大規模ウェブデータから適合率 0.40 で感動を与える文を取り出すことができることがわかった。

今後は、本研究の技術を利用してより多くの感動を与える文を収集したいと考えている。さらに、感動を与える文の分析として、構文パターン、修辭的表現などの単語以外の分析を行い、さらに多くの感動を与える文の言語的特徴を明かにしたいと考えている。この言語的特徴を利用することで、人に感動を与えることのできる文を作成する際に役立つ文作成支援システムを構築できるようになると考えている。



# 謝辞

本研究を進めるにあたり，種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  
計算機工学講座Cの村田真樹教授に心から御礼申し上げます．本研究を進めるにあたり，  
御指導を頂きました村上仁一准教授に心から御礼申し上げます．また，徳久雅人講師に  
は，研究の進め方や本論文の書き方など，御指導を頂きました．ここに深く感謝いたし  
ます．

その他様々な場面で御助言を頂いた計算機工学講座C研究室の皆様には感謝の意を表し  
ます．

## 参考文献

- [1] 石岡恒憲, 鷺坂由紀子, 二村英幸. Jess: 日本語小論文の自動採点システム.  
<http://coca.rd.dnc.ac.jp/jess/>, 2004.
- [2] Daisuke Kawahara and Sadao Kurohashi. Case frame compilation from the web using high-performance computing. *In Proceedings of the 5th International Conference on Language Resources and Evaluation*, pp. 1344–1347, 2006.
- [3] TinySVM : <http://chasen.org/taku/software/TinySVM/>
- [4] ChaSen : <http://chasen-legacy.sourceforge.jp/>
- [5] 村田真樹, 西村涼, 金丸敏幸, 土井晃一, 松岡雅裕, 井佐原均. 情報の重要度を定める要因の抽出・分析と重要度の自動推定. 言語処理学会第 14 回年次大会, pp. 907–910, 2008.
- [6] 村田真樹, 西村涼, 金丸敏幸, 土井晃一, 鳥澤健太郎. ユーザ個人の興味の影響を考慮した情報の重要度を定める要因の抽出・分析. 言語処理学会第 15 回年次大会, pp. 554–557, 2009.
- [7] 馬場こづえ, 藤井敦, 石川徹也. 小説テキストを対象としたジャンル推定と人物抽出. 言語処理学会第 11 回年次大会, pp. 574–577, 2005.
- [8] 松本和幸, 任福継, 黒岩眞吾. 語の意味情報を考慮した感情推定アルゴリズム. 言語処理学会第 11 回年次大会, pp. 145–148, 2005.
- [9] 横野光. 登場人物の感情表現に着目した物語要約. 言語処理学会第 13 回年次大会 発表論文集, pp. 1141–1144, 2007.
- [10] 徳久良子, 乾健太郎. 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.
- [11] 石川智治, 宮原誠. リビング環境において“深い感動”を喚起させる「場」の実現方法の検討. 芸術科学会論文誌, Vol. 2, No. 3, pp. 91–93, 2003.
- [12] 大出訓史, 今井篤, 安藤彰男, 谷口高士. 音楽聴取における“感動”の評価要因 — 感動の種類と音楽の感情価の関係. 情報処理学会論文誌, Vol. 50, No. 3, pp. 1111–1121, 2009.

# 付録

3.2 節の分析結果から得た 110 単語全てをそれぞれ表 6.1, 表 6.2, 表 6.3 に示す.

表 6.1: 正例に出現する割合が1かつ頻度が5以上の単語

単語	正例に出現する割合	正例の頻度	負例の頻度
幸福	1.00	83	0
友情	1.00	29	0
青春	1.00	18	0
悲しみ	1.00	12	0
存在	1.00	10	0
なさい	1.00	9	0
生涯	1.00	8	0
別れる	1.00	7	0
貧乏	1.00	7	0
素晴らしい	1.00	7	0
幸運	1.00	7	0
欠点	1.00	7	0
恐れ	1.00	7	0
みる	1.00	7	0
古い	1.00	6	0
闘い	1.00	6	0
耐え	1.00	6	0
失敗	1.00	6	0
賢明	1.00	6	0
愛情	1.00	6	0
矛盾	1.00	5	0
恥	1.00	5	0
女房	1.00	5	0
詩人	1.00	5	0
一瞬	1.00	5	0
むしろ	1.00	5	0
お互い	1.00	5	0
いたい	1.00	5	0

表 6.2: 正例に出現する割合が0.8以上の単語

単語	正例に出現する割合	正例の頻度	負例の頻度
我々	0.973684210526316	37	1
不幸	0.96969696969697	32	1
愛さ	0.958333333333333	23	1
恋愛	0.956521739130435	44	2
恋	0.945736434108527	122	7
孤独	0.941176470588235	32	2
この世	0.941176470588235	16	1
愛し	0.939393939393939	31	2
愛する	0.9375	30	2
あらゆる	0.933333333333333	14	1
お前	0.928571428571429	13	1
瞬間	0.916666666666667	11	1
人生	0.911949685534591	145	14
未来	0.909090909090909	20	2
幸せ	0.909090909090909	20	2
喜び	0.909090909090909	10	1
女	0.905511811023622	115	12
運命	0.904761904761905	19	2
死ぬ	0.902439024390244	37	4
知る	0.9	9	1
短い	0.9	9	1
結婚	0.898148148148148	97	11
生まれる	0.888888888888889	8	1
みんな	0.888888888888889	8	1
はじめて	0.888888888888889	8	1
なぜなら	0.888888888888889	8	1
どんなに	0.888888888888889	8	1
決して	0.88	22	3

表 6.3: 正例に出現する割合が0.8以上の単語

単語	正例に出現する割合	正例の頻度	負例の頻度
一生	0.878787878787879	29	4
恋する	0.875	7	1
愚か	0.875	7	1
永遠	0.875	7	1
われわれ	0.875	7	1
だが	0.875	14	2
すぎる	0.866666666666667	13	2
他人	0.863636363636364	19	3
死	0.863636363636364	57	9
馬鹿	0.857142857142857	6	1
静か	0.857142857142857	6	1
生きる	0.857142857142857	30	5
真実	0.857142857142857	6	1
酒	0.857142857142857	6	1
現在	0.857142857142857	6	1
つねに	0.857142857142857	6	1
しかも	0.857142857142857	6	1
恋人	0.846153846153846	11	2
もともと	0.846153846153846	11	2
男	0.840336134453782	100	19
買え	0.833333333333333	5	1
乗っ	0.833333333333333	5	1
兄弟	0.833333333333333	5	1
恐ろしい	0.833333333333333	5	1
たいてい	0.833333333333333	5	1
いつか	0.833333333333333	5	1
人々	0.80952380952381	17	4

表 6.4: 正例に出現する割合が0.8の単語

単語	正例に出現する割合	正例の頻度	負例の頻度
連中	0.8	4	1
立派	0.8	4	1
離婚	0.8	8	2
暮らす	0.8	8	2
忍耐	0.8	4	1
底	0.8	4	1
退屈	0.8	4	1
続く	0.8	4	1
受け入れる	0.8	4	1
失う	0.8	4	1
財布	0.8	4	1
交際	0.8	4	1
苦悩	0.8	4	1
恐怖	0.8	4	1
強く	0.8	4	1
間違っ	0.8	4	1
感動	0.8	8	2
感謝	0.8	4	1
過ごす	0.8	4	1
過ごし	0.8	4	1
チャンス	0.8	8	2
やり方	0.8	4	1
やがて	0.8	4	1
だれ	0.8	4	1
だって	0.8	8	2
あき	0.8	4	1