4I1-R-9-4

# Weblogからの失敗談抽出

Automatic extraction of failure stories from Weblogs

小野 博紀[*1] 内海 彰[*1]
Hiroki Ono Akira Utsumi

[*1]電気通信大学大学院 情報理工学研究科 総合情報学専攻
Department of Informatics, The University of Electro-communications

We propose a method for extracting from weblogs failure stories that provide sufficient information to prevent a failure. Some studies focus on the extraction of experiences, but no studies have addressed the extraction of failure stories. Failure stories consist of an action, a result, and causes, which are extracted using a Japanese thesaurus, connective expressions, and a lexicon for sentiment analysis. Our method extracts a result, an action, and causes for failure stories, in this order. If an action and a result can be extracted from a weblog, it is determined that this weblog includes a failure story. We tried to discover failure stories with our method. The F-measure score of extracting failure stories is 0.32, and that of extracting causes is 0.67. Furthermore, we found that we can discover failure stories more efficiently by using our method than by the traditional keyword search method.

## 1. はじめに

近年，アメーバブログ[*1]や yahoo ブログ[*2]などのブログサイトの普及により，個人の経験を発信する機会が増えている．これらに投稿される記事には，企業や公共機関が発信する情報からは得られない有益な情報が多く含まれている．

その有益な情報のひとつに失敗談がある．失敗談を知ることで，失敗を未然に防ぐことができる．失敗談を検索するにあたり，検索エンジンを用いて Web 全体から「失敗談」というクエリで検索すれば，失敗談を見つけることができるが，ここで見つけられる失敗談以外の失敗談も約 13 億 5,000 万件[*3]のブログ記事内に多く存在するはずである．しかし，現在主に利用されているキーワード検索では，失敗談が述べてある記事を効率良く見つけ出すのは困難である．例えば「失敗談」というクエリで検索を行った場合，失敗談以外の情報が検索結果に多く出力されてしまう．また，失敗談というキーワードがない文章に対しては，出力さえされない．

Weblog を対象としたテキストマイニングの研究の多くは話題情報，評判情報に関する研究であり，失敗談の上位概念である体験談に注目した研究は少ない．その数少ない研究のひとつに倉島ら [1] の研究がある．倉島らは記事から経験を表現する最小要素である時間，空間，動作，対象，評価，感情を抽出し，経験という観点から構造化することで知識発見を試みている．また，原ら [2] は経験を時間情報，極性，話者態度の観点から抽象化する枠組みを提案している．Weblog からの知識獲得という観点からみると，高見ら [3] の研究がある．高見らは，因果関係知識を獲得する対象として新聞記事などの比較的自然言語処理に適した文章ではなく，ブログ記事を選択した．そして，同じ話題に関するバースト特性がコミュニティ毎に異なる点に着目し，ある話題と因果関係にある事象を発見するための手法を提案している．

以上のように，失敗した経験の抽出を目的とした研究は行われていない．また，抽出した経験情報に対して成功/失敗の情報を付与する研究は存在するものの，失敗した経験が失敗を未然に防ぐための情報として十分かどうかまでは考慮されていない．そこで本研究では，情報量のある失敗談を Weblog から抽出する手法を提案する．

連絡先: 小野 博紀，電気通信大学大学院 情報理工学研究科 総合情報学専攻，〒 182-8585 東京都調布市調布ヶ丘 1-5-1，ono1119@utm.se.uec.ac.jp

∗1 http://ameblo.jp/
∗2 http://blogs.yahoo.co.jp/
∗3 2009 年 3 月総務省発表

## 2. 失敗談の定義

本研究では「ある行動を起こした結果，思惑通りに物事が運ばなかったときの体験談」と定義する．失敗に関する体験談には，行動が含まれないことがある．そのような体験談の情報量は，他人がそこから情報を得て失敗を未然に防ぐためには不十分である．物が破損した事に関する体験談を例に挙げると，その破損の原因が利用者の行動であるならば失敗談であるが，単なる時間経過が原因ならば失敗談ではないとして，本研究では抽出の対象としない．図 1 に weblog における失敗談の例を挙げる．丸い枠で囲まれた部分が順に行動，結果である．

## 3. 失敗談抽出手法

### 3.1 概要

本研究では，2 章の定義から失敗談は「行動，結果，原因」という要素で基本的に構成されると考え，本手法はこの 3 つの要素の有無を判定することによって失敗談を抽出する．ブログの本文に対して，以下のような手順で失敗談が存在しているかどうか判定する．

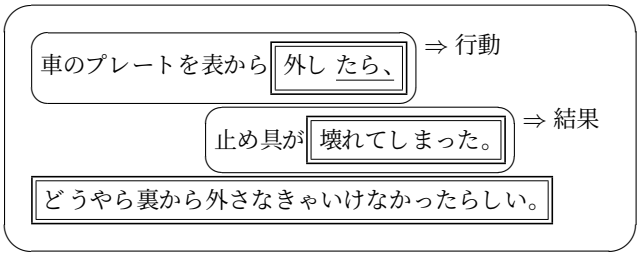


図 1: weblog における失敗談の例

**1. 結果の判定**

ネガティブな語句を含む文節を特定して，その文節の態度が断定であれば結果とする．

結果が存在していた場合，手順 2 へ進む．

**2. 行動の判定**

結果の文節へ係る文節のうち，因果関係を表す接続表現を含む文節を見つけ，その文節が 3.3 節で述べる条件を満たした場合，その文節を行動とする．

行動が存在していた場合，手順 3 へ進む．

**3. 原因の判定**

行動と結果が存在した場合，結果が含まれる文以降の文から 3.4 節で述べる判定方法を用いて原因を述べている文を探す．

図 1 において，2 重枠で囲まれた部分が順に行動，結果，原因にあたる．そして下線付きの部分が因果関係を表す接続表現にあたる．原因がなくても失敗談は成立すると考え，行動と結果があれば失敗談として抽出する．

### 3.2 結果の判定方法

あらかじめ用意した断定以外の表現が，文節内に含まれていたら態度は断定ではないとする．そして断定ではないと判定されなかった場合，態度が断定であるとする．

ネガティブな語句については，失敗知識データベース[*4]を参考にして本研究が人手で作成した極性評価辞書で判定を行っている．

### 3.3 行動の判定方法

まず，結果の文節へ係る文節のうち，因果関係を表す接続表現を含む文節を見つけ，その文節を行動の候補とする．因果関係を表す接続表現には，機能表現辞書 [4] において，順接確定・逆接確定を意味するとされている「～たら」,「～のに」などの機能語と,「そして」,「しかし」といった順接・逆接の接続詞，さらに因果関係を表す際によく使われる「その結果」,「～ところ」などの接続表現を用いる．

そして，行動の候補となった文節が以下の 2 つの条件を満たす場合，行動を述べている文節と判定する．

1. 文節の態度が断定である．
2. 人の行動を表す語句が含まれる．

条件 1 の断定の文であるかどうかは，3.2 節と同様の手法で判定する．条件 2 の人の行動を表す語句には，日本語語彙大系 [5] の用言意味体系において,「身体動作，利用，結合動作」と分類されていた語句を用いる．これを利用することによって「知る」や「覚える」といった知覚や思考動作を表す語句を除き,「使う」や「重ねる」といった人の物理的な動作を表す語句を選択できる．

### 3.4 原因の判定方法

行動と結果と同じ文内で「～ため」といった理由を表す機能語がある場合，その機能語を含む文節と，その文節に係る文節は明らかに原因であるため，その部分を原因とする．文内での原因の有無に関わらず，結果が述べられている文以降の文からも原因の文を探す．3.4.1 節で述べる手法でスコアを計算し，スコアが閾値より高い文すべてを原因の文と判定する．

*4 科学技術振興機構が科学技術分野の事故や失敗の事例を収集したデータベース

#### 3.4.1 スコアの計算方法

ブログ記事において，原因に関する記述は結果の記述の後にくるのがほとんどであるため，結果の文以降の文のみを対象にして，以下の 2 つの指標からスコアを計算する．

1. 結果の文の近くに存在している．
2. 原因の文によく出現する語句が存在している．

指標 1 は，結果の文に近ければ近いほど原因の文らしいという考えからスコアを計算する．したがって，結果の文以降の文をそれぞれ $C_1, C_2, C_3, \cdots$ とおいたとき，ある文 $C_i$ の指標 1 に関する重み $D(C_i)$ を式 (1) で計算する．ただし，$a$ は定数である．

$$D(C_i) = \frac{1}{i} + a \tag{1}$$

指標 2 で用いる原因の文によく出現する語句には，原因の文の中で頻度の高い形態素単位の 1-gram（ひとつの形態素）と 2-gram（2 つの形態素の並び）を利用する．図 1 における「どうやら」や「らしい。」などにあたる．指標 2 ではまず，このようなよく出現する語句の重要度を計算する．ある 1-gram を $u_j$，すべての 1-gram の集合を $U$，原因の文から求めた 1-gram の出現回数を $O_c(u_j)$ と表す．また，原因でない文から求めた 1-gram $u_j$ の出現回数を $O_g(u_j)$ と表す．このとき，$u_j$ の重要度 $f(u_j)$ は式 (2) で計算される．原因でない文の出現回数を引くことで，原因の文特有の 1-gram の重要度は高くなり，一般的な文にも使われる 1-gram の重要度は低くなっている．

$$f(u_j) = \{O_c(u_j) - O_g(u_j)\} / \sum_{u \in U} \{O_c(u) - O_g(u)\} \tag{2}$$

この重要度 $f(u_j)$ を使い，ある文 $C_i$ の指標 2 のスコア $F_u(C_i)$ を式 (3) で計算する．ただし，$n(u_j, C_i)$ は $C_i$ に含まれる 1-gram $u_j$ の個数とする．

$$F_u(C_i) = \sum_{u_j \in U} f(u_j) n(u_j, C_i) \tag{3}$$

2-gram も同様に考え，ある 2-gram を $b_j$，すべての 2-gram の集合を $B$ としたとき，ある文 $C_i$ における指標 2 に関する重み $F_b(C_i)$ は式 (4) で計算する．ただし，$n(b_j, C_i)$ は $C_i$ に含まれる 2-gram $b_j$ の個数とする．

$$F_b(C_i) = \sum_{b_j \in B} f(b_j) n(b_j, C_i) \tag{4}$$

ある文 $C_i$ における最終的な重み $G(C_i)$ は式 (5) で計算する．ただし，$w$ は定数（$0 \leq w \leq 1$）とする．

$$G(C_i) = D(C_i)\{wF_u(C_i) + (1-w)F_b(C_i)\} \tag{5}$$

式 (5) において，指標 1 のスコアと指標 2 のスコアの積をとっているのは，原因の文は指標 1 の要素があり，かつ指標 2 の要素もあるのが望ましいからである．また，$F_b(C_i)$ と $F_u(C_i)$ は，同じ観点から原因の文らしさを測るため，和をとっている．式 (5) で計算されたスコアが閾値を超えた文すべてを原因の文とする．

表 1: 行動と結果の抽出精度

| | 再現率 | 適合率 | F 値 |
|---|---|---|---|
| ランダム | 0.500 | $0.099\times10^{-1}$ | $0.194\times10^{-1}$ |
| 日本語評価極性辞書 | 0.010 | 0.013 | 0.011 |
| 単語感情極性対応表 | 0.170 | $0.077\times10^{-2}$ | $0.153\times10^{-2}$ |
| 用意した極性辞書 | 0.240 | 0.480 | 0.320 |

表 2: 発見した失敗談の個数

| | キーワード検索 | システムを利用 |
|---|---|---|
| 被験者 A | 0 | 7 |
| 被験者 B | 3 | 14 |
| 被験者 C | 2 | 7 |
| 被験者 D | 9 | 18 |
| 被験者 E | 1 | 16 |
| 平均 | 3 | 12.4 |

## 4. 評価

### 4.1 行動・結果の抽出手法の評価

ここでの評価は，行動と結果が両方正しくないと正解とみなさないとする．また，失敗談は行動と結果があれば成立すると考えるため，この節での評価が失敗談抽出の精度の評価となる．評価データとして，カテゴリが「車」に分類される記事の中から，失敗談を含む記事を 100 件，含まない記事を 10000 件用意した．カテゴリを「車」としたのは,「恥ずかしかった」や「1 日何もしないで終わってしまった」などの失敗談は対象とせず,「壊れた」や「割れた」などの物理的な被害を伴うような失敗に関する失敗談を対象としているからである．また，正例と負例の比が 1:1 でないのは，実際のブログサイト全体における失敗談の割合に合わせたからである．これらの記事を使い，F 値を用いて手法を評価した．F 値は再現率と適合率の調和平均であり，以下の式で表せる．

$$F\text{ 値} = \frac{2PR}{P+R} \quad (6)$$

$$\text{適合率 } P = \frac{\text{システム正解数}}{\text{システム出力数}} \quad (7)$$

$$\text{再現率 } R = \frac{\text{システム正解数}}{\text{正解数}} \quad (8)$$

評価極性を対象とした研究でない限り，単語の極性判定には公開されている極性辞書を用いるのが普通である．そこで極性の判定を公開されている 2 つの極性辞書を用いた場合と，本研究で用意した極性辞書を用いた場合を比較した．さらにランダムに記事を失敗談と判定した場合の精度とも比較した．比較実験で用いた極性辞書のデータは以下の通りである．

- 日本語評価極性辞書 [6]
  - 「ネガティブ（経験）」に分類されている語句
- 単語感情極性対応表 [7]
  - 感情属性値が -0.8 以下の単語

その結果を表 1 に示す．表 1 より，本手法の失敗談の抽出精度を表す F 値は 0.32 となった．また，用意した極性辞書を用いた場合が最も精度良く抽出できた．

### 4.2 原因の抽出手法の評価

失敗談で，なおかつ原因の文を含む記事 100 件と含まない記事 100 件を合わせた 200 件の記事を 5 分割し，4 セットで式 (1) と式 (5) における定数 $a, w$ と閾値を学習，そして残りの 1 セットでテストを行う実験を繰り返す交差検定を行った．原因の判定に用いる 1-gram と 2-gram は，人手で集めた評価データと異なる原因の文 250 文と原因でない文 250 文から求めた．

その結果，5 回のテストの平均は再現率が 0.674，適合率が 0.679 であり，F 値は 0.668 となった．

### 4.3 失敗談収集の効率性の評価

本手法を実装したシステムと，今日主に利用されているキーワード検索で，どちらが失敗談収集の効率が良いかを調べるために実験を行った．

約 5 万件のブログ記事と，それに対するシステムの出力結果約 600 件を用意した．ブログ記事には記事の題名とその本文，システムの出力結果には本手法によって取り出された「行動，結果，原因」と，その抽出元の記事の本文が書かれている．ブログ記事とシステムの出力結果は，それぞれテキストファイルとして用意した．これらを利用し，ブログ記事から失敗談を探す場合はキーワード検索を用いて，システムの出力結果から探す場合はキーワード検索を用いずに，それぞれ 10 分間の間に何件失敗談を見つけられるかを，学生 5 人に対して実験を行った．キーワード検索と本手法を F 値で比較するのではなく，時間内に発見できる個数で比較したのは，キーワード検索の F 値を一意に決めることが難しいからである．F 値を一意に決めづらい理由として，通常キーワード検索を使って目的の記事を見つける際，複数回検索を行うことが挙げられる．

実験の結果を表 2 に示す．発見した失敗談の個数の平均はキーワード検索を使った場合が 3.0 個，システムを利用した場合が 12.4 個となり，キーワード検索よりもシステムを利用して失敗談を検索した方が同じ時間内で多くの失敗談を発見できた．

## 5. 考察

4.1 節より，失敗談の抽出には特有の極性辞書が有効だということがいえる．本研究で用意した極性辞書は「動かなくなる」,「割れる」,「ガタが発生する」など，事象自体の極性を表す語句が多い．一方，一般的な極性辞書は「呆れる」,「困る」,「恨む」など，感情語に至るまで極性をつけてしまうために，余計な抽出が多くなってしまっていた．

**感情語に極性がついていることで抽出されてしまう例**

- 何年も乗ってるけど、まだ 不安 だ。
- 板を外すと、ちょっと 寂しい。

失敗談抽出の精度の値そのものがあまりよくなかった理由として，以下の例文の（a）,（b）のようにブログ特有の表現に対応できなかったことや，例文（c）のように行動がネガティブな結果につながっていない文を失敗談として誤って抽出していたことなどが挙げられる．

表 3: 原因の抽出における指標の有効性の比較

| | 再現率 | 適合率 | F 値 |
|---|---|---|---|
| ランダム | 0.500 | 0.116 | 0.189 |
| 指標 1 のみ | 0.609 | 0.236 | 0.334 |
| 指標 2 のみ | 0.590 | 0.631 | 0.564 |
| 提案手法 | 0.674 | 0.679 | 0.668 |

**誤って抽出してしまった例**

**（a）**カッコ良くドリフトで入ろうとしたら、振りっ返しに失敗ドアンダーを出して農作業用の洗濯機に …。

**（b）**ちょっと力んでしまったら「ピキッ」嫌な音が …。ミラー 1 枚割っちゃった。

**（c）**外して みると、ベルトが 切れていた。

ブログでよく用いられる「…」といった記号をそのまま記号として扱うのではなく，極性や接続表現などの情報を持った記号として扱うことで，再現率を上げることができると考えられる．また，誤った抽出を防ぐには行動の判定を 1 文節よりも広げて行うことが有効だと考えられる．例えば例文（c）と同じ「外す」という動作でも，図 1 の場合は「表から」という情報が加わることによってネガティブな結果につながる行動となっている．

4.2 節に関して，指標別に精度を比較した結果を表 3 に示す．それぞれの値は，4.2 節と同様に交差検定を行った結果である．表 3 をみると，提案手法は指標 1 と指標 2 を組み合わせることで，適合率を大きく向上させていることがわかる．したがって，ブログにおける原因の文の抽出には，語彙情報に加えて位置情報を用いることは有効であると考えられる．

4.3 節から，キーワード検索よりもシステムを利用して失敗談を検索した方が，効率が良いといえる．システムの出力結果には行動と結果という因果関係を持つ文が抽出されていることで，単語ベースの検索よりも情報量が多い記事が見つけやすい．そのため，システムを利用した方が効率良く発見できたと考えられる．

## 6. おわりに

本研究では，失敗談を構成する「行動，結果，原因」という 3 つの要素に着目して，Weblog から失敗談を抽出する手法を提案した．失敗談の抽出精度を表す F 値は 0.32 となった．また，失敗談の抽出には特有の極性辞書が有効だということを明らかにした．原因の文の抽出には，提案手法が有効であることがわかった．そして，本手法を用いることで失敗談を効率良く収集できることを示した．

今後の課題として，行動と結果の抽出精度が十分でない点がある．精度の向上には，これまで意味を与えてこなかった記号に対して焦点を当てることや，行動の判定の際に利用する範囲を 2 文節以上に広げることが必要である．

## 参考文献

[1] 倉島 健，藤村 考，奥田 英範: 大規模テキストからの経験マイニング, 電子情報通信学会論文誌, Vol.J92-D, No.3, pp.301-310 (2009).

[2] 原　一夫，乾 健太郎: 事態抽出のための事実性解析, 情報処理学会 研究報告, 2008-FI-89，2008-NL-183 (2008).

[3] 高見真也，田中克己: ブログのコミュニティ分析による因果関係事象の抽出，情報処理学会 研究報告，2006-DBS-140 (2006).

[4] 松吉俊, 佐藤理史, 宇津呂武仁: 日本語機能表現辞書の編纂，自然言語処理， Vol.14，No.5，pp.123-146 (2007).

[5] 池原悟，宮崎正弘，白井諭，横尾昭男，中岩浩巳，小倉健太郎，大山芳史，林良彦：日本語語彙大系，岩波書店 (1997).

[6] 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，Vol.12，No.2，pp.203-222 (2005).

[7] 高村大也, 乾孝司, 奥村学: スピンモデルによる単語の感情極性抽出，情報処理学会論文誌，Vol.47，No.02，pp.627–637 (2006).