

# 製品コンセプトと製品評価文章の関係の分析

長江 朋 望月 源 白井 清昭 島津 明

北陸先端科学技術大学院大学

t-nagae@jaist.ac.jp

## 1 はじめに

製品を購入する際、購入の判断材料となる情報源のひとつにカタログがある。我々は、企業のサイトにアクセスすれば製品のカタログを見ることができる。しかし、カタログなどから得られる情報は、主に企業からの情報や製品のコンセプトである。企業は、製品コンセプトを記載することにより、ユーザに対して製品の良さをアピールしている。これら以外の情報源としては雑誌やWeb上の製品評価サイトなどがある。評価サイトでは、実際にその製品を使用したユーザの生の声を聞くことができる。この時、企業の製品コンセプトと、それに対するユーザの評価を同時にユーザに提示できれば、ユーザにとって有益な情報となる。

Web上にある情報源からのデータ抽出に関連した研究として、有賀ら [1] は、カテゴリ別の製品情報の自動収集により、製品カテゴリ毎に製品一覧表を生成して、ユーザの製品購入の意思決定を支援するシステムを作成した。一方、立石ら [2] は、Webページを対象に評判情報検索として、商品の評価意見の抽出を行っている。

本研究では、Web上の製品評価サイトから製品評価文章を収集し、評価語辞書と評価判定規則により、ユーザの評価している対象と肯定、否定の判定を行う。そして、製品のカタログから企業が製品を販売する際の売り文句などを製品コンセプトとして抽出す。ユーザが評価している対象が製品コンセプトと一致するのかを判別する。製品コンセプトが含まれていない文については、肯定的、あるいは否定的評価文に分けて出力する。この様にして製品コンセプ

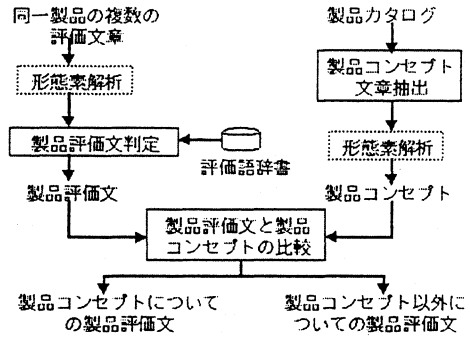


図 1: システム構成

トとユーザの評価文章を比較することにより、製品コンセプトがユーザに認識されているのかを判定するシステムの構築を考えている (図 1)。本稿ではシステム構築に関し、製品評価サイトから取出した評価文章を分析、それに基づく評価語辞書の作成、製品コンセプトの分析について述べる。

## 2 製品評価文

分析の対象とする評価文章とは、ユーザが実際にある製品を使用したり、あるサービスを利用したときの評価を記述した文章のことである。評価文章の例を図 2 に示す。このような評価文章中に存在する属性と評価語の組を抽出することにより、肯定、否定の評価判定を行うことを考えている。

属性とは、評価の対象となる単語のことであり、各製品を構成する要素のことである。評価語は属性を修飾する肯定的、あるいは否定的な

【インターネット証券取引の評価文章】  
 なんとと言っても手数料が馬鹿高いことです。それにしては他社と比較してお粗末なサポート、ホームページの使いにくさ。私も含めたゴミ投資家ならジェット証券、スーパーセットを求めるならDLJダイレクトがいいですよ。

【ノートパソコンの評価文章】  
 機能的には良いのだが、ちょっと重いのが難点。

図 2: 評価文章の例

意味合いを表す単語のことである。

例えば図 2 のノートパソコンの評価文章では、以下の属性と評価語が読み取れる。

- (機能：良い：肯定)
- (?: 重い：否定)
- (属性：評価語：評価)
- 「?」は属性がないことを表す。

評価文章から評価語と属性を組として取出されるものは以下のように分類できる。

- 属性と評価語
- 評価語のみ
- その他 (評価語が含まれていない文章)

その他の例として「PC-PJ2です。」がある。これは製品名を述べた文である。

分析する評価文章は製品評価サイト「owner's voice」[3] から収集した。同サイトは、製品やサービス別に 89 のカテゴリに分類されている。評価語の種類は製品カテゴリに依存すると思われる。そこでカテゴリに依存する評価語があるのかを調べてみた。その結果を表 1 に示す。割合は、カテゴリ数を全カテゴリ数 89 で割ったものである。評価語は、角川類語辞典の項目《良否》の見出し語から、肯定否定の判定ができる 110 単語を選んだ。

表 1: 多くのカテゴリに用いられている評価語

カテゴリ数	評価語	割合
70	良い	0.79
68	満足	0.76
54	高い	0.61
53	悪い	0.60
46	最高	0.52

評価語「高い」は、属性により評価が変わる評価語である。カテゴリ「インターネット証券取引」における評価文の具体例をあげる。

評価文 A：信頼度が高い

評価文 B：手数料が高い

評価文 A は、信頼度を修飾し、肯定的評価である。一方、評価文 B は手数料を修飾し、否定的評価である。評価語「高い」が属性によりどのような評価をするのか、全てのカテゴリの評価文章について調べた。調査評価文章数は 4775 件である。評価語「高い」が肯定的な評価であったカテゴリと属性の一例を以下に示す。

- <カテゴリ>：<属性>
- インターネット銀行取引：人気
- 自動車保険：満足度
- ベビーラック：安全性
- R V (ラージ)：完成度
- デスクトップ：コストパフォーマンス

否定的な評価であったカテゴリと属性を以下に示す。

- インターネット証券取引：手数料
- 南の島リゾート：旅費
- 旅行会社：キャンセル料
- R V (コンパクト)：税金
- プリンタ (カラー)：ランニングコスト

カテゴリがノートパソコンのとき、属性により評価が変わる評価語の出現数を表 2 に示す。

しかし、評価語と属性の組だけでは肯定、否定の判定ができない。以下にその考慮すべき点について例をあげ検討する。「xx」は評価語と属性の組を表す。

表 2: 属性により評価が変わる評価語の出現数

評価語		出現数	属性	評価
おおきい	大きい	9	画面	肯定
	大きく	2	持ち運び	否定
おちない	おちない	1	システム	肯定
	落ちない	1	汚れ	否定
すくない	少ない	2	故障	肯定
	少ない	3	HDD	否定
たかい	高い	5	値段	否定
	高い	1	完成度	肯定
ちいさい	小さい	4	本体	肯定
	小さい	4	画面	否定
ない	ない	2	装備	否定
	ない	1	エラー	肯定

表 3: 評価語 + 「すぎる」の出現数

評価語	出現数	属性	評価
大きすぎ	2	持ち運び	否定
多すぎる	4	ソフト	否定
重たすぎ	1	持ち運び	否定
良すぎる	1	操作	否定

● xx + すぎる

評価語に「すぎる」がつくと、評価が肯定であっても、その評価は否定になる場合が多い。表3に「すぎる」が付いた評価語と出現数を示す。これ以外にも否定的（要求表現も含む）評価になる場合が多い。

- もう少し xx ほうがいい
- もう少し xx ればいい
- xx できないか

肯定、否定の判定ができないものに以下の評価文があった。

スキー場：標高が高い

- 評価文 -

御岳スキー場は標高が高いためシーズン前より滑れますが、R19号から遠いため行くのが結構しんどい事と食堂を充実して欲しいです。

R'V (コンパクト)：車高が高い

- 評価文 -

車高が高いため運転しやすく、見た目も良い。

RV (ミドル)：車高が高い

- 評価文 -

ちょっと、車高が高い分ロールがキツイのが難点かも…。

カテゴリがスキー場の場合、「シーズン前より滑れます」という表現によりはじめて肯定的な評価と判定できる。RVコンパクトとRVミドルはどちらも「車高が高い」という属性と評価語の組である。しかし、その組に続く表現により、評価が肯定と否定に分かれてしまう。

属性と評価語の組で評価をすることは、評価対象を絞り込むためには有効な手段である。しかし、属性と評価語の組だけで評価の判定をすることができない場合もある。この組の前後の表現にも着目して評価を判定しなければならない。また、評価判定に前後の文脈が必要な場合も考えられる。

### 3 評価語辞書

前節の分析の基づき評価語辞書を作成した。評価語と属性の組だけでは判定できないものについては、評価判定規則を適用して対処することを考えている。

評価語辞書の項目は、製品評価文章を茶筌[5]で形態素解析し、評価語である単語に人手で肯定、否定の評価を付け収集した。辞書に登録した評価語の数は565である。評価語辞書の項目は、カテゴリ、評価語、属性、評価からなる。その例を表4に示す。評価語辞書の項目「-」は、どのカテゴリや属性に対しても評価が変化しないことを表す。

### 4 製品コンセプト文章

製品カタログのトップページに記述されている文が最も見る者の興味を引く表現がされている

表 4: ノートパソコンの評価語辞書

カテゴリ	評価語	属性	評価
ノートパソコン	大きい	持ち運び	否定
		画面	肯定
	小さい	画面	否定
		HDD	否定
		サイズ	肯定
-	おすすめ	-	肯定
-	悪い	-	否定

高性能を引き出すための機能性と美しいスタイルを両立。フラッシュユーザーフェスデザインインタークーラーフラップ機構。  
快適な操作性を追求したジョグダイヤルステンレスメカキー。

図 3: コンセプト文章

ると考えられる。しかし、その表現は抽象的な表現であり、ユーザの評価文章との比較において、抽出された単語との比較をすることが困難である。また、抽象度を上げることにより一致する単語の増加が考えられる。これも一つの製品コンセプトと評価文章との関連になるが、より細かい対象を捉えるためにも、あまり抽象的な表現は好ましくないと考える。製品カタログの構成に基づき本研究では製品紹介ページのうち「製品の特長」などを手がかり語として、その手がかり語を含むページなどを解析の対象と考えている。

製品コンセプトを企業側の「売り文句」であると定義する。そして、製品コンセプト文章を製品カタログ中の文章で視覚に訴える表現をしている文章とする。また、製品コンセプトワードを製品コンセプト文章中にあるその企業独自の言葉とする。製品コンセプト文章の一部を図 3 に示す。

製品コンセプトワードは、製品コンセプト文章を形態素解析し、その情報をもとに抽出を試みた。具体的には評価者の評価対象となりやすい単語の品詞は名詞であることから、ある程度

フラッシュユーザーフェスデザインインタークーラーフラップ機構ジョグダイヤルステンレスメカキー

図 4: 製品コンセプトワード

自動的に製品コンセプトワードを取出すことが考えられる。そこで、名詞を第 1 の候補として抽出し、その抽出された名詞を、他の製品の製品コンセプトワードとマッチングをとる。このマッチングでマッチしなかった単語は、その企業独自に作成した単語である可能性が高いと考え、実際に行った結果、一般名詞が含まれる問題があるが、取出すことができた。図 4 に図 3 の製品コンセプト文章から得られる製品コンセプトワードを示す。

## 5 おわりに

本稿では、企業の製品コンセプトとユーザの製品評価文章とを比較するシステムについて提案し、そのために製品評価文章の分析、評価語辞書の作成、製品コンセプト文章の分析について述べた。現在、これらの分析結果をもとにシステムを構築している。

## 参考文献

- [1] 有賀忠徳, カテゴリに基づく製品情報の組織化, 北陸先端科学技術大学院大学修士論文, 2000
- [2] 立石健二, 石黒義英, 福島俊一, インターネットからの評判情報検索, 情報処理学会研究報告, NL-144-011, 2001
- [3] URL: <http://www.ownersvoice.ne.jp/>
- [4] URL: <http://www.toppage.ne.jp/main.html>
- [5] URL: <http://chasen.aist-nara.ac.jp/>