

商品の属性と属性値を利用したキャッチコピーの自動作成*

森本 直樹 玉川 洋輔 牧野 正治 韓 東力
 日本大学文理学部 情報システム解析学科

1 はじめに

近年、様々な商品やサービス、さらには事業やイベントなどあらゆるところでキャッチコピーが使われており、キャッチコピーの自動作成を支援する研究も進められている。松平らはユーザーにその商品に関するキーワードを入力してもらい、そこからの連想をもとにキャッチコピーを作成するシステムを開発した[1, 2]。

本研究では、我々は、ユーザーの負担を極力少なくし、気軽に素早くキャッチコピーの作成を可能にすることを研究目的とする。そこで、本研究では商品の特徴を表すことができる「属性」と「属性値」に着目し、それらを用いてキャッチコピーを作成する手法を考案した。次章以降では本システムの概要と各モジュールについて説明していく。

2 システム概要

キャッチコピーには、1行程度のものから複数行に渡る物まであり、使用用途も企業に対するものから、小売店が宣伝に使う POP など、幅広く存在している。本研究ではメーカーが商品に付与するキャッチコピーを対象を絞ることにした。

キャッチコピー作成をシステム化するにあたって、キャッチコピーの作成には何かの法則があるのではないかと我々は考え、「売れるキャッチコピーがスラスラ書ける本」[3]をもとにキャッチコピーでよく使われる法則を調査した。以下はこの調査で得られた法則の例である。

(例)

- 損をさせないと訴える方法
 - ・豆乳クッキーダイエットヤセなかったら全額返金 (健康コーポレーション株式会社 / 豆乳クッキーダイエット)
- 数字をコピーに入れる方法
 - ・限定 10 個! (QUESTINA / キティプリザーブドリース)

しかし、これらの法則にはシステム化することが困難なものが多数存在し、一部の法則だけをシステム化してもできあがるキャッチコピーに偏りが現れる可能性がある。そこで、これらの法則は本研究では使わず、次で述べる商品の属性と属性値を利用する独自の方法を考案した。

本研究では商品のキャッチコピーを作成するので、その商品のセールスポイントが必要になる。セールスポイントには、商品の特徴や特性などを

表す属性とその属性を修飾するための属性値が必要であると考えた。

我々は、Web 上の大量のデータを解析して属性と属性値のデータベースをあらかじめ作成しておき、これを利用する手法を考案した。図 1 はシステムの流れを示す。

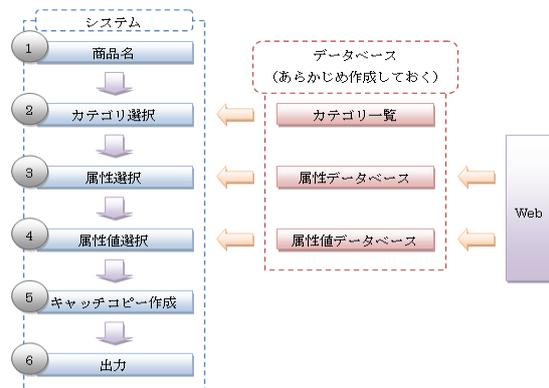


図 1: システムの流れ

まず、ユーザーが商品名を入力する。これはキャッチコピーに商品名を入れる場合に使用するもので、商品名を入れない場合はこの作業は省略が可能である。次に、カテゴリをカテゴリ一覧から選択する。カテゴリ一覧はあらかじめ作成しておく。選ばれたカテゴリに対して、ユーザーがキャッチコピーに使う属性を属性データベースから、選択された属性に対応する属性値を属性値データベースから選択する。属性と属性値のデータベースはそれぞれ Web を解析してあらかじめ用意しておく。

ここまでの作業で得られた商品名(省略可)、カテゴリ、属性、属性値を後述のテンプレートに当てはめキャッチコピーを作成する。最後に、キャッチコピーが複数出力されるので、ユーザーはの中から気に入ったものを選ぶ。

3 カテゴリ分け

我々の作成したシステムでは、キャッチコピーを作成する商品ごとにカテゴリ選択を行っている。キャッチコピーを作成するときには属性を選択する必要がある。商品には様々な特徴があり、妥当な属性を選択するためには、商品ごとにカテゴリ選択する必要があると考えた。

我々は図 2 のように商品のカテゴリを細分化し、

* Automatic Generation of Advertising Slogan Based on Products' Properties

カテゴリごとに Web から属性を抽出してみた。対象カテゴリを深く設定すると有効な属性が多く取得できることに気づいた。

例えば、カテゴリを『食品』とすると、属性は『味』くらいしか取得できないが、もう少し深くしてカテゴリを『飲料』とすると属性は『味』、『喉越し』、さらにカテゴリを『コーヒー』とすると属性は『味』、『苦み』、『香り』などのように、より有効な属性が増えてくる。

我々は「キャッチコピー大百科」[4]とYahoo!オークション¹の商品カテゴリを参考に「デジタル関連製品」、「家電製品」、「食品」に限定し、商品のカテゴリを細分化した。図2はカテゴリ分けの一部(食品)である。



図2:カテゴリ分けの一部(食品)

4 商品の属性と属性値の抽出

本章では、商品の属性と属性値の自動抽出について述べる。

4.1 属性の抽出

我々は属性を自動抽出するに当たって文章中の「○○(商品カテゴリ)の・・・」にある「・・・」に当たる語の中に属性があると考え、そこに焦点を当てた。

- (例) コーヒーの芳醇な香り
 → 属性: 香り
 コーヒーの苦味が～
 → 属性: 苦味

次に属性の抽出元として、常にリアルタイムで更新され、人々の意見が簡易に取得できる Web 上にある情報を利用する。しかし、Web 上における意味のなさないサイトも検索対象となるので、Yahoo オークションをドメイン指定し、サイト内検索をすることにより検索の効率を向上させる。Web 検索を行うときの検索キーワードとして「○○(商品

¹ <http://auctions.yahoo.co.jp/>

カテゴリ)の」を利用する。

また、検索キーワードとなる○○の部分に当たる商品カテゴリについては図2で示したカテゴリ分けにおける一番深いものを使用する。

属性抽出システムでは、Yahoo API²を用いてWeb上(インターネット上)のテキストから属性を抽出する。検索キーワードに「○○の」と入力し、検索されたWebページから「○○(商品カテゴリ)の・・・」を含むテキストを一文ずつ抽出し、その文に対し形態素解析(Juman³)と構文解析(Knp⁴)を行い条件に当てはまったものだけを商品の属性として抽出する。ここで、入力した商品カテゴリの係り受け先の単語が普通名詞ならば属性として抽出するという条件を利用する。

- (例) 文1) コーヒーの味が苦い
 文2) コーヒーの芳醇な香りが素晴らしい

係り受け解析を導入することにより、文1と文2の両方から属性が抽出できるようになる。次に「コーヒー」の属性を抽出した過程を図示する。

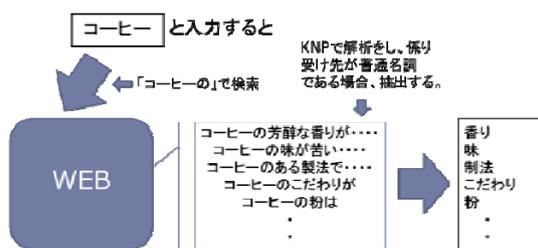


図3:属性抽出システム

4.2 属性の評価

図4の左側にあるように抽出された属性を出現頻度で並べかえると、「生」や「まわり」などコーヒーの属性としては不適切な語が混じることがある。

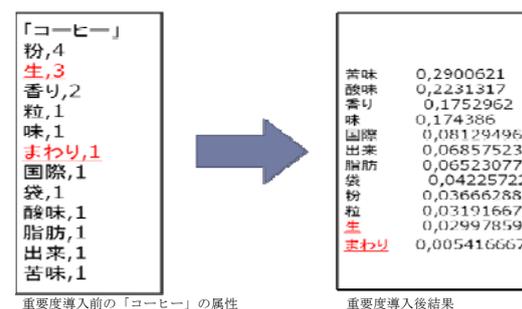


図4:「コーヒー」の属性対重要度導入した結果

² <http://developer.yahoo.co.jp/>

³ <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman.html>

⁴ <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/knp.html>

そこで、我々は有効な属性を見つけるために、各属性についての重要度を式(1)によって計算した。

$$\text{抽出語の重要度} = \frac{\text{カテゴリと抽出語が共起しているページ数}}{\text{抽出語が出現しているページ数}} \dots\dots (1)$$

式(1)は属性がどのくらいカテゴリにかかわっているかを示すことができる。重要度の数値が大きいものほどそのカテゴリと抽出語の関係度合が高いと考えられる。抽出された属性を重要度の高い順にソートし、上位 10 件を各商品カテゴリの属性として採用する。

4.3 属性値の抽出

属性値の多くは「形容詞」や「形容動詞」である。そこで、我々は「△△(属性)が・・・」の「・・・」に当たる表現の中に属性値があると考えた。

(例)

このテレビは映像がすこぶる綺麗
→ 属性値： 綺麗
テレビの画面が大きい
→ 属性値： 大きい

属性値は前々節で示した手法と同様に Web 上から取得する。Web 検索を行うときの検索キーワードを「△△(属性)が」とする。重要度計算を導入し抽出された上位 10 件の属性についてそれらの属性値を取得する。

属性値の抽出も係り受け解析の結果をもとに行われる。これは、属性を取得する時の理由と同様に「△△(属性)が」のすぐ後ろの形容詞を取得するだけでは、前例の「すこぶる」の様に間に修飾語などが入ってしまう文には対応が出来ないからである。

4.4 属性値評価

次に、属性と同じように属性値にも重要度を導入する。

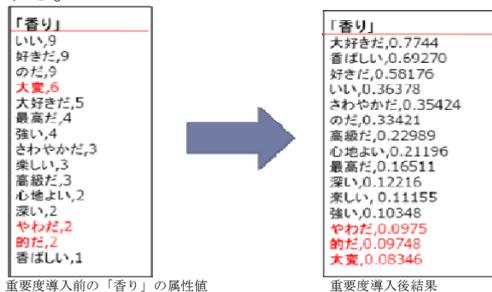


図 5: 「香り」の属性値に対する重要度導入結果

図 5 から分かるように抽出された属性値にも属性の時と同様に「大変」、「やわだ」、「的だ」など、「香り」の属性値としてふさわしくないものや不適切な語が混じって抽出されてしまうことがあるので、式(2)を使い各属性値に重要度計算を行った。

$$\text{抽出語の重要度} = \frac{\text{属性と抽出語が共起しているページ数}}{\text{抽出語が出現しているページ数}} \dots\dots (2)$$

5 テンプレートの自動作成

そもそも、キャッチコピーの多くは文章である。文章をコンピューターで作ることは高度な技術が必要とされており、容易ではない。また、本研究では商品の属性と属性値に着目しているため、こういった商品の特徴を効率的に反映させる作成法を考案する必要がある。そこで我々は最も単純で確かな方法としてテンプレートを利用することとした。その一例を下に示す。

(例)

商品名(名詞)+助詞+属性1(名詞)+助詞+属性値1(形容詞)
テンプレートには「属性」、「属性値」、「商品名」、「カテゴリ」、「助詞」などの要素が含まれており、各要素に情報をあてはめることでキャッチコピーが自動作成される。たとえば、上のテンプレートに各種の情報を適用すると、「〇〇コーヒーの香りが素晴らしい」のような簡単なキャッチコピーを生成することができる。本章では、このようなテンプレートを作成する過程を述べる。

5.1 テンプレートの品詞構造

まず、我々はキャッチコピーに使われる品詞列の構造に一定の法則性があるのではないかと考え、「カタログ・チラシ キャッチコピー大百科」[4]の一部のカテゴリに対して、形態素解析を行い、品詞列の構造解析を行った。

表 1: キャッチコピーに対する品詞構造解析結果 (一部)

デジタル関連製品 全594より 集計 上位13位		出現頻度
解析結果		
名詞 助詞 動詞 名詞 動詞		3
名詞 助詞 名詞 助詞 動詞		3
名詞 助詞 動詞		3
形容詞 名詞 助詞 形容詞 動詞		3
名詞 助詞 動詞 形容詞		3
形容詞 名詞 助詞 形容詞 名詞 助詞		3
名詞 カタカナ		3
名詞 助詞 形容詞		3
名詞 助詞 名詞 助詞 カタカナ		2
名詞 助詞 助詞 名詞		2
名詞 名詞		2
カタカナ 助詞 形容詞		2
名詞 助詞 動詞 名詞		2

表 1 に示される品詞解析の結果から品詞列の構造に最も多く出現している品詞列の頻度が 3 であり、総数 594 に対してあまりにも少ないと感じた。この結果よりキャッチコピーから直接抽出された品詞列の構造に一定の傾向あるいは法則性を見出すことは難しいと判断した。

5.2 テンプレート作成手順

本節では、表 1 の品詞列の構造をある程度まとめることで、より一般性の高い法則を抽出し、テ

ンプレートを自動的に作成する方法を述べる。

5.2.1 第一段階

まず、テンプレート自動作成の第一段階として品詞解析を行ったものに対して、次の4種類の交換規則を適用する。

- ・変換規則① 連続する同一の品詞を、一つの品詞にする。
- ・変換規則② 動詞を形容詞に変換する。
- ・変換規則③ 「その他」、「記号」、「連体詞」を削除する。
- ・変換規則④ 助詞を表層表記のまま残す。

変換規則は①~④の順に適用する。これは、どの変換から行っても結果は同じだからである。4つの変換規則の全てが適用できなくなるまで続ける。実際に5.1で品詞解析を行ったものに対し第一段階の変換を行った結果を表2に示す。

表2: 表1の品詞列構造解析結果に対する変換結果(一部)

変換結果(一部)	出現回数
形容詞 名詞 形容詞	8
形容詞	7
名詞	6
形容詞 名詞 を 形容詞	5
名詞 で 形容詞	5
名詞 カタカナ	4
名詞 形容詞	4
名詞 を 形容詞	4
名詞 に 形容詞 名詞 を 形容詞	4
形容詞 名詞 で 形容詞	3
形容詞 名詞 は 形容詞	3
形容詞 名詞 へ	3
名詞 の 形容詞 名詞	3
形容詞 名詞 で 形容詞 名詞	3
名詞 も 形容詞	3
形容詞 名詞	3
名詞 も 名詞 も 形容詞	3
名詞 を 形容詞 名詞 カタカナ	3

5.2.2 第二段階

次に、テンプレート自動作成の第二段階として、第一段階で変換された結果に対して、さらに次の表3に示されるような局所換言ルールを適用する。

表3: 局所換言ルール

局所換言ルール	ルールが適用できる品詞列の構造と付与する情報
ルール①	【名詞 助詞 形容詞 [助詞] × n ⇒【属性値 助詞 属性値 [助詞]】
ルール②	【形容詞 名詞 [助詞] × n ⇒【属性値1 属性1 [助詞]】
ルール③	文末が【形容詞 名詞 助詞】 ⇒【属性値 カテゴリ名 助詞】
ルール④	【名詞 女子(が・の) [形容詞] 名詞】 ⇒【カテゴリ名(商品名) 助詞 [属性値] 属性】
ルール⑤	文末が【助詞(の) 形容詞】 ⇒【助詞 カテゴリ名(商品名)】
ルール⑥	【形容詞 名詞】 ⇒【属性値 属性】
ルール⑦	【名詞 助詞(で・に)】 ⇒【カテゴリ名(商品名) 助詞】 【名詞 助詞(も・の・が・とを)】 ⇒【属性 助詞】
ルール⑧	【形容詞】 ⇒【属性値】
ルール⑨	【カタカナ】 ⇒【カテゴリ名(商品名)】

(i=1, 2, 3... n=2, 3, 4...)

第一段階で2回以上の頻度で出現したものに対して、以上の局所換言ルールを①から一回ずつ順に適用する。どのルールも適用できないものは除去する。局所換言ルールの適用例は以下に示す。

形容詞 名詞 が 形容詞 名詞 の 名詞

↓ ←換言ルール②を適用
属性値1 属性1 が 属性値1 属性2 の 名詞

↓ ←換言ルール⑤を適用
属性値1 属性1 が 属性値1 属性2 の カテゴリ名

実際に表2に対して第二段階の局所換言ルールを適用して生成されたテンプレートの一部を表4に示す。

表4: 作成されたテンプレート(一部)

ルール適用結果(一部)
属性値1 属性1 属性値2(連続可)
属性値(連続可)
属性値1 属性1 を 属性値2(連続可)
商品名(カテゴリ名) で 属性値(連続)
属性 を 属性値(連続)
属性1 に 属性値2 属性2 を 属性値3(連続)
属性値1 属性1 で 属性値2(連続)
属性値1 属性1 は 属性値2(連続)
属性値 カテゴリ名 へ
カテゴリ名(商品名) の 属性値 属性
属性値1 属性1 で 属性値1 属性2
属性 も 属性値(連続可)
属性値 属性
属性 も 属性 も 属性値(連続可)
属性1 を 属性値2 属性2 カテゴリ名(商品名)
属性 は カテゴリ名(商品名) に 属性値
属性1 を 属性値1 属性2 を 属性値2
商品名(カテゴリ名) の 属性値1 属性1 に 属性値(連続可)
属性 の 属性値
商品名(カテゴリ名) が 属性値 属性
属性値 を 属性値(連続可)
属性 が 属性値(連続可)
属性値1 属性1 が 属性値2(連続可)
属性 を 属性値 カテゴリ名(商品名)
カテゴリ名(商品名) の 属性 へ
カテゴリ名(商品名) に 属性値 属性
カテゴリ名(商品名) を 属性値(連続可)
カテゴリ名(商品名) に 属性値(連続可)
カテゴリ名(商品名) が 属性値 属性

6 おわりに

本研究では商品の属性と属性値に着目し、それらを用いてキャッチコピーを作成するシステムを構築した。しかし、本システムで作成されたキャッチコピーが適格であるか、そしてシステムとしての操作性やユーザーにかける負担に関する評価はまだ行われていないため、今後の課題としてアンケートを行っていく予定である。

参考文献

- [1] 松平智史, 萩原将文: 電子化辞書と遺伝的プログラミングを用いたキャッチコピー作成支援システム, IEEJ Trans, Vol.124, No.1, pp.164-169. (2004)
- [2] 松平智史, 萩原将文: 対話型遺伝的プログラミングと電子化辞書を用いたキャッチコピー作成支援システム, IEEJ Trans, Vol.125, No.4, pp.616-622. (2005)
- [3] 酒井とし夫: 売れるキャッチコピーがスラスラ書ける本. 日本能率協会マネジメントセンター. (2007)
- [4] カタログ・チラシ キャッチコピー大百科, ビエブックス. (2008)