

評判情報検索システムの試作と評価

立石健二 石黒義英 福島俊一
NECインターネットシステム研究所

2V-1

1. はじめに

筆者らは、ユーザの意思決定に有効な情報を提供することを目的として、インターネットから評判情報を取得する技術を開発している。本研究では、インターネットに存在する人々の意見をさまざまな観点で分類・分析したものを評判情報と呼んでおり、このような形式に意見をまとめることがユーザの意思決定に役立つと考える。すなわち、インターネット上に分散して存在する人の意見を見つけ出して、その意見を分類・分析する技術が本研究の重要な課題となる。分類・分析としては、例えば、各意見を肯定又は否定の評価毎に分類することなどをさす。

インターネット上の意見をユーザが見つかる従来の方法として、まず、一般のサーチエンジンでキーワード検索する方法がある。しかしこの方法では、効率的に意見を収集することが難しい。また、商品情報に限定した検索エンジンとして ShopBot[2]等があるが、この技術で抽出対象とするものは価格・スペックのような定型の情報であり、同じ手法を意見のような非定型情報に適用することはできなかった。

筆者らは、上記の課題を解決するシステムとして、インターネット上の Web ページから意見を含むページを探し出し意見を抽出し、それを肯定又は否定に分類するシステムを試作し、前回報告した[1]。本稿では、この試作システムの評価実験を中心に報告する。

2. 試作した評判情報検索システム

試作システムは、トップページで商品名と商品カテゴリを入力すると、その商品に関する人の意見を一覧で表示する。商品に関する意見は商品購入やマーケティングの際に重要である。各意見には肯定否定別に分類したアイコンを表示し、また、検索結果は意見の構文的な意見らしさ(適正值)の順にソートして表示する。

試作システムの処理の流れを図 1 を用いて、簡単に説明する。なお、この詳細は[1]に記載している。まず、

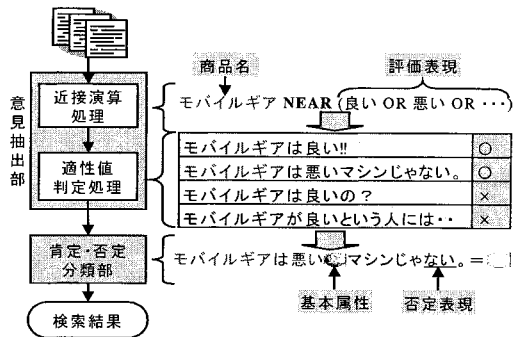


図 1 処理の流れ

ID	適正值判定ルール
1	商品名_*(は が も)*評価表現_ (商品名と評価表現が「は」「が」「も」を挟んで存在する)
2	商品名_*(. . ? !)*評価表現_評価表現_*(。 . ? !)*商品名_ (商品名と評価表現が別の文にある)

表 1 適正值判定ルールの例

意見抽出部は、入力された商品名と、あらかじめ用意した「良い」や「悪い」等の評価を示す表現である評価表現を近接演算して、両者が一定の距離以内にある場合は、両者を含む文字列を意見として抽出する(近接演算処理)。評価表現は、商品カテゴリ毎にあらかじめ辞書として用意しておく。

次に、意見抽出部は、抽出した意見の構文的な特徴から意見らしさを適正值として求める(適正值判定処理)。表 1 のような正規表現で表記された各適正值判定ルールを用いている。図 1 の例では、上の二つは入力された商品に対して意見らしいとして高い適正值が、逆に下の二つには低い適正值が与えられる。

最後に、肯定・否定分類を行い、検索結果を表示する。肯定・否定分類は、あらかじめ定めた評価表現の基本属性に対して、評価表現の近くに「ない」のような否定表現が出現する場合は属性を反転するようにしている。

3. 評価実験

2 節で述べた試作システムの意見抽出部の精度評価のために、9 個の商品(コンピュータ 5、書籍 4)について、従来のキーワード検索と試作システムの適合率と再現率を比較した。

3. 1 実験方法

母集団の Web ページは、一般のサーチエンジンで商品名をキーワードとした検索結果から、各商品についてランダムに 100 前後の Web ページを選択した合計 1087 ページとした。この母集団の Web ページの内、その商品に対する意見を含むページ (適合ページ) は 173 ページ存在した。したがって、キーワード検索の適合率は 15.9%(173/1087)、再現率は 100%である。

評価表現辞書に登録されている評価表現の数は、コンピュータの分野で 140 種類、書籍の分野で 97 種類であり、両分野とも、肯定・否定の表現を同程度用意した。また、近接演算処理の近接距離は、評価表現を基準として前方 80Byte、後方 40Byte とした。さらに、適性値判定処理における適性値判定ルールは 12 個である。

3. 2 実験結果と考察

図 2 に、適性値を 1 以上、2 以上、3 以上、3.3 以上、3.5 以上、4.0 以上とした場合の試作システムの適合率と再現率をプロットしたグラフを示す。

適性値 1 以上のプロットは、試作システムによるすべての検索結果を意味し、近接演算処理の段階における適合率は 48.0%、再現率は 48.6%であった。適合率は、適性値判定処理により適性値が大きくなるにつれ上昇した。表 2 の検索結果の数から適性値が 3.3 以上の検索結果は、試作システムの全検索結果の上位 17.1%を示すことがわかる。したがって、本システムは、検索結果の上位 17.1%において適合率が 86.6%という精度を持つ。同様に、検索結果の上位 38.3%において適合率が 76.1%という精度を持ち、適合率の面ではキーワード検索を大きく上回っている。特に、近接演算検索に適性値判定処理を加えた効果大きい。

一方、再現率については、48.0%であり、キーワード検索で見つかる意見の約半分が試作システムでは検索もれとなった。近接演算処理によって抽出できなかった意見を分析すると、検索もれの原因は、用意した評価表現辞書に存在しない表現が現れた場合と、近接演算の近接距離内に商品名と評価表現が存在しなかった場合の 2 つであった。前者については、「★★★」「80 点」のような記号・数値表現や、「やばい」「とろい」のような口語調の表現で評価を示すことによる登

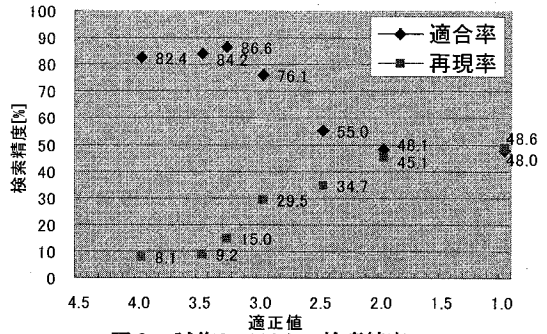


図 2 試作システムの検索精度

適性値	4.0	3.5	3.3	3.0	2.5	2.0	1.0
(a)	17	19	30	67	109	162	175
(b)	9.7%	10.9%	17.1%	38.3%	62.3%	92.6%	100%

(a): 表記の適性値以上の試作システムの検索結果数
 (b): 試作システムの全検索結果数(175)に対する(a)の割合

表 2 試作システムの検索結果数

録もれが多く、今後効率的に評価表現を収集する仕組みが必要である。一方後者については、Web ページのタイトルやページ内の段落の見だしに商品名が存在する時に、そのタイトル配下のページ全体又は見だし配下のある段落全体で、商品名の表記が省略されるために近接演算では抽出できない意見が多かった(ex 掲示板サイト)。このような意見を抽出するためには文章の構造を考慮した近接演算に代わる方式が必要である。

4. おわりに

本稿では、試作した評判情報検索システムの評価実験について述べた。実験から、試作システムの意見抽出部は、適性値判定処理により検索結果の上位 17.1%で適合率 86.6%という、キーワード検索よりも高い精度を持つことが確認できた。今後の課題は、収集した意見を分類・分析する機能の強化および、考察で述べた意見収集部の再現率の向上である。

参考文献

- [1] 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", 情報処理学会第 62 回全国大会予稿集(3), p.123-124, 2001
- [2] Doorenbos, R., Etzioni, O., and Weld, D. "A scalable comparison-shopping agent for the World-Wide Web.", In Proceedings of the First International Conference on Autonomous Agents Agents'97(Marina del Rey, Calif., Feb. 5-8). ACM Press, N.Y., 1997, pp.39-48.