4P-1

感情推定に基づくコンテンツ推薦システムに向けた小説書評感情調査*

平良浩嗣† 當間愛晃‡
琉球大学工学部情報工学科†

1 はじめに

本研究における感情推定とは、あるコンテンツを利用した時に利用者の心情に現れる感情種別を推定することを意味する。コンテンツ推薦システムは、利用者にどのような感情を得たいかを入力してもらうことにより感情推定し、その感情を得られるであろうコンテンツを推薦することを目的とする。

今回のシステムでは、推薦対象であるコンテンツには小説、どのような感情を得たいかという入力はテキストを予定している。テキストを入力とした理由は、本研究には感情推定の一手法として、テキストから感情推定が行えるのか検証するという目的があるためである。また、テキストから感情推定を行うことができると、そのテキストの集合である小説文自体への感情推定を適用しやすいと考えたためである。

本稿では、感情推定に基づく推薦システムの概要と、コンテンツから得られる感情を獲得するための書評 (レビュー) を利用した感情アンケート調査について記述する。

2 システム概要と研究内容

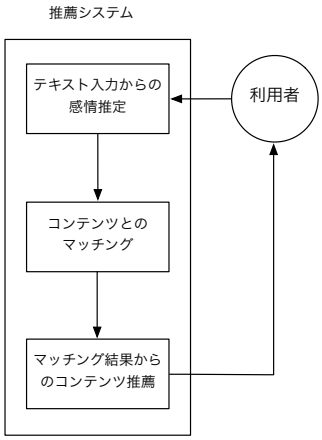

図 1: システム構成案

推薦システムの構成案を図 1 に示す。はじめに利用者に「泣ける本」や「ワクワクする本」など、どういう感情を得たいかテキスト入力してもらう。利用者が入力したテキストを感情推定し、その感情パターンを事前に感情推定しておいた小説のパターンデータとマッチングし、その結果を推薦するというシステムを考えている。なお推薦システムは、コンテンツからの感情パターンの蓄積、短文からの感情推定、データのマッチングといった三つの段階に分けて開発を行う。

現時点では感情パターンの蓄積をする前段階として、書評を書いた人がどういう感情を抱いて記述したと考えられるかを第三者へアンケート調査し、その感情パターンに規則性などが見られるかどうかについての調査を行っている。

利用者の入力したテキストからの感情パターン推定は行っていないが、感情推定の先行研究・関連研究 [1] や、コンテンツに対する感情パターンデータを応用するなどの方法を予定している。

3 小説書評 (レビュー) 感情調査

3.1 調査目的

小説の書評 (レビュー) を第三者が読み、書評を書いた人が小説を読んでどういう感情を得たかを推測し、感情パターンを回答して貰った。複数名で同じ書評を読み、それに対して客観的な視点で作成した感情パターンがどの程度一致するか、あるいはどういった規則性があらわれるかを観察することを目的としている。この調査は書評を使用しても今後の研究が行えるかどうかをはかるものでもある。作品そのものではなく書評の感情パターンを使うため、その書評の感情パターンがばらけてしまうようであれば、今後のシステムに利用するのは難しいと予想され、他の方法を考える必要があると思われるためである。

3.2 調査方法

調査はアンケート形式で行った。調査人数は 6 名である。回答項目には、感情を十種類で表した中村 [2] の十感情を使用している。これは感情を「喜・怒・哀・怖・恥・好・厭・昂・安・驚」の十種類に分けたものである。

回答方法は 0 から 3 までの離散値で、0 が「全く感じていない」、1 が「まあ感じている」、2 が「感じて

*A questionnaire study on novel's reviews for an emotional content recommender
†Hirotsugu TAIRA, Department of Information Engineering, Faculty of Engineering, University of the Ryukyus
‡Naruaki TOMA

Copyright ©2014 Information Processing Society of Japan.
All Rights Reserved.

表 1: 書評 7 と 9 の平均と分散

書評番号		喜	怒	哀	怖	恥	好	厭	昂	安	驚
7	平均	2.83	0	0	1.5	0	1.67	1.17	1.83	0.17	1
	分散	0.14	0.0	0.0	0.93	0.0	1.89	0.47	1.14	0.14	0.67
9	平均	0.33	0.33	0.5	0.17	0	0.33	0.5	0.17	0	0
	分散	0.22	0.22	0.25	0.14	0.0	0.22	0.25	0.14	0.0	0.0

いる」、3 が「強く感じている」というようになっている。いくつの感情をどの値で選んでも良いとした。また、それぞれどういう感情なのかという説明も行っている。

小説 1 冊について書かれた書評 (レビュー)20 件を筆者が収集し、各書評について回答してもらった。この時、回答者の主観で答えるのではなく、書評を書いた人がどういう感情を得たと思われるかを客観的に回答してもらった。

3.3 調査結果および考察

各書評の各感情ごとに平均値を出し、どの程度結果がばらけているのかを見るために、各感情に注目して、分散を出した。結果の分かりやすかった書評 7 と 9 の平均と分散を表 1 に示す。

書評の全体的に、分散が比較的大きかったのは好・昂の 2 つの感情であった。この 2 つは平均も高めの値が出ており、このことから、数人は高い感情値をつけたものの、他の人が低い感情値や感情値 0 をつけているため、分散が大きくなると予想される。好・昂よりは平均も分散も小さめであったが、厭・驚も同様の傾向が見られた。逆に、怒・哀・安などは、書評全体で平均値が小さく、分散も小さめであった。これは回答者がこれらの感情に小さな値をつけているためであると考えられる。喜は、平均値は大きいが、分散は少なかった。これは殆どの回答者が喜に大きな感情値をつけて回答しているためであると考えられる。このことから、書評においてメインとなって現れている感情であると推測できる。

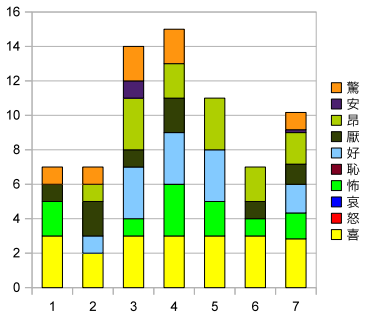

図 2: 書評 7 の回答者別と平均の積み上げ棒グラフ

次に、書評 7 と 9 の、回答者別の積み上げ棒グラフを図 2 と図 3 に示す。1 から 6 が各回答者の、7 は平均値の感情値の積み上げである。表 1 では書評 7 は分散

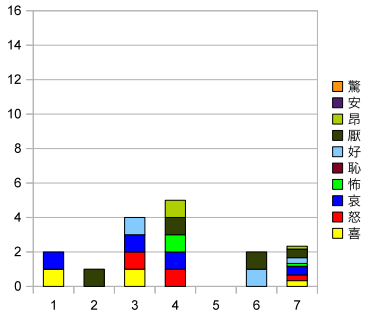

図 3: 書評 9 の回答者別と平均の積み上げ棒グラフ

が大きめに、書評 9 は分散が小さくなっているが、より一様な結果は書評 7 であることが分かる。ここから、分散の大小だけでは結果のばらけかたを観察するには不十分であることが考えられる。

この結果より、ばらけかたの程度についてもう少し調査してみる必要があると考える。主観ではあるが、回答を棒グラフ化したものから、ある程度の一致や規則性はあると感じているので、今後の研究に書評 (レビュー) を使用することはできると考えている。しかし、今回の結果のように分散が大きくなることが考えられるため、あまりにも他と違う回答ははじくなどの処理が必要になると考えられる。

4 今後の予定

今回の調査分析では書評 (レビュー) の使用は妥当であるとまでは判断できなかった。今後は更に分析を続け、使用が妥当であると判断できた際には、書評へ感情パターンを付加するデータ収集の自動化を目指す。

また、テキストからの感情推定の部分も、先行研究などでは単一感情の推定が主流であり、複数感情がどの比率で存在するかといった研究はなされていないため、この問題の解決方法を模索する必要がある。

参考文献

[1] 江崎大嗣, 小町守, 松本裕治:“感情軸における感情極性制約を用いたマルチラベル感情推定”, 言語処理学会第 19 回年次大会 発表論文集, pp.244-247(2013)

[2] 中村明:“感情表現辞典” 東京堂出版 (1993)

Copyright ©2014 Information Processing Society of Japan.
All Rights Reserved.