

タグ付けされた評価表現の分析ツールの開発

唯野良介[†] 嶋田和孝[†] 遠藤 勉[†][†]九州工業大学情報工学部知能情報工学科

1 はじめに

近年、Weblog や Web 掲示板等の普及により、個人が製品やサービスなどに対するレビューを投稿する機会が増えている。これに伴い、自然言語処理において、評価情報を対象とした処理や分析の重要性が増している [1]。Kaji ら [2] は大量の HTML 文書から大規模でかつ汎用的な評価表現辞書を構築している。この辞書はドメイン非依存であるため、非常に有効であるが、例えば、あるドメインに特化したシステムを構築したい場合には、そのドメインに依存した辞書を作成することも重要である。

本研究では、タグ付けされたコーパスを利用し、そのコーパスを分析することで、効率的に、より多くの新たなドメイン依存のタグ付けを支援するシステムの構築を目指す。

2 データ

本論文での対象は、Shimada ら [3] が評価文書分類で使ったゲームレビュー文書である。この評価文書は、各記事ごとに 7 つの評価項目が含まれている。評価項目は「オリジナリティ (o)」、「グラフィックス (g)」、「音楽 (m)」、「熱中度 (a)」、「満足度 (s)」、「快適度 (c)」、「難易度 (d)」である。この文書の一部 (約 5,000 文) について 2 名のアナテータ (A_1, A_2) が、その文に各評価項目に関連する評価表現が含まれているかをタグ付けした。タグ付けされた文章数はそれぞれ A_1 :3,446 文、 A_2 :1,589 文である¹。

3 分析ツール

システムに現在実装されている分析ツールの機能は以下の 3 つである。

3.1 一致度の算出

複数のアナテータによって評価表現とされた文章や同一のタグが付けられた表現を得ることで、より信頼度の高い評価情報を得ることができる。そこで、各アナテータがどのような文に対してどの種類のタグを振ったかを分析し、アナテータ間のタグ付けの一致度を算出する。一致度に関しては Cohen の κ 値を採用した。一致度のみでなく、各評価文の比較結果も表形式で出力する (図 1)。また、評価項目ごとの特徴を見るために、評価項目ごとに一致度とタグが振られた数を分析し、グラフ化することが可能である。

3.2 頻出表現の抽出

評価項目に関連した表現を抽出するための分析を行う。抽出する際の特徴として単語の頻度に注目する。評価表現に対して形態素解析を行い、各評価項目別に単語の出現数を求める。ツール上では出現した単語とその出現数を表形式で見ることが可能である。

3.3 タグの自動付与

3.2 節で得られた単語の頻度情報を用いて、簡易的なタグの自動付与を行う。処理の流れは以下ようになる。

1. 得られた単語の中から、ある閾値以上の頻出単語のみを抽出し、頻出単語群を作成する。
2. 作成した頻出単語群と一致する単語を含む文を、新たな評価情報としてタグ付けを行う。

¹ A_1 と A_2 の文章数に差があるのは、 A_1 はこのデータを他研究で使用した際に何度か修正・追加を行ったためである。

番号	行	ドキュメント1	ドキュメント2	一致タグ	非一致タグ
1	12	画像がカクカクしない	画像がカクカクしない	[op]	
2	12	スムーズなのでこれは嬉しい	スムーズなのでこれは嬉しい	[op]	[ap cp]
3	12	アイテムの使いが第で、最後の最...	最後の最後まで順位がわからないから...	[op]	[ap]
4	12	W i e f i での対戦は白熱	W i e f i での対戦は白熱	[ap]	
5	12	C P U 対戦とは一味違うバトルを...	C P U 対戦とは一味違うバトルを激し...	[op]	
6	12	D S 本伴しか持っていない人も対...	D S 本伴しか持っていない人も対戦が...	[op]	
7	12	音のマリオカードシリーズにあっ...	音のマリオカードシリーズにあった...	[op]	[op]
8	13	ミッションランがある	ミッションランがある	[op]	[op sp]
9	13	秘は好きだった		[sp]	
10	16	アイテムのトゲゾー甲羅が、翼に...	アイテムのトゲゾー甲羅が、翼に入ら...	[on sn]	
11	16	1位とその周辺のカーとのみ相う...	1位とその周辺のカーとのみ相う今回...	[on sn]	
12	16	6 4 にあった、1位がけて器面	甲羅に離れたらクラッシュするパター...	[on]	[sn]

図 1: 各評価表現の比較結果画面

4 評価実験

頻出表現の抽出、タグの自動付与によって得られた結果の有効性を検証した。具体的には、付与された新たな評価項目 (タグ) が、妥当であるか否かを著者が判断した。妥当であれば、タグの付け漏れを発見したことになり、このシステムが有効であることがいえる。実際のタグの自動付与は図 2 のようになる。

既にアナテータによって熱中度タグ (a) が付与されている 291 文²について本ツールを適用したところ、107 文に新たなタグが付与された。この結果より、提案手法がタグの付け漏れ検出に有効であることが確認された。しかし、107 文のうち、妥当であると判断された文の数は 49 文 (57.6%) であり、その正解率は十分とはいえない。この問題を解消するために、作成したツールへの適合性フィードバックの適用などが考えられる。

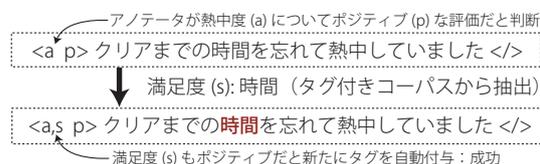


図 2: タグの自動付加例

5 おわりに

本研究ではタグ付け支援システムのためのコーパス分析ツールを作成した。タグ付け支援として、分析結果から得られた単語の頻度情報を基に新たなタグの自動付与を行った。しかし、その正解率は十分とはいえない。今後は適合性フィードバックの適用といった手法を用いて、より正確で有効なシステムを構築したい。

参考文献

- [1] 乾孝司, 奥村学. テキストを対象とした評価情報の分析に関する研究動向. 自然言語処理, Vol. 13, No. 3, pp. 201-242, 2006.
- [2] N. Kaji and M. Kitsuregawa. Building lexicon for sentiment analysis from massive collection of HTML documents. In Proceedings of EMNLP-CoNLL, pp. 1075-1083, 2007.
- [3] K. Shimada and T. Endo. Seeing several stars: a rating inference task for a document containing several evaluation criteria. Proceedings of PAKDD 2008, pp. 1006-1014, 2008.

² 本来はタグ付けされていない文や他のタグについても評価すべきだが、本論文では、その予備実験として、特定のタグについてのみを評価対象とした。