Sentiment Classification Techniques for Tracking Literary Reputation*

Maite Taboada[1], Mary Ann Gillies[2] and Paul McFetridge[1]

[1]Department of Linguistics, [2]Department of English

Simon Fraser University

8888 University Dr., Burnaby, BC, V5A 1S6, Canada

E-mail: mtaboada@sfu.ca, gillies@sfu.ca, mcfet@sfu.ca

Abstract

The initial stages of a project tracking the literary reputation of authors are described. The critical reviews of six authors who either rose to fame or fell to obscurity between 1900 and 1950 will be examined and we hope to demonstrate the contribution of each text to the evolving reputations of the authors. We provide an initial report on the use of the semantic orientation of adjectives and their rough position in the text to calculate the overall orientation of the text and suggest ways in which this calculation can be improved Improvements include further development of adjective lists, expansion of these lists and the consequent algorithms for calculating orientation to include other parts of speech, and the use of Rhetorical Structure Theory to differentiate units that make a direct contribution to the intended orientation from those that are contrastive or otherwise make an indirect contribution.

1. Introduction

The objective of our research is to extract information on the reputation of different authors, based on writings concerning the authors. The project aims to create a database of texts, and computational tools to extract content automatically.

Research on opinion and subjectivity in text has grown considerably in the last few years. New methods are being created to distinguish objective from subjective statements in a text, and to determine whether the subjective statements are positive or negative with respect to the particular subject matter. We believe that the methods currently being used to extract subjective opinion, or sentiment, from movie and consumer product reviews (e.g., Gamon, 2004; Hu & Liu, 2004; Turney, 2002) can be applied to literary reviews and other texts concerning author's works.

In this paper, we describe some of the methods currently being used to extract sentiment from text, and explain how we are applying those methods to literary reviews, letters to the editor, newspaper articles, and critical and scholarly publications concerning six authors who were active in the 1900-1950 period. Section 2 provides some background on literary reputation, and how we plan to quantify it. Section 3 discusses sentiment detection, as it has been applied to movie reviews and other present-day reviews of consumer reports. In Section 4, we address the issue of document structure: how important it is to identify the most important parts of the text, and what methods we can use to that end. This project is in its initial stages, and we do not have conclusive results yet. We present, however, the current state of the system in Section 5, and illustrate it with two examples in Section 6. Finally, conclusions and a discussion of future work are found in Section 7.

2. Background

The question of why writers' works, and by extension their literary reputations, fall in and out of critical and popular favour has long fascinated literary critics. In 1905, Marie Corelli was the best-known and most successful novelist in Britain. By 1950 she had been consigned to literary obscu BrTc 1.6823 3L 0.c 0.72823 -188.259.1406 0 Tw (i

* In Proceedings of LREC 2006 Workshop "Towards Computational Models of Literary Analysis", pp. 36-43.

systematic analysis of the records in order to derive conclusions about the "literary opportunities" of women at the turn of the century. Tuchman & Fortin admit, however, "Although our data about the literary opportunities of most women novelists are substantial, our conclusions are based on inferences." (Tuchman & Fortin, 1989: 18). Our project asks similar questions to Tuchman & Fortin and Herrnstein Smith, but we have designed it so that it permits us to combine the aesthetic and evaluative concerns raised by the former with the kinds of quantitative methodology employed by the latter.

The quantitative aspects of the project are based on research in information retrieval and text categorization. We are scanning documents pertaining to the authors in this study into a computer database designed to store them, and we will then analyze these documents automatically for positive and negative content, i.e., the document's overall *sentiment*. This problem has been characterized as one of determining whether the text is "thuh8 60.5 TD 0rTw t e7,çV|5 TD i ,h 4ning whether the text is

Appraisal is a functional framework for describing evaluation in text: how personal feelings, judgement about other people, and appreciation of objects and art are expressed (Martin & White, 2005; White, 2003). Whitelaw and colleagues compiled a list of appraisal words from the literature on appraisal, and extended it automatically by extracting synonyms and related words from WordNet (Fellbaum, 1998) and on-line thesauri. Other researchers have explored this avenue, extracting synonyms using either Pointwise Mutual Information (Turney, 2001) or Latent Semantic Analysis (Landauer & Dumais, 1997). It is unclear which method provides the best results; published accounts vary(Rapp, 2004; Turney, 2001). Word similarity may be another way of building dictionaries, starting from words whose SO we already know. For this purpose, WordNet is a valuable resource, since synonymy relations are already defined (Kamps et al., 2004). Esuli and Sebastiani (2005) also use synonyms, but they exploit the glosses of synonym words to classify the terms defined by the glosses.

Manual and semi-automatic methods, although highly accurate, are not ideal, given that it is time-consuming ththe glo are 0.589 0 .75 -10.5 TD6 (-) Tj 3080.5 TD can Twsibl Tc 087 -0.ndealnt12 Tc -200438 Tc 0 2133 Tj 35. TD TD -0R Tc 1.19

satellite. Some relations are also multinuclear, consisting of two spans that are equal in importance. The nucleus

towards the end of the text. We also take negation into account, changing the sign of an adjective in the scope of a negating word (e.g., *not*, *no*, *nor*, *neither*

0i2T 25-9.75 T714O.015 Tc60.0087 signnparticular a

results but is limited by the size and accuracy of the list of adjectives used, the accuracy of the algorithm used to identify adjectives, the ability of the algorithm to recognize the context in which the adjective appears (including the presence of negating elements and where the adjective appears in the text), the contribution to the sentiment of the text by words of other parts of speech, and the overall discourse structure of the text. Each of these limitations suggests fruitful avenues of research.

We are engaged in developing algorithms for automatically developing adjective dictionaries. Future research will expand this effort to include semantic orientation dictionaries for nouns and verbs as well. As these are developed, algorithms for integrating their contribution to the orientation of the text as a whole can be investigated.

An accurate identification of semantic orientation requires analysis of units larger than individual words; it requires understanding of the context in which those words appear. To this end, we intend to use Rhetorical Structure Theory to impose on the text a structure that indicates the relationships among its rhetorical units. In particular, we want to distinguish units that are nuclei from those that are satellites so that their respective contributions can be appropriately calculated.

Finally, since the overall structure of a text is often Tw (the cT72ng alg3 of a text .c05kFe3io8.0352 ss91l941ements a203ctive)

Evaluation. New York: Palgrave.

Pang, B., Lee, L. & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using Machine Learning techniques. Proceedings of Conference on Empirical Methods in NLP (pp. 79-86).

Rapp, R. (2004). A freely available automatically generated thesaurus of related words. Proceedings of 4th International Conference on Language Resources and Evaluation (LREC 2004). Lisbon, Portugal.

Remplin, C. (1995). Feminism and the Politics of Literary Reputation. Lawrence: University of Kansas Press.

Soricut, R. & Marcu, D. (2003). Sentence level discourse parsing using syntactic and lexical information. Proceedings of Human Language Technology and North American Association for Computational Linguistics Conference (HLT-NAACL'03). Edmonton, Canada.

Stone, P.J. (1997). Thematic text analysis: New agendas for analyzing text content. In C. Roberts (Ed.), Text Analysis for the Social Sciences. Mahwah, NJ: Lawrence Erlbaum.

Stone, P.J., Dunphy, D.C., Smith, M.S. & Ogilvie, D.M. (1966). The General Inquirer: A Computer Approach to Content Analysis. Cambridge, MA: MIT Press.

Taboada, M. (2004a). Building Coherence and Cohesion: Task-Oriented Dialogue in English and Spanish. Amsterdam and Philadelphia: John Benjamins.

Taboada, M. (2004b). The genre structure of bulletin board messages. Text Technology, 13(2), 55-82.

Taboada, M., Anthony, C. & Voll, K. (2006). Creating semantic orientation dictionaries. Proceedings of 5th International Conference on Language Resources and Evaluation (LREC) (to appear). Genoa, Italy.

Taboada, M. & Grieve, J. (2004). Analyzing appraisal automatically. Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text (AAAI Technical Report SS-04-07) (pp. 158-161). Stanford University, CA.

Tuchman, G. & Fortin, N.E. (1989). Edging Women Out: Victorian Novelists, Publishers, and Social Change. New Haven: Yale University Press.

Turney, P. (2001). Mining the Web for synonyms: PMI-IR versus LSA on TOEFL. Proceedings of the 12th European Conference on Machine Learning (ECML-2001). Freiburg, Germany.

Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of 40th Meeting of the Association for Computational Linguistics (pp. 417-424).

Turney, P. & Littman, M. (2002). Unsupervised learning of semantic orientation from a hundred-billion-word corpus (No. ERB-1094, NRC #44929): National Research Council of Canada.

Turney, P. & Littman, M. (2003). Measuring praise and criticism: Inference of semantic orientation from association. ACM Transactions on Information Systems, 21(4), 315-346.

White, P.R.R. (2003). An Introductory Course in Appraisal Analysis, from http://www.grammatics.com/appraisal

Whitelaw, C., Garg, N. & Argamon, S. (2005). Using Appraisal groups for sentiment analysis. Proceedings of ACM SIGIR Conference on Information and Knowledge Management (CIKM 2005) (pp. 625-631). Bremen, Germany.

Wiebe, J., Wilson, T., Bruce, R., Bell, M. & Martin, M. (2004). Learning subjective language. Computational Linguistics, 30(3), 277-308.