

HEEM, a Complex Model for Mining Emotions in Historical Text

Inger Leemans,

Faculty of Humanities, VU University Amsterdam

Email: i.b.leemans@vu.nl

Janneke M. van der Zwaan,

Netherlands eScience Center, Amsterdam

Email: j.vanderzwaan@esciencecenter.nl

Isa Maks,

Faculty of Humanities, VU University Amsterdam

Email: e.maks@vu.nl

Erika Kuijpers,

Faculty of Humanities, VU University, Amsterdam,

Email: erika.kuijpers@vu.nl

Recently, emotions and their history have become a focus point for research in different academic fields (Matt and Stearns 2013; Plamper 2015; Boddice 2014). At the same time, sentiment analysis and opinion mining have become important research areas, both within and outside academia. So far, however, many techniques are aimed at fitting relatively simple emotion models (positive/negative emotion, or limited sets of at most of six or seven ‘basic’ emotions (e.g., ‘anger’, ‘disgust/contempt’, ‘fear’, ‘interest’, ‘joy’, ‘love’, ‘sadness’, and ‘surprise’) (Buitinck et.al. 2015). In addition, these simple models are almost exclusively applied to contemporary, and/or web-based texts (e.g., Strapparava and Mihalcea 2008; Yang, Lin, and Chen 2007).

However, simple emotion models are not sufficient for Digital Humanities scholars who are interested in research questions about changes in emotional expressions over time. Answering these questions requires more complex, historically motivated emotion models. Also, because historical (literary) text differs significantly from contemporary, and/or web-based text, e.g., with respect to spelling, it is not clear to what extent modern approaches to emotion mining work on historical text. These two issues are addressed in this paper.

This paper presents a new model for emotion mining, resulting from the research project ‘Embodied Emotions’. This project aims: 1. to trace historical changes in emotion expression and in the embodiment of emotions, and 2. to develop methods to trace these changes in sizeable corpuses of digitized texts.

To meet these challenges, we have developed the Historic Embodied Emotion Model (HEEM), built on a test case of 29 plays Dutch language theatre plays written between 1600 and 1800 and

annotated manually with HEEM labels for emotions and body terms. In this paper we present this model and compare it with other sentiment mining techniques, e.g. off the shelf linguistic analysis software LIWC (Linguistic Inquiry and Word Count) (Pennebaker 2001) and a version of LIWC that has been adapted for the analysis of Dutch historical texts

We will conclude that, although different forms of sentiment mining have their value and use, for emotion mining and analysis of embodied emotions in historical texts, HEEM sets a new standard.

References

Boddice, R. (2014) 'The affective turn: historicising the emotions'. In: C. Tileag and J. Byford (ed.), *Psychology and history: Interdisciplinary explorations*, pp. 147–156. Cambridge: Cambridge University Press.

Buitinck, L., Van Amerongen, J., Tan, E. and De Rijke (2015), M. 'Multi-Emotion Detection in User-Generated Reviews', *ECIR 2015: 37th European Conference on Information Retrieval*.

Matt, Susan J., and Peter N. Stearns (eds.) (2013) *Doing Emotions History*. University of Illinois Press.

Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001) *Linguistic Inquiry and Word Count: LIWC2001*. Mahwah, NJ: Erlbaum Publishers (www.erlbaum.com).

Plamper, Jan. 2015. *The History of Emotions: An Introduction*. Oxford: Oxford University Press.

Strapparava, C. and Mihalcea, R. (2008) 'Learning to identify emotions in text', *Proceedings of the 2008 ACM symposium on Applied computing*, pp.1556–1560.

Yang, C., Lin, K.H. and Chen, H.H. (2007) 'Emotion classification using web blog corpora', *IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 275–278.