

A cognitive study of subjectivity extraction in sentiment annotation

Abhijit Mishra¹ Aditya Joshi^{1,2,3} Pushpak Bhattacharyya¹

¹IIT Bombay, India

²Monash University, Australia

³IITB-Monash Research Academy, India

{abhijitmishra, adityaj, pb}@cse.iitb.ac.in

Abstract

Existing sentiment analysers are weak AI systems: they try to capture the *functionality* of human sentiment detection faculty, without worrying about how such faculty is realized in the *hardware* of the human. These analysers are agnostic of the actual cognitive processes involved. This, however, does not deliver when applications demand order of magnitude facelift in accuracy, as well as insight into characteristics of sentiment detection process.

In this paper, we present a cognitive study of sentiment detection from the perspective of strong AI. We study the sentiment detection process of a set of human “sentiment readers”. Using eye-tracking, we show that on the way to sentiment detection, humans first extract subjectivity. They focus attention on a subset of sentences before arriving at the overall sentiment. This they do either through “anticipation” where sentences are skipped during the first pass of reading, or through “homing” where a subset of the sentences are read over multiple passes, or through both. “Homing” behaviour is also observed at the sub-sentence level in complex sentiment phenomena like sarcasm.

1 Introduction

Over the years, supervised approaches using polarity-annotated datasets have shown promise for SA (Pang and Lee, 2008). However, an alternate line of thought has co-existed. Pang and Lee (2004) showed that for SA, instead of a document in its entirety, an extract of the subjective sentences alone can be used. This process of generating a subjective extract is referred to as subjectivity extraction. Mukherjee and Bhat-

tacharyya (2012) show that for sentiment prediction of movie reviews, subjectivity extraction may be used to discard the sentences describing movie plots since they do not contribute towards the speaker’s view of the movie.

While subjectivity extraction helps sentiment classification, the reason has not been sufficiently examined from the perspective of strong AI. The classical definition of strong AI suggests that a machine must be perform sentiment analysis in a manner and accuracy similar to human beings. Our paper takes a step in this direction. We study the cognitive processes underlying sentiment annotation using eye-fixation data of the participants. Our work is novel in two ways:

- We view documents as a set of sentences through which sentiment changes. We show that the nature of these polarity oscillations leads to changes in the reading behavior.
- To the best of our knowledge, the idea of using eye-tracking to validate assumptions is novel in case of sentiment analysis and many NLP applications.

2 Sentiment oscillations & subjectivity extraction

We categorize subjective documents as *linear* and *oscillating*. A *linear* subjective document is the one where all or most sentences have the same polarity. On the other hand, an *oscillating* subjective document contains sentences of contrasting polarity (*viz.* positive and negative). Our discussions on two forms of subjectivity extraction use the concepts of linear and oscillating subjective documents.

Consider a situation where a human reader needs to annotate two documents with sentiment. Assume that the first document is linear subjective - with ten sentences, all of them positive. In

case of this document, when he/she reads a couple of sentences with the same polarity, he/she begins to assume that the next sentence will have the same sentiment and hence, skips through it. We refer to this behavior as *anticipation*. Now, let the second document be an oscillating subjective document with ten sentences, the first three positive, the next four negative and the last three positive. In this case, when a human annotator reads this document and sees the sentiment flip early on, the annotator begins to carefully read the document. After completing a first pass of reading, the annotator moves back to read certain crucial sentences. We refer to this behavior as *homing*.

The following sections describe our observations in detail. Based on our experiments, we observe these two kinds of subjectivity extraction in our participants: subjectivity extraction as a result of *anticipation* and subjectivity extraction as a result of *homing* - for linear and oscillating documents respectively.

3 Experiment Setup

This section describes the framework used for our eye-tracking experiment. A participant is given the task of annotating documents with one out of the following labels: positive, negative and objective. While she reads the document, her eye-fixations are recorded.

To log eye-fixation data, we use Tobii T120 remote eye-tracker with Translog(Carl, 2012). Translog is a freeware for recording eye movements and keystrokes during translation. We configure Translog for reading with the goal of sentiment.

3.1 Document description

We choose three movie reviews in English from IMDB (<http://www.imdb.com>) and indicate them as D0, D1 and D2. The lengths of D0, D1 and D2 are 10, 9 and 13 sentences respectively. Using the gold-standard rating given by the writer, we derive the polarity of D0, D1 and D2 as positive, negative and positive respectively. The three documents represent three different styles of reviews: D0 is positive throughout (linear subjective), D1 contains sarcastic statements (linear subjective but may be perceived as oscillating due to linguistic difficulty) while D2 consists of many *flips* in sentiment (oscillating subjective).

It may seem that the data set is small and

may not lead to significant findings. However, we wished to capture the most natural form of sentiment-oriented reading. A larger data set would have weakened the experiment because: (i) Sentiment patterns (linear v/s subjective) begin to become predictable to a participant if she reads many documents one after the other. (ii) There is a possibility that fatigue introduces unexpected error. To ensure that our observations were significant despite the limited size of the data set, we increased the number of our participants to 12.

3.2 Participant description

Our participants are 24-30 year-old graduate students with English as the primary language of academic instruction. We represent them as P0, P1 and so on. The polarity for the documents as reported by the participants are shown in Table 1. All participants correctly identified the polarity of document D0. Participant P9 reported that D1 is confusing. 4 out of 12 participants were unable to detect correct opinion in D2.

3.3 Experiment Description

We obtain two kinds of annotation from our annotators: (a) sentiment (positive, negative and objective), (b) eye-movement as recorded by an eye-tracker. They are given a set of instructions beforehand and can seek clarifications. This experiment is conducted as follows:

1. A complete document is displayed on the screen. The font size and line separation are set to 17pt and 1.5 cm respectively to ensure clear visibility and minimize recording error.
2. The annotator verbally states the sentiment of this sentence, before (s)he can proceed to the next.
3. While the annotator is reading the sentence, a remote eye-tracker (Model: Tobii TX 300, Sampling rate: 300Hz) records the eye-movement data of the annotator. The eye-tracker is linked to Translog II software (Carl, 2012) in order to record the data. A snapshot of the software is shown in figure 1. The dots and circles represent position of eyes and fixations of the annotator respectively. Each eye-fixation that is recorded consists of: coordinates, timestamp and duration. These three parameters have been used to generate sentence progression graphs.

Document	Orig	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
D0	+ve	+ve	+ve	+ve	+ve	+ve	+ve	+ve	+ve	+ve	+ve	+ve	+ve
D1	-ve	-ve	+ve	-ve	-ve	-ve	-ve	-ve	-ve	-ve	Neu/-ve	-ve	-ve
D2	+ve	+ve	+ve	-ve	+ve	+ve	Neu	+ve	Neu	Neu	+ve	+ve	+ve

Table 1: Polarity of documents as perceived by the writer (original) and the participants +ve, -ve and Neu represent positive, negative and neutral polarities respectively.

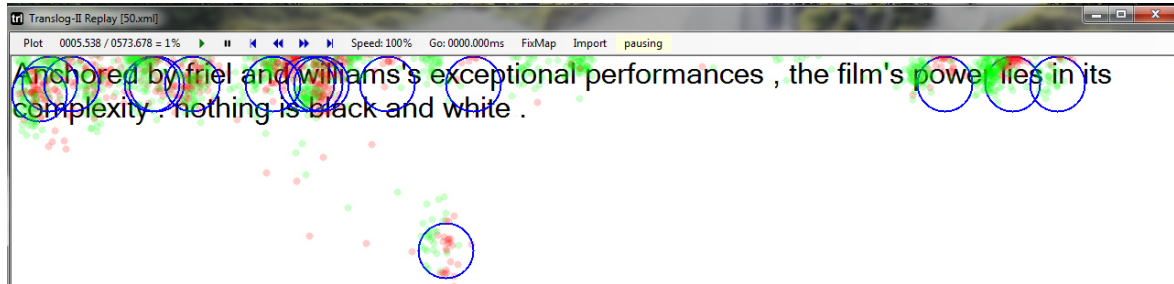


Figure 1: Gaze-data recording using Translog-II

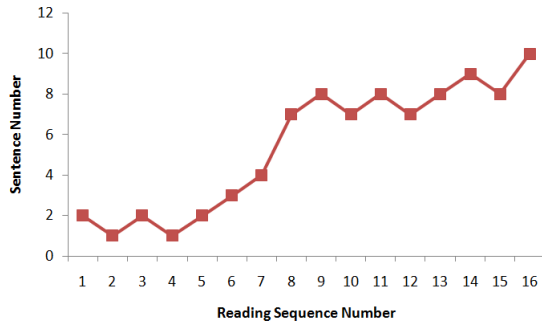


Figure 2: Sentence progression graph for participant P7 document D0

4 Observations: Subjectivity extraction through anticipation

In this section, we describe a case in which participants skip sentences. We show that anticipation of sentiment is linked with subjectivity extraction.

Table 2 shows the number of unique and non-unique sentences that participants read for each document. The numbers in the last column indicate average values. The table can be read as: participant P1 reads 8 unique sentences of document D0 (thus skipping two sentences) and including repetitions, reads 26 sentences. Participant P0 skips as many as six sentences in case of document D1.

The number of unique sentences read is lower than sentence count for four out of twelve participants in case of document D0. This skipping is

negligible in case of document D1 and D2. Also, the average non-unique sentence fixations are 21 in case of D0 and 33.83 for D1 although the total number of sentences in D0 and D1 is almost the same. This verifies that participants tend to skip sentences while reading D0.

Figure 2 shows sentence progression graph for participant P7. The participant reads a series of sentences and then skips two sentences. This implies that anticipation behaviour was triggered after reading sentences of the same polarity. Similar traits are observed in other participants who skipped sentences while reading document D0.

5 Observations: Subjectivity extraction through homing

This section presents a contrasting case of subjectivity extraction. We refer to a reading pattern as *homing*¹ when a participant reads a document completely and returns to read a selected subset of sentences. We believe that during sentiment annotation, this subset is the subjective extract that the user has created in her mind. We observe this phenomenon in reading patterns of documents D1 and D2. The former contains sarcasm because of which parts of sentences may appear to be of contrasting polarity while the latter is an oscillating subjective document.

¹The word is derived from missile guidance systems. The definition² of homing is “the process of determining the location of something, sometimes the source of a transmission, and going to it.”

Document		P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	Avg.
D0	Non-unique	9	26	23	17	18	18	35	16	33	19	15	23	21
	Unique	8	8	10	10	10	10	10	8	10	8	10	10	
D1	Non-unique	5	23	46	13	15	44	35	26	56	57	40	46	33.83
	Unique	3	9	9	9	9	9	8	9	9	9	9	9	
D2	Non-unique	36	29	67	21	23	51	64	48	54	59	73	80	50.42
	Unique	13	13	13	13	13	13	13	13	13	13	13	13	

Table 2: Number of unique and non-unique sentences read by each participant

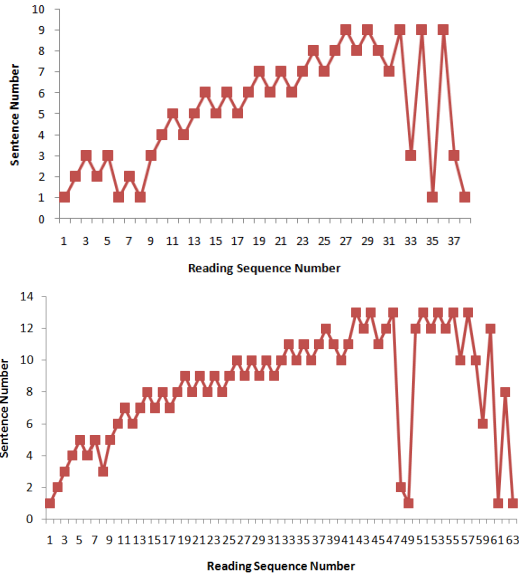


Figure 3: Sentence progression graph of participant P2 for document D1 (left) and document D2 (right)

Figure 3 shows sentence progression graphs of participant P2 for documents D1 and D2. For document D1, the participant performs one pass of reading until sequence number 30. A certain subset of sentences are re-visited in the second pass. On analyzing sentences in the second pass of reading, we observe a considerable overlap in case of our participants. We also confirm that all of these sentences are subjective. This means that the sentences that are read after sequence number 30 form the subjective extract of document D1.

Similar behaviour is observed in case of document D2. The difference in this case is that there is less overlap of sentences read in the second pass among participants. This implies, for oscillating subjective documents, the subjective extract is user/document-specific.

It may be argued that fixations corresponding

Participant	TFD-SE (secs)	PTFD (%)	TFC-SE
P5	7.3	8	21
P7	3.1	5	11
P9	51.94	10	26
P11	116.6	16	56

Table 3: Reading statistics for second pass reading for document D1; TFD: Total fixation duration for subjective extract; PTFD: Proportion of total fixation duration = (TFD)/(Total duration); TFC-SE: Total fixation count for subjective extract

to second pass reading are stray fixations and not subjective extracts. Hence, for the second pass reading of document D1, we tabulate fixation duration, fixation count and proportion of total duration in Table 3. The fixation duration and fixation count are both recorded by the eye-tracker. The fixation counts are substantial and the participants spend around 5-15% of the total reading time in the second pass reading. We also confirm that all of these sentences are subjective. This means that these portions indeed correspond to subjective extracts as a result of homing.

6 A note on linguistic challenges

Our claim is that regression after reading an entire document corresponds to the beginning of a subjective extract. However, we observe that some regressions may also happen due to sentiment changes at the sub-sentence level. Some of these are as follows.

1. **Sarcasm:** Sarcasm involves an implicit flip in the sentiment. Participant P9 does not correctly predict sentiment of Document D1. On analyzing her data, we observe multiple regressions on the sentence ‘Add to this mess some of the cheesiest lines and concepts, and

there you have it; I would call it a complete waste of time, but in some sense it is so bad it is almost worth seeing.’ This sentence has some positive words but is negative towards the movie. Hence, the participant reads this portion back and forth.

2. **Thwarted expectations:** Thwarted expectations are expressions with a sentiment reversal within a sentence/snippet. Homing is observed in this case as well. Document D2 has a case of thwarted expectations from sentences 10-12 where there is an unexpected flip of sentiment. In case of some participants, we observe regression on these sentences multiple times.

7 Related Work

The work closest to ours is by Scott et al. (2011) who study the role of emotion words in reading using eye-tracking. They show that the eye-fixation duration for emotion words is consistently less than neutral words with the exception of high-frequency negative words. Eye-tracking³ technology has also been used to study the cognitive aspects of language processing tasks like *translation* and *sense disambiguation*. Dragsted (2010) observe co-ordination between reading and writing during human *translation*. Similarly, Joshi et al. (2011) use eye-tracking to correlate fixation duration with polysemy of words during *word sense disambiguation*.

8 Conclusion & Future work

We studied sentiment annotation in the context of subjectivity extraction using eye-tracking. Based on how sentiment changes through a document, humans may perform subjectivity extraction as a result of either: (a) anticipation or (b) homing. These observations are in tandem with the past work that shows benefit of subjectivity extraction for automatic sentiment classification.

Our study is beneficial in three perspectives: (i) Sentiment classifiers may use interaction between sentiment of sentences. Specifically, this can be modeled using features like sentiment run length (i.e. maximal span of sentences bearing same

sentiment) or sentiment flips (i.e. instances where consecutive sentences bear opposite polarity), (ii) Crowd-sourced sentiment annotation can devise variable pricing models based on our study. Based on anticipation and homing information about documents, documents can be grouped into difficulty categories and priced accordingly.

Acknowledgment

We thank Tobii Corporation for lending us their eye-tracker for this study, and our annotators from CFILT, IIT Bombay. Aditya is funded by the TCS Research Fellowship Program.

References

- Bo Pang and Lillian Lee. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts In Proceedings of the *ACL*, 271-278.
- Bo Pang and Lillian Lee. 2008. Opinion Mining and Sentiment Analysis *Foundations and Trends in Information Retrieval*, 2008, vol. 2, nos.12 1135.
- B Dragsted. 2010. Co-ordination of reading and writing processes in translation. *Contribution to Translation and Cognition*. Shreve, G. and Angelone, E.(eds.) *Cognitive Science Society*.
- Michael Carl. 2012. *Translog-II: A Program for Recording User Activity Data for Empirical Reading and Writing Research*. In Proceedings of the Eight International Conference on Language Resources and Evaluation, European Language Resources Association.
- Scott G. , ODonnell P and Sereno S. 2012. Emotion Words Affect Eye Fixations During Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2012, Vol. 38, No. 3, 783792.
- Salil Joshi, Diptesh Kanojia and Pushpak Bhattacharyya. 2013. *More than meets the eye: Study of Human Cognition in Sense Annotation*. NAACL HLT 2013, Atlanta, USA.
- Subhabrata Mukherjee and Pushpak Bhattacharyya. 2012. WikiSent : Weakly Supervised Sentiment Analysis Through Extractive Summarization With Wikipedia *European Conference on Machine Learning (ECML PKDD 2012)*, Bristol, U.K.,

³Related Terms:

Eye-fixation: Long stay of visual gaze on a single location

Regression: Revisiting a previously read segment

Sentence Progression Graph: Graph showing reading sequence of sentences