# A Pragma-Semantic Analysis of the Emotion/Sentiment Relation in Debates

Valerio Basile[1], Elena Cabrio[2], Serena Villata[2], Claude Frasson[3], and Fabien Gandon[1]

[1] Université Côte d'Azur, Inria, CNRS, I3S, France
firstname.lastname@inria.fr
[2] Université Côte d'Azur, CNRS, Inria, I3S, France
firstname.lastname@unice.fr
[3] University of Montreal, Canada
frasson@iro.umontreal.ca

**Abstract.** In the last years, emotions recognition tools have become more and more popular, aiming at detecting the emotions of human actors while performing different intelligent tasks by means of headsets and facial emotions detection tools. In addition to this kind of technology, when participants interact with each others by means of textual exchanges, sentiment analysis techniques, from the natural language processing research area, are exploited to detect the polarity of the exchanged messages. Investigating how these two connected components interacts and can support each other towards a better emotions and sentiment detection is a relevant but unexplored research challenge. In this paper, we start from a dataset of debate interactions annotated with the emotions of the involved participants, captured by means of EEG headsets and a facial emotions recognition tool, and the argumentative structures of the debates, and we compare this information to the polarity of the proposed textual arguments, retrieved through a sentiment analysis algorithm. A pragma-semantic analysis of the obtained results is provided, along with a discussion of the potential future work.

## 1 Introduction

Analyzing and detecting the emotions felt by people engaged in a debate is becoming more and more important in Artificial Intelligence. Reasoning techniques such as argumentation theory [8], based on rational postulates and critical thinking, have started to be connected with personal and emotional information, like for instance in [6, 2, 4]. The idea is that, to have an overall view of a debate, several components have to be considered at the same time, i.e., argumentation, emotions, sentiment, and they influence each other in a mutual way. In this paper, we start from a dataset of textual arguments annotated with the emotions felt by the participants of an experimental session through an Emotiv EPOC EEG headset and the facial expression real-time frame-by-frame analysis software FaceReader [1], and we apply sentiment analysis techniques to analyse the natural language textual arguments proposed in the debates. More precisely, we address a pragma-semantic analysis of the the obtained mismatches, e.g., captured emotion *happy* and polarity of the argument *negative*, and we discuss the preliminary results of this study. The analysis we propose has the aim to challenge current sentiment analysis techniques over a gold standard of textual argument whose associated *real* emotion is annotated.

Up to our knowledge, this is the first time such a kind of analysis is proposed, while it is important to start considering the interplay of these three components, i.e., argumentation, emotions and sentiment, that are indispensable to model cognitive agents.

## 2 The Pragma-Semantic Analysis

In this section, we first describe the dataset of textual arguments we use to address our pragma-semantic analysis (Section 2.1), and second, we provide some insights about sentiment analysis techniques (Section 2.2). Finally, we discuss the results of our ongoing pragma-semantic analysis (Section 2.3).

### 2.1 Dataset

In [1], we presented an open dataset to compare and analyze emotion detection in an argumentation session. More precisely, the goal of our empirical analysis was to study the link between the argumentation people address when they debate with each other, and the emotions they feel during these debates. We conducted an experiment aimed at verifying our hypotheses about the correlation between the positive/negative emotions emerging when positive/negative relations among the arguments are put forward in the debate. For more details about the participants and the results of this study, we refer the reader to [1].

The dataset of debates consists of 10 debates carried out by 4 participants at a time (20 total participants), excluding the moderator. The dataset is composed of three main layers: *(i)* the basic annotation of the arguments proposed in each debate (i.e. the annotation in xml of the debate flow downloaded from the debate platform); *(ii)* the annotation of the relations of support and attack among the arguments; and *(iii)* starting from the basic annotation of the arguments, the annotation of each argument with the emotions felt by each participant involved in the debate. Table 1 shows some statistics on the dataset[1].

An example, from the debate about the topic "Religion does more harm than good" where arguments are annotated with emotions (i.e., the third layer of the annotation of the textual arguments we retrieved), is as follows:

```
<argument id="30" debate_id="4" participant="4"
  time-from="20:43" time-to="20:43"
  emotion_p1="neutral" emotion_p2="neutral"
```

[1] The dataset of annotated arguments is available here: http://project.inria.fr/seempad/datasets/

Table 1: Distribution of the length and number of words across the debates in the dataset.

| **ID** | Topic | # of arguments | # of words |
|---|---|---|---|
| 1 | Ban animal testing | 49 | 1219 |
| 2 | Go nuclear | 39 | 972 |
| 3 | Housewives should be paid | 42 | 1001 |
| 4 | Religion does more harm than good | 46 | 970 |
| 5 | Advertising is harmful | 71 | 1033 |
| 6 | Bullies are legally responsible | 71 | 1208 |
| 7 | Distribute condoms in schools | 68 | 1190 |
| 8 | Encourage fewer people to go to the university | 55 | 1341 |
| 9 | Fear government power over Internet | 41 | 886 |
| 10 | Ban partial birth abortions | 41 | 1014 |
| 11 | Use racial profiling for airport | 31 | 562 |
| 12 | Cannabis should be legalized | 43 | 605 |


```
  emotion_p3="neutral" emotion_p4="neutral">
    Indeed but there exist some advocates of the devil
    like Bernard Levi who is decomposing arabic countries.
</argument>
<argument id="31" debate_id="4" participant="1"
  time-from="20:43" time-to="20:43"
  emotion_p1="angry" emotion_p2="neutral"
  emotion_p3="angry" emotion_p4="disgusted">
    I don't totally agree with you Participant2: science
    and religion don't explain each other, they tend to
    explain the world but in two different ways.
</argument>
<argument id="32" debate_id="4" participant="3"
  time-from="20:44" time-to="20:44"
  emotion_p1="angry" emotion_p2="happy"
  emotion_p3="surprised" emotion_p4="angry">
    Participant4: for recent wars ok but what about wars
    happened 3 or 4 centuries ago?
</argument>
```


Figure 1 shows one of the visualizations of the debates that we used to explore the data set (number 9, “Fear government power over Internet”). The participants are color-coded (gray always indicates the moderator). The number inside the nodes represents the identifiers of the arguments, in chronological order. A single directed edge indicates a support relation, where the tail node supports the head node, while a double-lined edge indicates an attack. The shape of the argument nodes shows the associated sentiment (see Section 2.2).

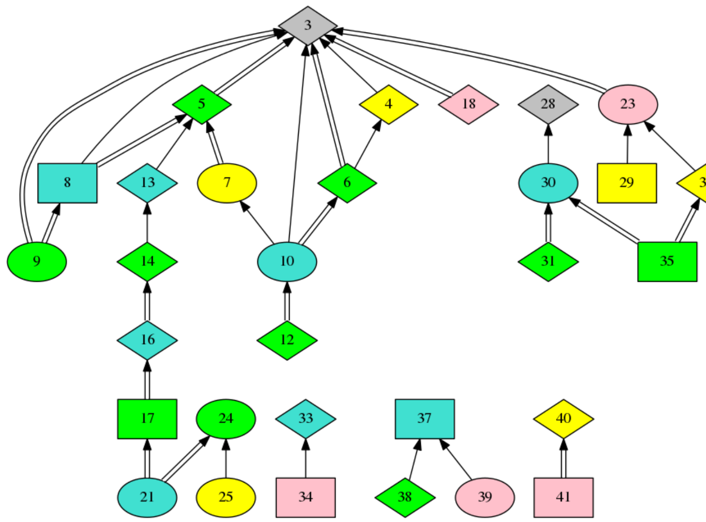

Fig. 1: Visual depiction of the debate "Fear government power over Internet", highlighting the relation between arguments and their detected sentiment.

### 2.2 Sentiment Analysis

The goal of *sentiment analysis* (sometimes referred to as *opinion mining*) is detecting the polarity, whether positive, neutral, or negative, of the attitude contained in a natural language utterance. A typical first step is to determine whether a statement is objective or subjective, and then only in the latter case one can proceed to identify its polarity. However, often only the second task is performed, thus collapsing objective statements and a neutral attitude.

The last years have seen an enormous increase in research on developing sentiment analysis systems of various sorts that employ several natural language processing techniques. Solutions range from simple lookups in polarity or affection resources, i.e., databases where a polarity score is associated to terms, to more sophisticated models built through supervised, unsupervised, and distant learning involving various sets of features [3].

Several approaches are found in literature for polarity detection. The simplest route is detecting the specific words which are known to express a positive, negative or neutral feeling. For example, [5] use a lexicon projection strategy yielding predictions which significantly correlate with polls. Deeper linguistic analysis has been proven to improve the performance of sentiment analysis systems [7]. However, accurate processing can be hard on texts such as social media messages and text chats, which are short, rich in abbreviations and often contain syntactic or spelling mistakes.

In order to extract the sentiment from the text of the debates, we submit every argument separately to Alchemy API[2], a free service provided by IBM via

[2] http://www.alchemyapi.com/api/sentiment-analysis

an online HTTP REST Web service. For any given text, Alchemy API returns a sentiment label, either *neutral*, *negative* or *positive*, and a sentiment score ranging from -1 (totally negative sentiment) to 1 (totally positive sentiment).

We correlate the sentiment of the arguments with the emotion of the participants who wrote them. Moreover, we only consider the primary emotion.

The confusion matrix in Table 2 shows how the sentiment compares to the emotions detected by the FaceReader. It is clear that, while the FaceReader is quite conservative in assigning non-neutral emotions, the output of Alchemy API reveals highly polarized sentiment, with 212 arguments categorized as negative, 140 as positive, and a relative minority of 104 as neutral. The emotion *happy* is observed in the dataset only twice as primary emotion, and in neither case is the emotion detected for the proponent of the argument.

Table 2: Correlation between the sentiment extracted by Alchemy API and the emotions of the participants in the debates.

| **Sentiment** | **Emotion** | | | | | |
|---|---|---|---|---|---|---|
| | angry | disgusted | happy | neutral | scared | surprised |
| negative | 19 | 17 | 0 | 172 | 1 | 3 |
| neutral | 7 | 14 | 0 | 78 | 3 | 2 |
| positive | 13 | 12 | 0 | 112 | 2 | 1 |

In order to study the correlation of sentiment and emotions, we proceed to manually map the polarity of emotions and sentiment, following this simple scheme:

- angry, disgusted, scared → negative
- happy, surprised → positive
- neutral → neutral

This mapping results in the confusion matrix shown in Table 3. The sentiment extracted by Alchemy API matches the polarity of the emotion in 116 cases (37 negative, 1 positive, 78 neutral, about 25% of all arguments). Out of all the cases of mismatch, in 30 cases the polarities are inverted: in 27 cases a positive sentiment is associated to a negative emotion, and in 3 cases the opposite happens.

We believe that the cases where a full mismatch is observed are the most interesting to explore the relation between the sentiment found in the text, and the emotions felt by the participants in a debate. Therefore, we inspected them one by one and report our findings in the next section.

### 2.3 Discussion

In the three cases where the sentiment is negative, the emotion is "surprised". The arguments seem to be genuinely of a negative nature ("Racial profiling is

Table 3: Correlation between the sentiment extracted by Alchemy API, and the polarity of the emotions of the participants in the debates.

| **Sentiment** | **Emotion polarity** | | |
|---|---|---|---|
| | negative | neutral | positive |
| negative | 37 | 172 | 3 |
| neutral | 24 | 78 | 2 |
| positive | 27 | 112 | 1 |

a prototype which is unacceptable I think. "), therefore the mismatch could be the result of either a misclassification of the emotion by the FaceReader or the naïve mapping between sentiment polarity and emotions shown in Section 2.2, where surprise is classified as a positive emotion.

In eight cases, the polarity of the argument has been wrongly predicted. Given the statistical nature of the majority of state-of-the-art language analysis tools, including the software for sentiment analysis, a certain rate of errors is always to be expected. Interestingly, here most of the errors are due to the word choice, in particular to the use of certain nouns that are typically associated to positive sentiment. Examples of this phenomenon include, for instance: "Do we know what would be a good way to make someone not a bully? i.e. to teach "respect"?". The word "respect" in particular is associated with a positive polarity by the system. While the sentiment score of the original message is 0.33, replacing the word "respect" with a neutral word like "math" gives a totally neutral message, according to Alchemy API. Another word that seem to confuse the automatic classification of sentiment is "thanks", in phrases such as "thanks to ...".

In nine cases, the argument is a reply (possibly an attack or a support) to an argument proposed by another participant. Since we fed isolated messages to the sentiment analysis component, it is natural that the analysis of such cases will not be accurate, since the system is missing important contextual information.

Two of the mismatching arguments are phrased in a quite convoluted way that contributes to confusing the classifier. One of such examples recites: "Of course, from university you can learn a lot of stuff, have better degree, but don't think that such degree will be helpful to get a better job later." Note that the pattern "of course X, but Y" is difficult to interpret by automatic language analysis without resorting to some logical interpretation of the text structure.

Finally, in three cases, the sentiment seems to be genuinely positive. In two of them, the corresponding emotion is "scared", which is seldom observed across the entire dataset. Since there is no element of fear in the text of these arguments, we tend to attribute these mismatches to noise in the original data.

The six remaining examples include a mix of the aforementioned phenomena or their features are too sparse to draw proper conclusions. It is worth noticing that one argument is interestingly of ironic nature ("RFID ALL THE PEOPLE!"), a case where the positive polarity of the literal meaning of the text is correctly associated with a negative emotion.

## 3 Conclusions

In this paper, we have presented some preliminary results of a pragma-semantic analysis over a dataset of textual arguments from human debates annotated with their emotions, and on which we have then applied sentiment analysis techniques. More precisely, we have studied the cases where a mismatch holds between the sentiment captured from the textual arguments through sentiment analysis, and the emotion(s) detected from the participants proposing such arguments in the debate. Some patterns emerge from this analysis. However, our intuition is that adding also the argumentative information into the loop, i.e., considering not only the detected emotions but also the attack and support relations among the arguments, would return useful information to enrich such patterns. This is our main direction for future work, i.e., to study the interplay of argumentation, sentiment analysis and emotions in debates, in order to detect patterns of information from the argumentation and emotion components to improve the performance of sentiment analysis techniques, and enrich their results.

## References


1. Benlamine, S., Chaouachi, M., Villata, S., Cabrio, E., Frasson, C., Gandon, F.: Emotions in argumentation: an empirical evaluation. In: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015. pp. 156–163 (2015)
2. Grosse, K., González, M.P., Chesñevar, C.I., Maguitman, A.G.: Integrating argumentation and sentiment analysis for mining opinions from twitter. AI Commun. 28(3), 387–401 (2015), http://dx.doi.org/10.3233/AIC-140627
3. Liu, B.: Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers (2012), http://dx.doi.org/10.2200/S00416ED1V01Y201204HLT016
4. Medellin, R., Reed, C., Hanson, V.L.: Spoken interaction with broadcast debates. In: Computational Models of Argument - Proceedings of COMMA 2014. pp. 51–58 (2014)
5. O'Connor, B., Balasubramanyan, R., Routledge, B.R., Smith, N.A.: From tweets to polls: Linking text sentiment to public opinion time series. In: Cohen, W.W., Gosling, S. (eds.) ICWSM. The AAAI Press (2010), http://dblp.uni-trier.de/db/conf/icwsm/icwsm2010.html#OConnorBRS10
6. Pak, A., Paroubek, P.: Twitter as a corpus for sentiment analysis and opinion mining. In: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (eds.) Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, 17-23 May 2010, Valletta, Malta. European Language Resources Association (2010), http://www.lrec-conf.org/proceedings/lrec2010/summaries/385.html
7. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retr. 2(1-2), 1–135 (Jan 2008), http://dx.doi.org/10.1561/1500000011
8. Rahwan, I., Simari, G.R.: Argumentation in Artificial Intelligence. Springer Publishing Company, Incorporated, 1st edn. (2009)