Toward a unifying model for Opinion, Sentiment and Emotion information extraction

Amel Fraisse, Patrick Paroubek

LIMSI-CNRS
Bât 508, Université Paris-Sud XI
fraisse@limsi.fr, pap@limsi.fr

Abstract
This paper presents a logical formalization of a set 20 semantic categories related to opinion, emotion and sentiment. Our formalization is based on the BDI model (Belief, Desire and Intention) and constitutes a first step toward a unifying model for subjective information extraction. The separability of the subjective classes that we propose was assessed both formally and on two subjective reference corpora.

Keywords: Sentiment Analysis, Opinion mining, Affective Lexicon

1. Introduction
There is no global model for opinions, sentiments and emotions (OSE) annotation and information extraction in texts. Existing models were either devised only for a part of the subjective expression spectrum, e.g. opinions about commercial products (Dave et al., 2003), or either with the aim to provide a biologically plausible explanation of human behaviour, either to serve as a basis for artificial agent specification (Meyer, 2006). In this paper we address the problem extracting from all the existing models a common ground suitable to cover opinions, sentiments and emotions that can be deployed in an information extraction task. After reviewing the literature, we present a generic formal model based on a minimal Belief, Desire and Intention agent model (Sadèk, 1992) for OSE. The proposed model provide a logical formalization of 20 semantic classes of opinions, sentiments and emotions. Next we show how on the one hand the emotion model designed for the I2B2 2011 task2 evaluation campaign about suicide notes analysis (Pak et al., 2012) and on the other hand the opinion model developed by the DOXA project (Paroubek et al., 2010) both map onto our model. Then we have proved the theoretical and practical disjunction of the 20 OSE semantic classes. We have used the Coq proof assistant (Bertot and Castéran, 2004) to prove the theoretical separability. Then we have evaluated the practical disjunction of different classes on the reference corpus of DOXA showing that the classes we have defined correspond to separable sets of linguistic lexicon.

2. Related work
In the early 1970s, Ekman found evidence that humans share six basic emotions: happiness, sadness, fear, anger, disgust and surprise (Ekman, 1970). Few tentative efforts to detect non-basic affective states, such as fatigue, anxiety, satisfaction, confusion, or frustration, have been also made (Kapoor et al., 2007). The dimensional approach (Osgood et al., 1957), in turn, represents emotions as coordinates in a multi-dimensional space. For both theoretical and practical reasons, more and more researchers like to dene emotions according to two or more dimensions. An early example is Russell’s circumplex model (Russell, 1979), which uses the dimensions of arousal and valence to identify 150 affective labels. Similarly, Whissell considers emotions as a continuous 2D space whose dimensions are evaluation and activation (Whissell, 1989).

(Cambria et al., 2012) proposed an affective categorisation model primarily inspired by Plutchik’s studies on human emotions (Plutchik, 2001). Such model represents affective states both through labels and through four independent but concomitant affective dimensions (Pleasantness, Attention, Sensitivity, Aptitude). In total, he identified 24 emotion labels. Other research works are focused on the formalization of such emotional categories. During the last 20 years, several logical models have been developed for modeling cognitive autonomous agents that are suited for this purpose. Most of these so-called agent logics belong to the class of belief-desire-intention, that describe autonomous agents on the intentional level in terms of beliefs, desires (goals), intentions and possibly other related attitudes. More precisely, BDI models are formal models that arise from the combination of several modal logics: a temporal or a dynamic logic used to capture the dynamic nature of agents, and logics for the mental states of belief, desire and intention. Each of the modal operators is given a precise syntactical definition in terms of a set of axioms, and a precise semantics in terms of possible worlds models. Formalizations of belief-desire analyses of emotions in agent logics are of relatively recent origin. Most of these formalizations focus on the cognitive and motivational preconditions of emotions.

(Castelfranchi and Lorini, 2003) formalized the belief-desire preconditions of a set of emotions related to expectations (hope, fear, disappointment, and relief) using one of the first BDI model, proposed by Cohen and Levesque (Cohen and Levesque, 1990).

(Meyer, 2006) proposes a logical model of emotions based on KARO, his logic of action, belief and choice. He uses this logic to write generation rules for four emotions: joy, sadness, anger and fear proposed in Oatley and Johnson-Laird’s theory of emotion.
More recently, (Steunebrink et al., 2012) used KARO to formalize the cognitive-motivational preconditions of the 22 emotions considered in the OCC theory. Another formalization of the OCC theory, using an extended version of the Cohen-Levesque logic, was proposed by (Adam et al., 2009).

3. Our model for Emotions, Opinions and Sentiments annotation and information extraction

As we have said above, there is no global model for opinions, sentiments and emotions annotation and information extraction in text. We propose a generic model that can be used to modelize and annotate the whole subjective expression spectrum. Our model divides subjective information into three main categories: affective expressions (emotions), affective-intellective expressions (sentiments) and intellective expressions (opinions). The model associates to each category a set of semantic classes; each semantic class is represented by a generic label and it contains a set of equivalent semantic classes. Such as the generic affective class LOVE that contains affection, care, tenderness, fondness, kindness, attachment, devotion, passion, envy and desire.

3.1. Fine-Grained Opinion/Sentiment/Emotion classes

To define the OSE semantic classes we are based on the DOXA model (Paroubek et al., 2010), it is one of the richest model proposed so far in terms of the number of OSE defined, with 17 semantic categories. We have added the semantic category Love to the OSE expressions of DOXA. We have also modified the semantic class RECOMMANDATION, SUGGESTION by INFORMATION and DEMAND_QUERY by INSTRUCTION in order to be more generic. In total we consider 20 semantic classes for OSE annotation and representation. As described in the table 3, we identified 8 negative emotions (e−): NEGATIVE SURPRISE, DISCOMFORT, FEAR, BOREDOM, DISPLEASURE, SADNESS, ANGER and CONTEMPT, 4 positive emotions (e+): PLEASURE, APPEASEMENT, POSITIVE SURPRISE and LOVE, 1 negative sentiment (s−): SATISFACTION, 1 positive sentiment (s+): INSATISFACTION, 2 positive opinions (o+): AGREEMENT and VALORIZATION, 2 negative opinions (o−): DISAGREEMENT and DEVALORIZATION and 2 neutral classes: INFORMATION and INSTRUCTION. For the Opinion, Sentiment and Emotion (OSE) annotator the notion of time is also very important to express private state. So, we added to the three modal operators, a time function t that associate to an object o one value from the set \{past, present, future\}. (in our case an object is either an event or an action).

| Modal Operator | Mapping |
|----------------|---------|
| Bel_p(E) | “person p believes that E" |
| Des_p(E) | “E is desirable for p” |
| Int_p(A) | “person p intends to do action a” |

3.2. Model Formalism

Our aim is to model opinions, sentiments and emotions in a logic of mental attitudes. Formal logic provides a universal vocabulary with a clear semantics and it allows explanation of person opinion, sentiment and emotion. A given formal definition of emotions may be criticized, but it still has the advantage to be unambiguous. The logic used here is based on the BDI model belief, desire and intention (Sadek, 1992).

We used the \( Bel_p(E) \) operator to express the expectedness or the knowledge of an event \( e \) by a person \( p \). In fact, this operator is important to formalize emotion that are triggered by an expected or non-expected event, such as Negative Surprise or Positive Surprise.

\[
\text{Expected}_p(e) \overset{\text{def}}{=} Bel_p(e) 
\]

\[
\text{Unexpected}_p(e) \overset{\text{def}}{=} \neg Bel_p(e) 
\]

The second operator, we used is \( Des_p(E) \), that expresses the polarity of an event \( e \) for a person \( p \). we consider that if a event \( e \) is positive for a person \( p \) than \( e \) is desirable for \( p \) and vice versa.

\[
\text{Positive}_p(e) \overset{\text{def}}{=} Des_p(e) 
\]

\[
\text{Negative}_p(e) \overset{\text{def}}{=} \neg Des_p(e) 
\]

We also used the \( Int_p(A) \), which is, mostly, associated with opinions, sentiments or emotions having a high intensity. For instance, anger triggers, often, an intention to do an action (run away, be hidden).

\[
t : O \rightarrow \{\text{past, present, future}\}
\]

\[
o \mapsto t(o)
\]

So, in total, we used four attributes (i.e. the three operators: belief, desire and intention and the time function \( t \)) to formalize each OSE of the table 3. For example (as described in table 2), we formalize the negative surprise (\( \text{Neg Surprise}_p \)): a person \( p \) is negatively surprised by an event \( e \) if \( e \) is not desirable for \( p \) (\( \neg Des_p(e) \)) and \( e \) happened (\( t(e) \leq \text{present} \)) and \( e \) is not expected by \( p \) (\( \neg Bel_p(e, t(e)) \)).

We formalize the positive surprise as opposite to the negative surprise: a person \( p \) is positively surprised by an event \( e \) if \( e \) is desirable for \( p \) (\( Des_p(e) \)) and \( e \) happened (\( t(e) \leq \text{present} \)) and \( e \) is not expected by \( p \) (\( \neg Bel_p(e, t(e)) \)).
Table 2: Logical formalization of Emotions, Sentiments and Opinions

3.3. Mapping annotation models: I2B2 and DOXA onto our annotation model

In this section we show that the proposed semantic classes are rich and complete enough to make the mapping possible between the three representations (DOXA to uComp and I2B2 to uComp). (See Table 3).

Table 3: I2B2 and DOXA classes proposed mapping to our annotation model classes.

4. Experiments and results

4.1. Data decription

To investigate how the 20 OSE classes proposed under this work are disjoint, we used two corpora:

DOXA corpus: we used the reference corpus of the DOXA project, such corpus consists of video game review annotated with their corresponding semantic category. For example, the Figure 1 shows an example of a user review (written in French) annotated with the semantic category Anger (in French Colère). Figure 1: An Example of an annotated paragraph extracted from the reference corpus of DOXA.

Table 4 represents the number of paragraph for each semantic category. For instance, there are 76 paragraphs annotated Sadness and 928 paragraphs annotated Displeasure. In the DOXA project, annotations are done in two levels:

- macro, which corresponds to the document level,
- meso, for the paragraph level.

Thus, each annotated document contains at least one paragraph. For our experimentation, we consider each paragraph as a document. Firstly, we construct a corpus with all paragraphs as well as their associated semantic category. Thus we obtained a corpus with 7162 paragraphs, some paragraphs may be annotated with up to 5 semantic categories. From the 7162 paragraphs, 612 of them had several annotations, 609 with 2 annotations and 3 with 3 annotations. And there is 1239 paragraphs annotated as neutral. Then, we grouped all documents per semantic category and extracted all words occurring in these documents. Thus, we construct a generic lexicon for each semantic category (figure 2). In order to estimate the separability of classes, we have plotted the obtained lexicons of the two largest classes, i.e. Valorization and Devalorization classes on a 2-dimension graph using principal component analysis for dimension reduction (Figure 2).

Affective Twitter corpus: we also used the affective twitter corpus constructed by (Fraisse and Paroubek, 2014). Such corpus is collected using the Twitter Search API. It contains subjective tweets annotated with the corresponding semantic category. Since, on twitter, anyone can express their opinions, sentiments or emotion about anything, this corpus is more generic then the DOXA and consequently the lexicon of the corpus has more coverage. Ta-

---

The image contains a page from a document with text and tables. The text appears to be discussing emotion categorization and mapping annotation models, specifically I2B2 and DOXA. The tables and text mention logical formalization of emotions, sentiments, and opinions, and experimental results. The document includes a figure illustrating an annotated paragraph from the DOXA corpus, along with a table representing the number of paragraphs for each semantic category. The text also discusses the construction of a corpus with 7162 paragraphs, some of which have annotations from 2 to 5 semantic categories, and grouped them per category to generate lexicons for each category. There is a note about the affective Twitter corpus constructed by Fraisse and Paroubek, 2014, collected using the Twitter Search API, which contains subjective tweets annotated with corresponding semantic categories. The text mentions that this corpus is more generic compared to the DOXA corpus, as it covers a broader spectrum of emotions and opinions.
Table 4: Characteristics of the reference corpus of DOXA.

| Neg_Surp. | Disc. | Fea. | Bor. | Disp. |
|-----------|-------|------|------|-------|
| Number of: 
parag.     | 31    | 110  | 23   | 182   | 928   |
| Sad. | Ang. | Cont. | Disat. | Dev. |
| Number of: 
parag.     | 76    | 128  | 175  | 0     | 984   |
| disag. | val. | Agr. | Sat. | Pos_Surp. |
| Number of: 
parag.     | 57    | 1814 | 858  | 299   | 144   |
| Appe. | Plea. | Lov. | Inf. | Inst. |
| Number of: 
doc.      | 122   | 432  | 0    | 161   | 13    |

Table 5: Characteristics of the twitter corpus

| Neg_Surp. | Disc. | Fea. | Bor. | Disp. |
|-----------|-------|------|------|-------|
| Number of: 
doc.      | 33    | 47   | 94   | 257   | 369   |
| Sad. | Ang. | Cont. | Disat. | Dev. |
| Number of: 
doc.      | 1042  | 430  | 424  | 406   | 4     |
| disag. | val. | Agr. | Sat. | Pos_Surp. |
| Number of: 
doc.      | 45    | 178  | 87   | 617   | 86    |
| Appe. | Plea. | Lov. | Inf. | Inst. |
| Number of: 
doc.      | 653   | 1527 | 1698 | 0     | 0     |

In order to evaluate the separability of the different semantic categories defined under our model, we wanted to compare different lexicons used by users to express their different affective states (emotions, sentiments and opinions). In fact, we consider that if there is an important overlapping between different lexicons then our classes are not sufficiently separated. As for the DOXA corpus, we constructed a generic lexicon per semantic category by extracting occurring words and removing stop words. Then, we have plotted the obtained lexicons on a 2-dimension graph using principal component analysis for dimension reduction.

4.2. Results

Although we do not make any preprocessing on the data before its analysis, as we can see from the figure 3, the distinction between Valorization lexicon (red crosses) and Devalorization lexicon (blue circle) is easier. The Valorization and Devalorization lexicons are build on a corpus of 400 documents.

In the same way, we plotted lexicons of different semantic classes of the twitter corpus. In order to be significant, we did a side-by-side comparison for all semantic classes. The figure 4 show that Love lexicon (red crosses) and Pleasure lexicon (blue circle) are disjoint.
5. Conclusion

We have presented a generic and formal model for opinions, sentiments and emotions annotation and information extraction in texts. After reviewing the state of the art in terms of opinion mining modeling and opinion mining evaluation, we have presented a logical and formal model to unify the OSE annotation and representation. The model is based on the BDI logics and it consists on 20 semantic classes that we have proved theoretically and practically their disjunction. In a future work, we plan to do a human validation of our model to show that the OSE classes we propose are distinguishable by human annotators, relying on the infrastructure for game with a purpose and crowd sourcing under the uComp project.

6. References

Carole Adam, Andreas Herzig, and Dominique Longin. 2009. A logical formalization of the occ theory of emotions. In Synthese, volume 168. Kluwer Academic Publishers.

Yves Bertot and Pierre Castéran. 2004. In The Coq’Art.

E. Cambria, A. Livingstone, and A. Hussain. 2012. The hourglass of emotions. In Cognitive behavioural systems. Springer Berlin Heidelberg.

Cristiano Castelfranchi and Emiliano Lorini. 2003. Cognitive anatomy and functions of expectations. In European Cognitive Science Conference, Osnabrück, September. Lawrence Erlbaum Associates.

Philip R. Cohen and Hector J. Levesque. 1990. Intention is choice with commitment. In Artificial Intelligence, volume 42. Elsevier Science Publishers Ltd. Essex, March.

Kushal Dave, Steve Lawrence, and David M. Pennock. 2003. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. In Proceeding WWW ’03 Proceedings of the 12th international conference on World Wide Web. ACM, New York, NY, USA.

P. Ekman. 1970. In Universal facial expressions of emotions. Digest, California mental Health Research.

A. Fraisse and P. Paroubek. 2014. Twitter as a comparable corpus to build multilingual affective lexicons. In The 7th Workshop on Building and Using Comparable Corpora, 26-31, May.

A. Kapoor, W. Burleson, and R. Picard. 2007. Automatic prediction of frustration. In International Journal of Human-Computer Studies, volume 65, pages 724–736.

John-Jules Ch. Meyer. 2006. Reasoning about emotional agents. In International Journal of Intelligent Systems, volume 21.

C. Osgood, G. Suci, and P. Tannenbaum. 1957. The measurement of meaning. In University of Illinois Press.

Alexander Pak, Delphine Bernhard, Patrick Paroubek, and Cyril Grouin. 2012. A combined approach to emotion detection in suicide notes. the limsi participation in the i2b2/va 2011 challenge. In In: Biomed Inform Insights. PubMed PMID: 22879766.

Patrick Paroubek, Alexander Pak, and Djamel Mostefa. 2010. Annotations for opinion mining evaluation in the uComp is CHIST-ERA project.
industrial context of the doxa project. In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC)*, Valetta, Malta, May. ELDA.

R. Plutchik. 2001. The nature of emotions. In *American Scientist*, volume 89.

J. Russell. 1979. Affective space is bipolar. In *Journal of Personality and Social Psychology*.

M. David Sadek. 1992. A study in the logic of intention. In B. Nebel, C. Rich, and W. Swartout (eds): *Proceedings third Int. Conf. on Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann Publishers.

Bas R. Steunebrink, Mehdi Dastani, and John-Jules Ch. Meyer. 2012. A formal model of emotion triggers: an approach for bdi agents. In *Synthese. An International Journal for Epistemology, Methodology and Philosophy of Science*, volume 185.

C. Whissell. 1989. The dictionary of affect in language. In *Emotion: Theory, Research, and Experience 4*.