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

Amel Fraisse, Patrick Paroubek

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

We propose in this paper a novel and global model for opinions, sentiments and emotions annotation and information extraction in texts. The proposed model consists in 20 semantic classes that we formalized using the BDI logic. In order to evaluate the disjunction between the different classes, we analyzed lexicons of two subjective corpus. The result shows that the semantic classes defined by our model are separable.

Keywords: Sentiment Analysis, Opinion Mining, Affective Lexicons

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 biologicaly 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 litterature, we present a generic formal model based on a minimal Belief, Desire and Intention agent model (Sadek, 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 developped by the DOXA project (Paroubek et al., 2010) both map onto our model.

Then we have proved the theoretical and pratical disjunction of the 20 OSE semantic classes. We have used the Coq proof assistant (Bertot and Castran, 2004) to prove the theoretical separability. Then we have evaluated the pratical disjunction of different classes on the reference corpus of DOXA showing that the classes we haved 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 Russells 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 Plutchiks 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 logic languages 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 logics*, that describe autonomous agents on the intentional level in terms of beliefs, desires (goals), intentions and possibly other related attitudes. More precisely, BDI logics are formal logic languages 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 logics, proposed by Cohen and Levesque (Cohen and Levesque, 1990).

(Meyer, 2006) proposes a logical modal 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 anf fear proposed in Oatley and Johnson-

Lairds 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 [1]. 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+): *INSATISFATION*, 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 complexity and number of items to consider depends on the annotation task. In its simplest form, the annotator may be asked to provide a binary answer whether a piece of text can be said to express a given OSE or not, in that case the model has only one class. But the annotator may be asked to identify various OSE instances in the text and, depending on the context of the experiment, the number and structural arrangement in the model of the various OSE classes can be quite different from one task to another.

[1]A Semantic Category refers to a ”meaning category”, i.e. a set of opinions, sentiments or emotions which are so semantically close as to be considered indiscernable

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 logics *belief, desire and intention* (Sadek, 1992).

Modal Opertor	Mapping
$Bel_p(E)$	”person p belives that E”
$Des_p(E)$	”E is desirable for p”
$Int_p(A)$	”person p intends to do action a”

Table 1: The Three Basic Modal Operators

We used the $Bel_p(E)$ operator to express the expectedness or the knowledgement 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*.

$$\mathrm{Expected_p}(e) \stackrel{def}{=} \mathrm{Bel_p}(e) \quad (1)$$

$$\mathrm{Unexpected_p}(e) \stackrel{def}{=} \neg Bel_p(e) \quad (2)$$

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.

$$\mathrm{Positive_p}(e) \stackrel{def}{=} \mathrm{Des_p}(e) \quad (3)$$

$$\mathrm{Negative_p}(e) \stackrel{def}{=} \neg Des_p(e) \quad (4)$$

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*).
The notion of time is also very important to express private state. So, we added to the three madal 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).

$$\begin{array}{lllll} t & : & O & \rightarrow & \{past, present, future\} \\ & & o & \mapsto & t(o) \end{array}$$

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 ($Neg\_Surprise_p$) as: a person p is negatively surprised by a an event *e* if *e* is not desirable for *p* ($\neg Des_p(e)$) and *e* happened ($t(e) \leq present$) and *e* is not expected by *p* ($\neg Bel_i(e, t(e))$).
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 ($t(e) \leq present$) and e is not expected by p ($\neg Bel_i(e, t(e))$).

Label	Type	Definition
$Neg\_Surprise_p$	e-	$\exists e, p \mid \neg Des_p(e) \wedge$ $t(e) \leq present \wedge$ $\neg Bel_i(e, t(e))$
$Discomfort_p$	e-	$\exists e, p \mid \neg Des_p(e) \wedge$ $t(e) \leq present \wedge$ $e \Rightarrow a \wedge$ $Int_p(a) \wedge$ $t(a) > present$
$Fear_p$	e-	$\exists e, p \mid \neg Des_p(e) \wedge$ $t(e) \leq present \wedge$ $Bel_p(e)$
$Boredom_p$	e-	$\nexists e, p \mid Des_p(e) \wedge$ $Bel_p(\neg(e))$
$Displeasure_p$	e-	$\exists e, p \mid \neg Des_p(e) \wedge$ $t(e) \leq present \wedge$ $Bel_p((e))$
$Sadness_p$	e-	$\exists e, p \mid Des_p(e) \wedge$ $t(e) \leq present \wedge$ $Bel_p(t(e) > present \vee t(e) = \emptyset)$
$Anger_p$	ex1-	$\exists e, p, a \mid \neg Des_p(e) \wedge$ $t(e) \leq present \wedge$ $Int_p(a)$
$Contempt_p$	e-	$\exists x, p \mid \neg Des_p(x) \wedge$ $Bel_p(x)$
$Pos\_Surprise_p$	e+	$\exists e, p \mid Des_p(e) \wedge$ $Des_p(e) \wedge$ $\neg Bel_i(e, t(e)) \wedge$ $t(e) \leq present$
$Appeasement_p$	e+	$\exists e, p \mid \neg Des_p(e) \wedge$ $t(e) \leq present \wedge$ $e \Rightarrow a \wedge$ $Int_i(a) \wedge$ $t(a) \leq present$
$Pleasure_p$	e+	$\exists e, p \mid Des_p(e) \wedge$ $t(e) \leq present \wedge$ $Bel_p(e)$
$Love_p$	e+	$\exists x, p \mid Des_p(x) \wedge$ $Bel_p(x)$
$Valorization_p$	o+	$\exists x, p \mid Des_p(x) \wedge$ $Bel_p(x) \wedge$ $x \rightarrow a \wedge$ $Inte_p(x) \wedge$ $t(a) \geq present$
$Devalorization_p$	o-	$\exists x, p \mid \neg Des_p(x) \wedge$ $Bel_p(x) \wedge$ $x \rightarrow a \wedge$ $\neg Inte_p(x) \wedge$ $t(a) \geq present$
$Satisfaction_p$	s-	$\exists a, p \mid Des_p(a) \wedge$ $Inte_p(a) \wedge$ $t(a) \leq present$
$Disatisfaction_p$	s-	$\exists a, p \mid Des_p(a) \wedge$ $Inte_p(a) \wedge$ $t(a) > present$
$Agreement_p$	o+	$\exists p_1, p_2, e \mid Des_p 1(e) \wedge$ $Des_p 2(e)$
$DisAgreement_p$	o-	$\exists p_1, p_2, e \mid Des_p 1(e) \wedge$ $\neg Des_p 2(e)$
$Instruction_p$	i	$\exists p_1, p_2, a \mid Int_p 1(Int_p 2(a))$
$Information$	i	$\exists p_1, p_2, e \mid Int_p 1(Bel_p 2(e))$

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).

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 *Colere_Agacement_Irritation_Enervement_Exasperation*).

	DOXA classes	corresponding classes in our model	I2B2 emotion classes
1	NEGATIVE SURPRISE	NEGATIVE SURPRISE	
2	DISCOMFORT	DISCOMFORT	guilt
3	FEAR	FEAR	fear
4	BOREDOM	BOREDOM	
5	DISPLEASURE	DISPLEASURE	abuse
6	SADNESS	SADNESS	sorrow
7	ANGER	ANGER	anger
8	CONTEMPT	CONTEMPT	blame
9	DISATISFACTION	DISATISFACTION	hopelessness
10	DEVALORIZATION	DEVALORIZATION	
11	DISAGREEMENT	DISAGREEMENT	
12	VALORIZATION	VALORIZATION	
13	AGREEMENT	AGREEMENT	
14	SATISFACTION	SATISFACTION	hopefulness pride
15	POSITIVE SURPRISE	POSITIVE SURPRISE	
16	APPEASEMENT	APPEASEMENT	thankfulness forgiveness
17	PLEASURE	PLEASURE	hapiness peacefulness
18		LOVE	love
19	RECOMMANDATION	INFORMATION	information
20	DEMANDE	INSTRUCTION	instruction

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

```
<PSOA polarity="negatif_fort" type="meso" id="d1002.5">
<LINK>
<SEM_CAT offsetStart="2486" offsetEnd="2504"
type="Colere_Agacement_Irritation_Enervement_Exasperation"><![CDATA[carrément
énervante]]></SEM_CAT>
<TOPIC offsetStart="2364" offsetEnd="2374" type="Gameplay"><![CDATA[jauge d’air]]></TOPIC>
</LINK>
<PARAG offsetStart="2169" offsetEnd="2728" id="texte_d1002.5"><![CDATA[ tas de tremplins
qui vous permettront de réaliser tout un tas de tricks et figures, en essayant de bien
retomber. Le but principal étant de finir premier, il faudra constamment surveiller votre
jauge d’air, qui descend aussi vite qu’une bière sur le comptoir d’un pub en Irlande.
Cette jauge est d’ailleurs au final carrément énervante, au point d’en jeter sa manette
ou de tout simplement éteindre la console pour se lancer dans un Mario Kart ou un
Oblivion . Sonic, c’est un payday Malgré les bonnes sensations de vitesse et la bonne
quantité de personnages]]></PARAG>
<ParagSegmentJustif offsetStart="2507" offsetEnd="2578"><![CDATA[au point d’en jeter sa
manette ou de tout simplement éteindre la console]]></ParagSegmentJustif>
</PSOA>
```

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 paragrpah 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 annotatd 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 ocuring 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 [2]. It contains subjective tweets annotated with the corresponding semantic category. Since, on twitter, anyone can express their opinions, sentiments or emotion about anything,

[2]https://dev.twitter.com/docs/using-search

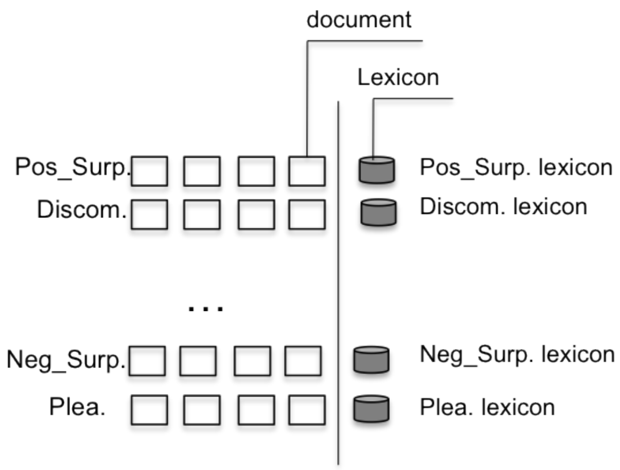

Figure 2: Data setting

	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: parag.	122	432	0	161	13

Table 4: Characteristics of the reference corpus of DOXA.

this corpus is more generic then the DOXA and consequently the lexicon of the corpus has more coverage. Table 5 represents the number of document for each semantic category.

	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

Table 5: Characteristics of the twitter corpus

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 occuring 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.

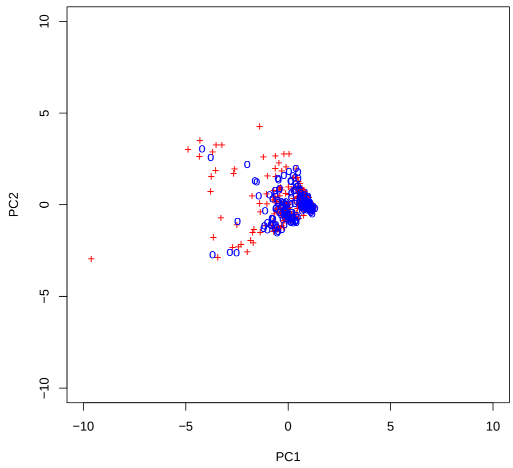

Figure 3: Visualizing samples in 2-dimensions: Valorization (red crosses) vs. Devalorization (blue circles)

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.

Despite, the easier distinction between the two lexicons through different semantic classes, there is always a little intersection between lexicons. Such intersection is explained by the fact that analysed lexicons are generic (i.e. since it contains all occuring words in the documents) and consequently they contain some common neutral words, Named entities, etc.

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 planify 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

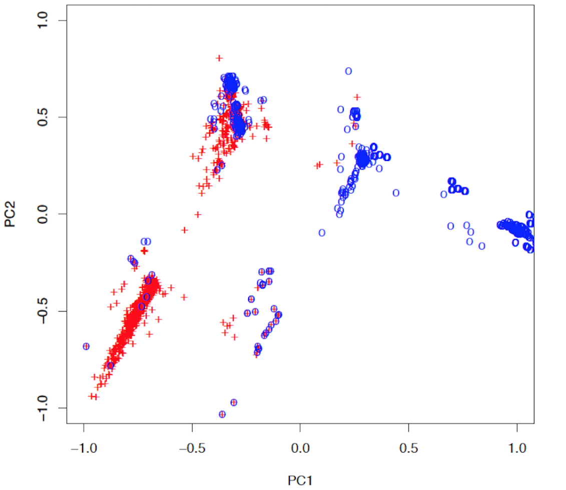

Figure 4: Visualizing samples in 2-dimensions: Love (red crosses) vs. Pleasure (blue circles)

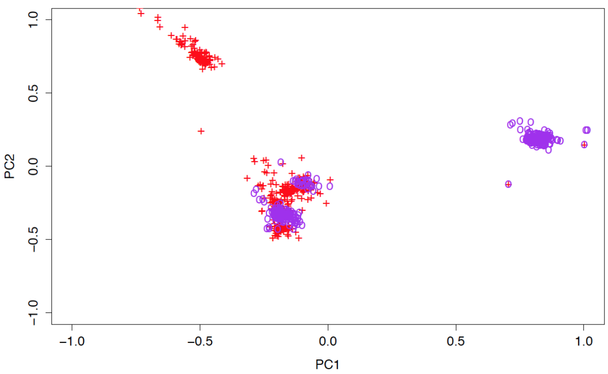

Figure 5: Visualizing samples in 2-dimensions: Anger (red crosses) vs. Displeasure (purple circles)

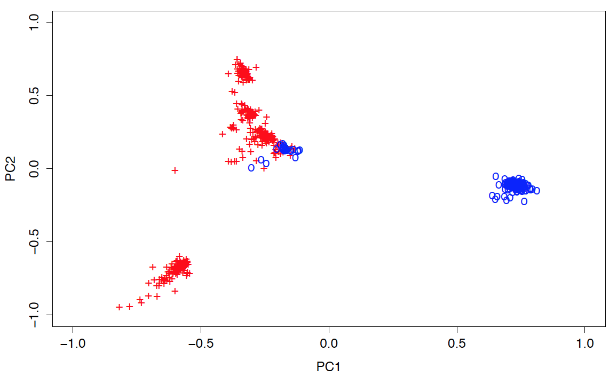

Figure 6: Visualizing samples in 2-dimensions: Anger (red crosses) vs. Boredom (blue circles)

sourcing under the uComp project [3].

[3] uComp is CHIST-ERA project

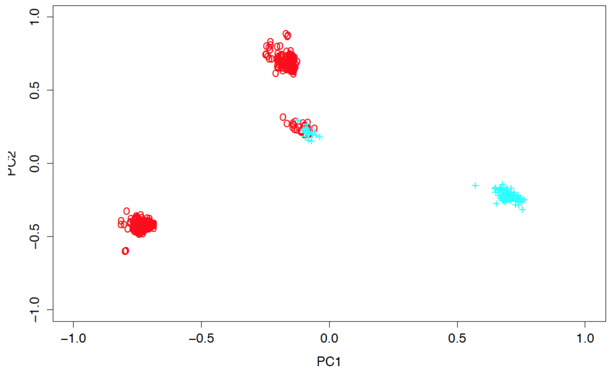

Figure 7: Visualizing samples in 2-dimensions: Boredom (red crosses) vs. Displeasure (cyan circles)

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 Castran. 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 Emilliano Lorini. 2003. Cognitive anatomy and functions of expectations. In *European Cognitive Science Conference*, Osnabrck, 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, page 724736.

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 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*.