The Affective Weight of Lexicon

Carlo Strapparava, Alessandro Valitutti, Oliviero Stock

ITC-irst, 38050, Povo, Trento, Italy
{strappa, alvalitu, stock}@itc.it

Abstract

This paper presents resources and functionalities for the recognition and selection of affective evaluative terms. An affective hierarchy as an extension of the WORDNET-AFFECT lexical database was developed in the first place. The second phase was the development of a semantic similarity function, acquired automatically in an unsupervised way from a large corpus of texts, which allows us to put into relation concepts and emotional categories. The integration of the two components is a key element for several applications.

1. Introduction

All words can potentially convey affective meaning. Each of them, even those more apparently neutral, can evoke pleasant or painful experiences. While some words have emotional meaning with respect to the individual story, for many others the affective power is part of the collective imagination (e.g. words "mum", "ghost", "war" etc.).

Therefore, it is interesting to individuate a way to measure the affective meaning of a generic term. To this aim, we studied the use of words in textual productions, and in particular their co-occurrences with the words in which the affective meaning is explicit. As claimed by Ortony et al. (Ortony et al., 1987), we have to distinguish between words directly referring to emotional states (e.g. "fear", "cheerful") and those having only an indirect reference that depends on the context (e.g. words that indicate possible emotional causes as "monster" or emotional responses as "cry"). We call the former *direct affective words* and the latter *indirect affective words*.

The main contributions of this work consist on (i) the organization of the direct affective words and synsets inside WORDNET-AFFECT, an affective lexical resource based on an extension of WORDNET, and on (ii) a selection function (named *affective weight*) based on a semantic similarity mechanism automatically acquired in an unsupervised way from a large corpus of texts (100 millions of words), in order to individuate the indirect affective lexicon.

Applied to a concept (e.g. a WORDNET synset) and an emotional category, this function returns a value representing the semantic affinity with that emotion. In this way it is possible to assign a value to the concept with respect to each emotional category, and eventually select the emotion with the highest value. Applied to a set of concepts that are semantically similar, this function selects subsets characterized by some given affective constraints (e.g. referring to a particular emotional category or valence).

As we will see, we are able to focus selectively on positive, negative, ambiguous or neutral types of emotions. For example, given "difficulty" as input term, the system suggests as related emotions: IDENTIFICATION, NEGATIVE-CONCERN, AMBIGUOUS-EXPECTATION, APATHY. Moreover, given an input word (e.g. "university") and the indication of an emotional valence (e.g. positive), the system suggests a set of related words through some positive emotional category (e.g. "professor" "scholarship" "achieve-ment") found through the emotions ENTHUSIASM, SYMPATHY, DEVOTION, ENCOURAGEMENT.

These fine-grained kinds of affective lexicon selection can open up new possibilities in many applications that exploit verbal communication of emotions.

2. WORDNET-AFFECT and the Emotional Categories

WORDNET-AFFECT is an extension of WordNet database (Fellbaum, 1998), including a subset of synsets suitable to represent affective concepts. Similarly to our method for domain labels (Magnini and Cavaglià, 2000), we assigned to a number of WordNet synsets one or more affective labels (*a-labels*). In particular, the affective concepts representing emotional state are individuated by synsets marked with the a-label EMOTION. There are also other a-labels for those concepts representing moods, situations eliciting emotions, or emotional responses. WORDNET-AFFECT is freely available for research purpose at http://wndomains.itc.it. See (Strapparava and Valitutti, 2004) for a complete description of the resource.

	# Synsets	# Words	# Senses
Nouns	280	539	564
Adjectives	342	601	951
Verbs	142	294	430
Adverbs	154	203	270
Total	918	1637	2215

Table 2: Number of elements in the emotional hierarchy.

Recently, we extended WORDNET-AFFECT with a set of additional a-labels (i.e. the emotional categories), hierarchically organized, in order to specialize synsets with a-label EMOTION. In a second stage, we introduced some modifications, in order to distinguish synsets according to emotional valence. We defined four addictional a-labels: POSITIVE, NEGATIVE, AMBIGUOUS, NEUTRAL. The first one corresponds to "positive emotions", defined as emotional states characterized by the presence of positive edonic signals (or pleasure). It includes synsets such as joy#1 or enthusiasm#1. Similarly the NEGATIVE a-label identifies "negative emotions" characterized by negative edonic signals (or pain), for example anger#1 or

A-Labels	Valence	Examples of word senses
JOY	positive	noun `joy#1`, adjective `elated#2`, verb `gladden#2`, adverb `gleefully#1`
LOVE	positive	noun `love#1`, adjective `loving#1`, verb `love#1`, adverb `fondly#1`
APPREHENSION	negative	noun `apprehension#1`, adjective `apprehensive#3`, adverb `anxiously#1`
SADNESS	negative	noun `sadness#1`, adjective `unhappy#1`, verb `sadden#1`, adverb `deplorably#1`
SURPRISE	ambiguous	noun `surprise#1`, adjective `surprised#1`, verb `surprise#1`
APATHY	neutral	noun `apathy#1`, adjective `apathetic#1`, adverb `apathetically#1`
NEGATIVE-FEAR	negative	noun `scare#2`, adjective `afraid#1`, verb `frighten#1`, adverb `horryfyingly#1`
POSITIVE-FEAR	positive	noun `frisson#1`
POSITIVE-EXPECTATION	positive	noun `anticipation#1`, adjective `cliff-hanging#1`, verb `anticipate#1`

Table 1: Some of emotional categories in WORDNET-AFFECT and some corresponding word senses

`sadness#1`. Synsets representing affective states whose valence depends on semantic context (e.g. `surprise#1`) were marked with the tag AMBIGUOUS. Finally, synsets referring to mental states that are generally considered affective but are not characterized by valence, were marked with the tag NEUTRAL.

Positive	Negative	Ambiguous	Neutral	Total
97	156	20	7	*280*

Table 3: Valence distribution of emotional categories.

An other important property for affective lexicon concerning mainly adjectival interpretation is the stative/causative dimension (Goy, 2000). An emotional adjective is said *causative* if it refers to some emotion that is caused by the entity represented by the modified noun (e.g. "amusing movie"). In a similar way, an emotional adjective is said *stative* if it refers to the emotion owned or felt by the subject denoted by the modified noun (e.g. "cheerful/happy boy").

3. Affective Semantic Similarity

A crucial issue is to have a mechanism for evaluating the similarity among generic terms and affective lexical concepts. To this aim we estimated term similarity from a large scale corpus. In particular we implemented a variation of Latent Semantic Analysis (LSA) in order to obtain a vector representation for words, texts and synsets.

In LSA (Deerwester et al., 1990), term co-occurrences in the documents of the corpus are captured by means of a dimensionality reduction operated by a Singular Value Decomposition (SVD) on the term-by-document matrix. For the experiments reported in this paper, we run the SVD operation on the British National Corpus[1].

The resulting LSA vectors can be exploited to estimate both term and document similarity. Regarding document similarity, Latent Semantic Indexing (LSI) is a technique that allows us to represent a document by means of a LSA vector. In particular, we used a variation of the *pseudo-document* methodology described in (Berry, 1992). This variation takes into account also a *tf-idf* weighting schema

[1]The British National Corpus is a very large (over 100 million words) corpus of modern English, both spoken and written (BNC-Consortium, 2000).

(see (Gliozzo and Strapparava, 2005) for more details). Each document can be represented in the LSA space by summing up the normalized LSA vectors of all the terms contained in it. Also a synset in WORDNET (and then an emotional category) can be represent in the LSA space, performing the pseudo-document technique on all the words contained in the synset. Thus it is possible to have a vectorial representation of each emotional category in the LSA space (i.e. the *emotional vectors*). With an appropriate metric (e.g. cosine), we can compute a similarity measure among terms and affective categories. We defined the *affective weight* as the similarity value between an emotional vector and an input term vector.

For example, the term "sex" shows high similarity with respect to the positive emotional category AMOROUS-NESS, with the negative category MISOGYNY, and with the ambiguous valence tagged category AMBIGU-OUS_EXPECTATION. The noun "gift" is highly related to the emotional categories: LOVE (with positive valence), COMPASSION (with negative valence), SURPRISE (with ambiguous valence), and INDIFFERENCE (with neutral valence).

4. Examples of Usage

In the previous section, we have seen that the vectorial representation in the Latent Semantic Space allows us to represent in a *uniform* way emotional categories, terms, concepts and possibly full documents. The affective weight function can be used in order to select the emotional categories that can best express or evoke valenced emotional states with respect to input term. Moreover, it allows us to individuate a set of terms that are semantically similar to the input term and that share with it the same affective constraints (e.g. emotional categories with the same value of valence).

For example, given the noun *university* as input-term, it is possible to ask the system for related terms that have a positive affective valence, possibly focussing only to some specific emotional categories (e.g. SYMPATHY). On the other hand given two terms, it is possible to check whether they are semantically related, and with respect to which emotional category. Table 4 shows a portion of affective lexicon related to "university" with some emotional categories grouped by valence.

In addition we also implemented a procedure for the automated generation of evaluative expressions. These expressions are composed by a part referring to the evaluated

Related Emotional Term	Positive Emotional Category	Emotional Weight
university	ENTHUSIASM	0.36
professor	SYMPATHY	0.56
scholarship	DEVOTION	0.72
achievement	ENCOURAGEMENT	0.76
	Negative Emotional Category	
university	DOWNHEARTEDNESS	0.33
professor	ANTIPATHY	0.46
study	ISOLATION	0.49
scholarship	MELANCHOLY	0.53
	Ambiguous Emotional Category	
university	AMBIGUOUS-HOPE	0.25
career	EARNESTNESS	0.59
rector	REVERENCE	0.57
scholar	REVERENCE	0.67
	Neutral Emotional Category	
university	WITHDRAWAL	0.12
faculty	APATHY	0.13
admission	WITHDRAWAL	0.31
academic	DISTANCE	0.35

Table 4: Some terms related to "university" through some emotional categories

object (i.e. *target*) and a part expressing the affective evaluation on it. For example, the target can be represented by a noun and the evaluation by a *causative* adjective (see Section 2.), generating an expression consisting of a noun phrase.

The procedure gets in input a generic term and a fixed value of valence and creates the corresponding LSA-vector. Then, the system selects the emotional category with the input valence and the maximum value of affective weight. Finally, depending on the type of required expression, target-term and and evaluative-term are selected and the corresponding expression is composed. For example, if we give in input the verb "shoot" with negative valence, the system individuates the emotional category HORROR. Then, it extracts the noun "gun" (similar to "shoot") and the causative evaluative adjective "frightening" and finally generates noun phrase "frightening gun".

5. Possible Applications

Computer Assisted Creativity. WORDNET-AFFECT and the affective-weight function can be useful for computer assisted creativity. The automated generation of evaluative expressions with a bias on some valence orientation are at the basis of various possible applications such as automatic personalized advertisement, computational humor (Stock and Strapparava, 2003) and persuasive communication.

Verbal Expressivity of Embodied Conversational Agents. Emotions expression by synthetic characters is considered now a key element for their believability. Intelligent dynamic words selection is crucial for realizing appropriate and expressive conversations.

Sentiment Analysis. The emotional weight function can be employed in text analysis as a sentiment analysis

technique (e.g. text categorization according to affective relevance, opinion exploration for market analysis, etc.)

6. Conclusions

In this work, we continued with the activity of organization of an affective lexicon, started with the development of WORDNET-AFFECT. In particular, we considered subset of senses tagged by EMOTION label. The additional annotation allows us to distinguish emotional senses by valence and causative/stative type. Then we used Latent Semantic Analysis to represent the emotional categories in vectorial form and to realize the *affective weight* function, through which characterize the affective meaning of a generic term.

Acknowledgement

7. References

M. Berry. 1992. Large-scale sparse singular value compuations. *International Journal of Supercomputer Applications*, 6(1):13–49.

BNC-Consortium. 2000. British national corpus http://www.hcu.ox.ac.uk/bnc/.

S. Deerwester, S. T. Dumais, G. W. Furnas, T.K. Landauer, and R. Harshman. 1990. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407.

C. Fellbaum. 1998. *WordNet. An Electronic Lexical Database*. The MIT Press.

A. Gliozzo and C. Strapparava. 2005. Domains kernels for text categorization. In *Proc. of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005)*, Ann Arbor, June.

A. Goy. 2000. Lexical semantics of emotional adjectives. In S. Feist, S. Fix, J. Hay, and J. Moore, editors, *: Proceedings of Student Conference in Linguistics. MIT Working Papers in Linguistics 37*. MIT Press.

B. Magnini and G. Cavaglià. 2000. Integrating subject field codes into wordnet. In *Proceedings of LREC-2000, the Second International Conference on Language Resources and Evaluation*, pages 1413–1418, Athens, Greece.

A. Ortony, G. L. Clore, and M. A. Foss. 1987. The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology*, 53:751–766.

O. Stock and C. Strapparava. 2003. Getting serious about the development of computational humour. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI-03)*, Acapulco, Mexico, August.

C. Strapparava and A. Valitutti. 2004. WordNet-Affect: an affective extension of WordNet. In *Proc. of 4th International Conference on Language Resources and Evaluation (LREC 2004)*, pages 1083 – 1086, Lisbon, May.