Changeable Polarity of Verbs through Emotions' Attribution in Crowdsourcing Experiments

Irene Russo[1] and Tommaso Caselli[2]

[1] Istituto di Linguistica Computazionale - CNR,
Via G. Moruzzi, 1 56123 Pisa
irene.russo@ilc.cnr.it
[2] Trento RISE
Via Sommarive, 18 38123 Povo (TN)
t.caselli@trentorise.eu

Abstract. Sentiment analysis and emotion detection are tasks with common features but rarely related because they tend to categorize the objects of their studies according to different categories, i.e. positive, negative and neutral values in SA, and emotion labels such as "joy", "anger" etc. in emotion detection. In this paper we try to bridge this gap, reporting on three crowdsourcing experiments to collect speakers' intuitions on emotion(s) associated with events denoted by verbs and propose to set contextual polarity values on the basis of the selected emotions. In this way we suggest a methodology to handle connotational meanings of verbs that can help to refine automatic sentiment analysis on social media, where shared contents are often short reports on pleasant or unpleasant events and activities.

Keywords: emotion attribution, connotations of verbs, empathy.

1 Introduction

Connotations of words are important in social media communication analysis, where shared contents are often just short reports on pleasant or unpleasant activities. For instance, in [1] connotation lexicon guarantees better performance than other sentiment analysis (SA henceforth) lexicons that don't encode connotations on sentiment Twitter data.

Going towards fine grained analyses requires sentiment analysis systems able to handle different aspects of subjective language, such as i.) the fact that the polarity of words in context can be reversed or intensified by specific linguistic constructions [2]; ii.) the identification of point of views in texts [3]; and iii.) the implicit sentiment conceived as syntactic "packaging" of the sentence [4].

Sentiment analysis systems based on dedicated lexical resources such as SentiWordNet (SWN, in [5]), Subjectivity Lexicon [6] and General Inquirer [7] do not take into account how pragmatic aspects of opinion (e.g. writer's and reader's perspective) cause shifts in words polarities that can acquire a subjective nuance, as "emissions" in 1a, or display changeable polarity on the basis of reader's stance as in 1b.

1a Geothermal replaces oil-heating; it helps reducing greenhouse emissions (from [1])
1b Obama attacks Snowden.

In this paper we discuss the hypothesis that reader's stance on 1b is influenced by his/her awareness of the feelings and emotions of the agent and the patient associated with the event denoted by *to attack*. Reader's stance, and consequently the occasional subjectivity of a sentence like 1b, also depends on his/her sympathizing for that specific agent and/or patient (at the moment we do not take into account this variable). Through three crowdsourcing experiments, we test if there is agreement on the emotion attribution to the agent and to the patient in decontextualized sentences such as "x VERB y". Sentences with the target verb as the main predicate are related to 6 basic emotions (love, joy, surprise, anger, sadness and fear), following [8] framework. We also ask for the attribution of an emotion to the whole sentence with the aim to test the relevance and the direction of empathic emotion attribution. Empathy can be briefly defined as the cognitive ability – supported by shared affective neuronal networks - to intuit what another person is feeling and as a consequence to share the other person's feelings without confusing feelings experienced by the self versus feelings experienced by the other person [9]. Empathy involves inferencing about the thoughts and the feelings of the others and has, among its mechanisms, perspective taking and role taking. It is not related to automatic processes but it depends on contextual appraisal and modulation, it is influenced by saliency and intensity of the emotional state, familiarity with the involved subject, and characteristics of the empathizer [10]. Such a selective role of information explains why the same situation (or sentence, in our study) could or could not elicit empathic responses and will turn useful to explain why the polarity arising from sentential contexts has to be intended as potential, though not always instantiated, and motivating our idea of connotational polarity of verb.

2 Towards Connotations of Words

Dealing with subjectivity at word level means managing connotations of lexical items that are usually considered neutral or unspecified in SA resources because there is not a clear, homogeneous polarity attached to them, although it is widely recognized that they can display occasionally implicit polarity for speakers that include them in their discourse - even in fact-reporting discourse. Sentiment analysis based on word occurrences in texts have focused at the beginning on adjectives and adverbs that, since first experiments in opinion mining [11], proved to be the most useful indicators of subjectivity in texts because they are used to synthetitically express judgments on entities. For other words, like nouns as *party* and *incident*, the subjective meaning is not the constitutive part. Nonetheless they can display in context polarized usages and as a consequence the can acquire a polarity as effect of semantic prosody [12].

Verbs are the neglected part of speeches when connotations are investigated. From a semantic point of view verbs play a key role in the organization of the information, usually help in the description of an action/situation/state of being and, by denoting events, processes or states that happen or are valid in the world [13], are not be included in SA lexicons with the same modality of purely evaluative words, such as adjectives and adverbs, that are used to convey speakers' stances in texts and discourses.

However, several verbs denoting events have positive or negative polarity values in lexical resources such as SWN and the OpinionFinder lexicon. A quantitative analysis on existing lexica for SA provided the following results: the OpinionFinder Lexicon has 5.2% of neutral values for verb lemmas; in SentiWordNet, it reaches 76,5% and, finally, the CoNLL 2011 Subjectivity Sense Annotation [14] 59.81% of verb senses are labelled as objective. For instance, the verb "to attack" with sense key attack%2:33:01:: in WordNet 3.1 (WN) has been labelled as objective in the CoNLL 2011 Subjectivity Sense Annotation. However, considering one of the examples reported in 2a which accompanies the gloss and how it would be perceived by a reader/speaker, it's clear that this sense of "to attack" is not always objective but could trigger judgments on the event described depending on the feelings and the attitudes of the reader toward the agent of the sentence. In a similar vein, 2b can be perceived as reporting a positive event, if, for instance, the reader is a social media user sympathizing with a close friend.

2a. The Serbs attacked the village at night.
2b. I attacked the burglar last night and saved my new laptop!

Moreover, WordNet senses for the verb "to attack" in two different SA resources display different polarities (see table 1):

Resource	to_attack#1	to_attack#2	to_attack#3	to_attack#4	to_attack#5	to_attack#6
ConLL2011 SSA	obj	subj	obj	both	obj	obj
SWN 3.0	P: 0 O: 1 N: 0	P: 0 O: 1 N: 0	P: 0 O: 0.5 N: 0.5	P: 0 O: 0.625 N: 0.375	P: 0 O: 1 N: 0	P: 0 O: 1 N: 0

Table 1. Comparison between two SA resources for the verb *to attack*.

In SWN the synset values (based on the quantitative analysis of the glosses associated to synsets and on vectorial term representations for semi-supervised synset classification) are different with respect to [14], which is a manually annotated gold standard. According to this evidence assigning polarity out of context don't provide homogeneous results.

In this paper we focus on 51 verb lemmas (such as *to hug, to abort, to wait, to hide* etc.) as a case study and we propose to list them as potentially polarized items on the basis of the emotions attributed to their participants. In particular, the first two polarity values of this new structure correspond to the polarity associated to the emotion attributed to the thematic roles of agent/experiencer and that of patient, while the third value is derived by the emotion(s) attributed by the hearer/reader to the whole sentence. Though similar in concept to polarity values of verbs in existing lexica, our encoding is different since it is grounded on and derived from the emotions attributed to event participants. We want to propose multiple values which could be activated in the reader/hearer mind. The main reason for this choice is linked to the working hypothesis that the participants of events can trigger different, even opposed, connotational polarity values and that the polarity value of the whole sentence is dependent on the empathic involvement of the reader/hearer.

3 Crowdsourcing Emotions Associated to Verbs

In order to investigate how verbal polarities can depend on emotions' attributions, we identify a set of Italian verbs on the basis of the following criteria: a.) frequency in the corpus La Repubblica [15]; b.) polarity values in SWN (neutral items vs. polarized items); and c) context of occurrence based on the verb syntactic and semantic frame (for transitive verbs – Subject[Human] Verb Direct_Object:[Animate|Object] - vs. for intransitive verbs – Subject:[Human] Verb; or Subject:[Human] Verb Preposition_NP:[Animate|Object]). In this way, we collected 51 different verb lemmas and a total of 60 verb frames. The data have been uploaded as three different crowdsourcing tasks on the CrowdFlower platform.

The first task aims at collecting judgments on the emotion(s) of the grammatical subject (Agent or Experiencer) involved in a certain situation. The second tasks aims at collecting the emotion(s) of the direct object when realized by an animate filler (Patient). Finally, the third task tries to identify the emotion(s) of an external observer, i.e. the reader/hearer of the reported situation. To clarify, consider the following example:

X [Human] *hugs* Y [Human]
Emotion of X: love
Emotion of Y: pleasure
Emotion of EO: joy

where X stands for the subject, Y for the patient and EO for the reader of the sentence.

One of the main issues in using crowd-sourcing techniques is related to quality control. In order to assure the goodness of the data collected we have adopted the following strategies, namely i.) we have created a Gold Standard, composed by 10% of the verb frames, by manually selecting among our data highly polarized items (e.g. the verbs *amare* [to love] and *odiare* [to hate]) for a total of ; ii.) we did not offer any compensations and recruited our workers by means of a campaign on social networks such as Facebook and Twitter. The first strategy will help us in assuring that the workers' answers are correct with respect to the instructions. On the basis of CrowdFlower settings, the trust thresholds was set to 75% of the Gold Standard, i.e. if a worker provides less that 75% of the correct answers in the Gold is considered as untrusted and its answers are not taken into account. On the other hand, the second strategy facilitates the recruitment of interested workers, thus avoiding the presence of spammers. The three tasks have a similar structure, based on three blocks of questions:

- the first question asks the workers if the subject, the direct object or an external observer, respectively, experience an emotion on the basis of the verb context. This question has been selected in order to develop the different Gold Standards. However, the Gold Standards apply only to the first and second tasks (subject and direct object emotion). As for the exploratory nature of the third task (external observer emotion) we did not provide any Gold Standard to avoid influencing the workers' judgments;

- the second question requires the workers to select one or more emotion(s). The workers were presented with the list of Parrot's basic emotions [8] (i.e. love, joy, surprise, anger, sadness and fear) plus an additional value "other". This underspecified value has been selected in order to elicit from the workers other emotions. Notice that only one value can be assigned to "other";
- the third question requires the workers to grade the magnitude/intensity of the selected emotion(s) on a scale ranging from 1 (lowest intensity) to 5 (highest intensity).

Following [16], a maximum of 5 judgments is required in order to finalize the analysis of each verb context.

4 Data Analysis

The analysis will be in two parts: first we will report on the data of the three tasks separately, and then we will provide a global analysis which comprise a method to identify and assign the connotational polarity of verbs. All three tasks were completed in a week. The judged contexts have been analysed on the basis of the agreements on: a.) the existence of an emotional reaction; and b.) the emotion value(s). We have identified 3 clusters of agreement: 1) below 0.5 (no agreement); 2) from 0.5 up to 0.6 (low agreement), and 3) from 0.7 to 1.0 (high or perfect agreement).

4.1 Emotions and Wisdom of the Crowd

The first task aimed at collecting judgments on the emotions of the subject/agent-experiencer of a set of specific actions. For 60 verb contexts we collected a total of 468 judgments. Only 396 judgments were retained. According to the Gold Standard, 291 judgments (73.48%) were provided by trusted workers and 105 (25.52%) by untrusted workers. Overall accuracy (i.e. the percentages of the agreed and non-agreed judgments on the existence of an emotion for the subject/agent-experiencer) of the trusted judgments is 94%. These figures suggest that the task is not trivial and people easily agree on the presence of an emotion when the subject performs certain actions. Most of the contexts were considered as emotional for the subject (52/60), while only 8 cases were considered as not emotional.

The second task aimed at collecting judgments on the emotions of the direct object/patient. In this case, the set of contexts was reduced to 42 (only transitive contexts with an [Animate] direct object). We collected a total of 456 judgments. As in the first task, only 396 judgments were retained as valid. In particular, 261 (65.9%) were from trusted workers and 135 (34.1%) were from untrusted ones. In addition to this, the overall accuracy is 88%. In this case the task, though easy, is more difficult with respect to the first one. Similarly, most of the contexts were considered as emotional contexts for the direct object (38/42), while only 4 cases were classified as not emotional.

The third task is the most complex. The workers were required to assign an emotional value to the event contexts as if they were an external participant, i.e. as being someone which assists to or reads about the action denoted by the verbs. Due to the nature of the task, and the fact that an emotional reaction to an event is extremely grounded on each person's experience, no Gold Standard for assessing trusted and untrusted workers was developed. We collected judgments on all 60 contexts, for a total of 365 judgments. All judgments were retained as good. The presence of spammers is excluded on the basis of the recruitment procedures of the workers (see Section 3). 50 contexts were considered as eliciting an emotion from an external observer, while only 10 of them are considered as non-emotional ones, i.e. neutral.

In Table 2 we report the frequency of the contexts with respect to their distribution in the three clusters of agreement on the emotional contexts and on the specific emotion for the three tasks. As for Task 1, 37 emotional contexts belong to the transitive pattern Subject[Human] Verb Direct_Object:[Animate], 12 belong to the transitive pattern Subject[Human] Verb Direct_Object:[Object] and 3 to the intransitive pattern. Concerning the non- emotional contexts, the distribution in the three cluster is quite similar for all the tasks, namely in Task 1 we have 5 items in the low agreement cluster and only 3 in the high agreement cluster. In Task 2 all items are in the low agreement cluster. In Task 3 we observe 4 items in the low agreement cluster and 6 in high agreement cluster.

Tasks	no agreement	low agreement	high agreement
Task 1: Subject emotion	0	12	40
Task 1: Emotion value	2	18	32
Task 2: D.O. emotion	0	6	32
Task 2: Emotion value	2	15	21
Task 3: Observer emotion	0	9	41
Task 3: Emotion value	9	17	24

Table 2. Distribution of the emotional contexts and the emotion value among the three clusters of agreement.

Table 3 reports the figures on the selection of a specific emotion for the three tasks. The computation of the preferred emotions based both on majority voting and on the magnitude/intensity. As for the value "other", we obtained different sets of elicited emotion nouns, which in large part can be mapped to Parrot's lists of secondary and tertiary emotions. In particular, in Task 1 we collected 56 unique emotion nouns (37 *hapax*, and the remaining with a frequency ranging from 2 to 9); in Task 2, 35 unique emotion nouns (21 *hapax*, and the remaining with a frequency ranging from 2 to 12); and in Task 3, 43 unique emotion nouns (31 *hapax*, and the remaining with a frequency ranging from 2 to 4).

Emotion Values	Preferred Emotion		
	Task 1	Task 2	Task 3
Love	19.2% (10 contexts)	5.26% (2 contexts)	21.67% (13 contexts)
Joy	15.38 (8 contexts)	23.68% (9 contexts)	6.67% (4 contexts)
Surprise	7.69% (4 contexts)	2.63% (1 contexts)	8.33% (5 contexts)
Anger	13.46% (7 contexts)	21.05% (8 contexts)	18.33% (11 contexts)
Sadness	5.76% (3 contexts)	5.26% (2 contexts)	3.34% (2 contexts)
Fear	19.2% (10 contexts)	15.78% (6 contexts)	10% (6 contexts)
Other	19.2% (10 contexts)	26.31% (10 contexts)	15% (9 contexts)

Table 3. Percentages of selection of the preferred emotions on the three tasks.

By observing the data, we checked if the emotion associated with the sentences depends on empathic involvement, without focusing on the agent or the patient. In 49% of the cases there is a kind of empathic involvement that cause a coincidence between emotions associated to the whole sentence and those attributed tone of the participant to the event. When this does not occur, the agreement on the presence of an emotion is low (i.e. the sentence is located in cluster 2) or the kind of event involves ambiguous emotions (e.g. *X cade*, "X falls down" is associated with fear and surprise).

As a preliminary method for dealing with connotational polarity of verbs, we propose to set numerical values for positive/negative emotions on the basis of the crowdsourced data. Among Parrot's (2001) six basic emotions two of them are positive ("love" and "joy"), three are negative ("anger", "sadness" and "fear") and one is ambiguous ("surprise"). Taking into account the average value of the emotion more often associated with the verb, we multiply it by the agreement value both on the emotion and on the fact that the sentence elicit an emotion in one of the event participants. A global polarity value for verbs can be obtained as the mean value for the same sentence evaluated in the three tasks (i.e. from the point of view of the agent, of the patient, and from a general external point of view); we scale this value between 0 and 1, as reported in Table 4. X stands for a human subject; Y stands for an animate direct object and Z for an inanimate one.

Sentence	Polarity value
X cura Y *[X heals Y]*	0.3787
X applaude Y *[X claps Y]*	0.3721
X scrive a Y/uno Z *[X writes to a Y/ writes a Z]*	0,1194
X abbraccia Y *[X hugs Y]*	0.6322
X difende Y *[X defends Y]*	0.1422
X ricorda Y *[X remembers Y]*	-0.0639
X nasconde Y/uno z *[X hides Y/hides a Z]*	-0,4996
X ammazza Y *[X kills Y]*	-0.5445
X discute con Y *[X argues with Y]*	-0.2360
X ferisce Y *[X wounds Y]*	-0.3722

Table 4. Global polarity values for some verbs in the data set.

The final result of our polarity analysis will have multiple values, ranging from -1 (negative polarity) to 1 (positive polarity). For instance, a transitive pattern of such as "X[Human] kills Y[Animate]" will have a tripartite valued structure, with a specific polarity value for X, one for Y and a proposed global value associated with the verb pattern (as in Table 4).

5 Conclusions and Future Perspectives

Sentiment analysis and emotion detection are tasks with common features but rarely related because they tend to categorize the objects of their studies according to different categories, i.e. positive, negative and neutral values in SA, and emotion labels such as "joy", "anger" etc. in emotion detection.

In this paper we try to bridge this gap, reporting on three crowdsourcing experiments to collect speakers' intuitions on emotion(s) associated with events denoted by verbs and propose to set contextual polarity values on the basis of the selected emotions. This approach needs testing to identify in contexts the polarity values of verbs. In particular, future work will concentrate on the elaboration of specific rules to map a set of optional polarized values that can be accepted or refused also depending on the textual genre considered (i.e. social media vs. newspapers).

We believe that taking into account the different perspectives involved in the emotional evaluation of an event described with a verb can help sentiment analysis systems to deal with the complexity of the role of verbs in expressing judgments and opinions, even starting with the analysis at the lexical level.

Better understanding of how subjective language works can improve artificial natural language intelligence, making language-based human-computer interaction more comfortable [17] and improving the modeling of emotional states in intelligent social agents that need to communicate with users in natural language [18].

References

1. Feng, S., Kang, J.S., Kuznetsova, P. and Choi, Y.: Connotation Lexicon: A Dash of Sentiment Beneath the Surface Meaning. In: Association for Computational Linguistics Proceedings (ACL), (2013)
2. Pang, B. and Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval. 2(1-2), (2008)
3. Sayeed, A.: An opinion about opinions about opinions: subjectivity and the aggregate reader. In: North American Association for Computational Linguistics (NAACL 2013) Proceedings, Atlanta, USA (2013)
4. Greene, S. and Resnik, P.: More Than Words: Syntactic Packaging and Implicit Sentiment, In: NAACL 2009 Proceedings, Boulder, CO, (2009)
5. Baccianella, S., Esuli, A. and Sebastiani, F.: SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), European Language Resources Association (ELRA), pp. 2200-2204. (2010)
6. Wilson, T., Wiebe, J. and Hoffmann, P.: Recognizing Contextual Polarity in Phrase- Level Sentiment Analysis. In: Proceedings of HLT/EMNLP 2005, Vancouver, Canada (2005)
7. Stone, P. J., Dunphy, D. C., Smith, M. S., and Ogilvie, D. M.: The General Inquirer: A Computer Approach to Content Analysis. MIT Press (1966)
8. Parrott, W. G.: The nature of emotion. In: A. Tesser & N. Schwarz (eds.), Blackwell handbook of social psychology: Vol. 1. Intraindividual processes. pp. 375-390. Basil Blackwell, Oxford (2001)
9. Decety, J. and Ickes, W. (eds.): The Social Neuroscience of Empathy. MIT Press, Cambridge (2009)
10. De Vignemont, F., & Singer, T.: The empathic brain: How, when and why? Trends in Cognitive Sciences. 10(10), 435-441 (2006)
11. Hatzivassiloglou, V. and McKeown, K.: Predicting the semantic orientation of adjectives. In: Proceedings of 35th Meeting of the Association for Computational Linguistics. Madrid, Spain, pp. 174-181 (1997)
12. Louw, B.:Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies. In: Baker, M., Francis, G. & Tognini-Bonelli, E. (eds.) Text and Technology. John Benjamins, Philadelphia/Amsterdam (1993)
13. Vendler, Z.: Linguistics in Philosophy. Cornell University Press, Itchaca, NY (1967)
14. Akkaya, C. Wiebe, J., Conrad, A. and Mihalcea, R.: Improving the Impact of Subjectivity Word Sense Disambiguation on Contextual Opinion Analysis. In: Proceedings of Conference on Computational Natural Language Learning (2011).
15. Baroni, M. S. Bernardini, F. Comastri, L. Piccioni, A. Volpi, G. Aston, M. Mazzoleni: Introducing the "la Repubblica" corpus: A large, annotated, TEI(XML)-compliant corpus of newspaper Italian. In: Proceedings of LREC 2004 (2004)
16. Mohammad, S. M. and Turney P. D.: Crowdsourcing a Word–Emotion Association Lexicon. Computational Intelligence. 59, 1-24 (2011)
17. Alm, C. O.: Subjective natural language problems: Motivations, applications, characterizations, and implications. In: Proceedings of ACL (2011)
18. Paiva, A.: Empathy in Social Agents. International Journal of Virtual Reality. 10-1, 65-68 (2011)