

A Lexical Grammatical Implementation of Affect

Matthijs Mulder^{1,2}, Anton Nijholt¹, Marten den Uyl², and Peter Terpstra²

¹ University of Twente, PO Box 217, 7500 AE, Enschede, The Netherlands
Email: a.nijholt@ewi.utwente.nl

² Parabots Services, Singel 160, 1015 AH, Amsterdam, The Netherlands
Email: {m.mulder, m.denuyl, p.terpstra}@parabots.nl

Abstract. In this paper we report about our research towards the use of affect in language wherein we have attempted to formalise the affective functionality at word and grammatical level for a fraction of Dutch and English. These formalisations have been demonstrated in a pilot experiment. The empirical background of the formalisation, and the results of the experiment constitute the basis for further research on a lexical, grammatical implementation of affect.

1 Introduction and Background

Affect is part of the non-literal, emotional meaning of language. Although the literal meaning of text is a well known topic within the field of Natural Language Processing (NLP), few research has been done on the non-literal meaning of language. In this paper we report about our research towards *affect* in language.

We have chosen to define affect in language as text having a *positive or negative orientation*, an *intensity*, and a *direction* towards an object. This definition of affect is based on the notion of *subjective sentences* taking a *psychological point of view*. These sentences present characters' thoughts and perceptions and express characters' beliefs, emotions, and other attitudes of affects [1,2]. They can be described using three attributes: the *experiencer*, the *attitude* and the *object*. The experiencer is the person being in a private state of the kind *attitude* towards the *object*.

Osgood's *Theory of Semantic Differentiation* suggests that meaning can be defined within a multidimensional semantic space, set out by descriptive scales, defined by pairs of antonymous adjectives being opposite in meaning [3]. A set of independent descriptive scales represent the meaning of a word. The *direction* and *distance* of a vector represent the *quality* and *intensity* of the property defined by the antonymous adjectives. The three most dominant factors of emotional meaning are *evaluation* (the attitudinal variable in human thinking, judging if the concept is either good or bad), *activity* (concerned with power and things associated with it: height, size, toughness, etc.), and *potency* (concerned with quickness, excitement, warmth, agitation, etc.).

Kamps et al. [4] propose the use the *evaluation* scale in combination with *thesauri* like WordNet [5] for the detection of the *semantic orientation* (the positive, negative, or neutral orientation) of adjectives (words having an "attitude"). Although it is suggested to use a more-dimensional definition of emotion in the field of emotion-research [6,7], the positive-negative scale suffices when we want to express the semantic orientation of language. This positive-negative orientation of phrases is a common approach in the field of affective language research [4,8,9,10].

The use of simple *grammatical models*, or templates, in the research towards affect have proven to be useful for the detection of subjective sentences, and of words having a semantic orientation [11,10]. These simplistic models describe how words with an affective meaning are being used within a sentence, but fail to offer a more general approach.

What seems to be lacking in the research towards affect in language from the NLP point of view is the relation between *attitude* and *object*. How does the attitude in a subjective sentence influence the way we think about the object? This question is relevant, because of the fact that opinions about persons or issues are often based upon the things we read or hear.

We have therefore studied how the relation between *attitude* and *object* may be formalised. We acknowledge the importance of the source and target when it concerns affect in language, but have restricted our research to plain text, without knowing the source, thus creating a more objective notion of the effect of affect. We have also chosen to avoid issues as humour, irony, and sarcasm in language. Affect is therefore defined as the *semantic orientation* of the subjective sentence, directed at an *object*, having an *intensity*.

Affect directed at an object may also influence other objects within the sentence when a semantic relation is established between the two. The sentence ‘The dog of the neighbours stinks’ does not only affect the ‘dog’, but also the ‘neighbours’. The preposition ‘of’ has created a relation between the two objects.

We will therefore use the word “*target*” to refer to the entity for which we want to investigate the affective effect of a sentence. A (possibly anaphoric) reference towards this “target” may be contained in the sentence, but does not necessarily have to. When we exclude pragmatic relations between entities we assume that targets not being referred to within a sentence receive the neutral affective meaning from this sentence; the opinion about the target is not influenced by the content of the sentence. This is illustrated by the sentences in 1. Suppose that the target for all three sentences is ‘dog’, then sentences 1a and 1b (through the pronominal anaphoric reference) are negative, while 1c is neutral. When the target would be defined as being the ‘cat’, however, sentence 1c would be positive, while the outcome for the other sentences would be neutral.

1. (a) The neighbours’ dog stinks. (dog: - | cat: 0)
- (b) The neighbours have a dog. It stinks. (dog: - | cat: 0)
- (c) The neighbours’ cat is very beautiful. (dog: 0 | cat: +)

A lexical, grammatical approach has been chosen in order to formalise the propagation of affect towards a target, thus dividing between the *lexical* and *grammatical* aspects of affective language. This solution is based on the assumption that affective meaning is contained in content-words, but intensified and propagated towards the target through function-words and grammatical constructions. This assumption is an attempt to combine previous research in one formal system. All research assumed the existence words having an affective meaning as the bearers of affect, and the results of experiments conducted with human judges have fortified this assumption [10]. Grammatical constructions in order to detect affective sentences have been used by several researchers [11,10]. An affective grammar would be an attempt to generalise these models.

2 The Affective Mechanisms

In this section we report about the formalisation of some of the mechanisms that seem to play a part in the propagation of affect within a sentence at grammatical and word level. This paper does not allow for a detailed description of the implementation and use of these mechanisms within a grammatical definition of a subset of English or Dutch. We refer to [12] for a detailed discussion of these mechanisms, and will report about their formalisation and implementation in a future paper.

The expression “*affective value*” is used as the means of expressing the affective meaning of a word. The affective value is a position on the positive-negative scale, represented by a rational number ranging from -1 (for negative evaluations) to 1 (for positive evaluations).

2.1 Intensification

The affective value of some adjectives, substantives, and verbs can be *intensified* by other words, or by pre- or postfixes. Words belonging to the set of *scalable adjectives*, and having a non-zero affective value (e.g. ‘good’, ‘bad’, ‘ugly’) are examples of such [13]. Adverbial modification intensifies or diminishes the affective value of these adjectives. The phrase ‘very good’ demonstrates how the adverbially used adjective ‘very’ intensifies the affective value of the adjective ‘good’. The word ‘reasonably’ in the phrase ‘*reasonably* good’ is an example of a word *diminishing* the affective value of the adjective ‘good’. Comparative and superlative formations like ‘nice-nicer-nicest’ are examples of intensification through postfixation. Adjectives like ‘super’ may be used as intensifying prefixes in words like ‘*supercool*’.

We have chosen to implement intensification as the multiplication of the affective value with the *potency value* of the intensifier. A potency value smaller than one diminishes the affective value, while a value greater than one intensifies this value.

The intensifier function can also be used as an implementation of the notion of *distance* between the *target* and the *phrase containing the actual affect*. Prepositional phrases, clause phrases and verb phrases containing an adverbial phrase or noun phrase create such a distance between affective language and target. Example 2 illustrates how the preposition ‘of’ creates a distance between the substantive ‘neighbours’ and the negative verb ‘stinks’. An implementation of the intensifier mechanism *diminishes* the affective value of verb if the preposition, pronoun, or verb has got a potency value smaller than one. The notion of distance is bidirectional, as is illustrated by example 3.

2. [_{NP} The dog [_{PP} of the neighbours]] [_{VP} stinks]
3. [_{NP} The dog [_{PP} of the friendly neighbours]] [_{VP} sits in front of the door]

2.2 Negation

Negation influences the *semantic orientation* of an expression: positive becomes negative, and negative becomes positive. Negation has already been used in his research towards affect in language by [11]. It may be realised by the word ‘not’, but also by the prefix ‘un-’, as is illustrated by the sentences in 4. Adjectives, but also verbs, may be negated. The negation of a verb does not only change its semantic orientation, but also changes the affective meaning of its direct object, as is illustrated by sentence 5.

4. (a) The friendly man
(b) The unfriendly man
5. (a) I think [VP artists are lazy]
(b) I don't think [VP artists are lazy]

Negation using the prefix 'un-' is different than negation using the word 'not'. The word 'unfriendly', for example, has got another meaning than the expression 'not friendly' [13]. In spite of this difference these two forms of negation have been implemented as the negation of the affective value of the lemma; the sentences in 4 suggest that the *absolute* affective value of the words 'friendly' and 'unfriendly' is more or less the same.

The effect of the negation on the verb with the clause phrase as its direct object has been formalised as not only the negation, but also a diminishing of the affective value of the clause phrase towards the target.

2.3 Perception

Is there a maximum and minimum affective evaluation of an affective expression? The examples in 6 illustrate how there is no end to the length of the expression *containing* affect. The *perception*, or *sensation* of such an expression, however, will not increase linearly and infinitely, and the reader will at some point experience a maximum to the sensation of affect.

6. (a) A *very* good idea.
(b) A $very_0, very_1, \dots, very_\infty$ good idea.

The relation between stimulus (the actual text) and sensation (the resulting affective value of the expression) may be formalised using a logarithmic function, analogous to for example the relation between the actual and perceived intensity of sound (decibel). The *sigmoid function* has the nice property of being point-symmetrical in the origin, and of having two horizontal asymptotes, defining the maximum and minimum affective value. This function may also be used to *scale* the affective value when changing from word to expression, or from expression to sentence level, since the effect of one single word may be greater than that of thousand others.

2.4 Quantification

The affective mechanism of *quantification* is different from the others, since it does not modify the affective value attributed to the target, but the target itself. A target may represent a *set* of entities, and a quantifier selects a subset of the complete set. The attribution of affect towards a *subset* of the set denoted by the target will not have the same effect as the attribution towards the *complete* set.

Quantification is difficult to formalise when it concerns set-denoting adjectives that do not only denote a subset, but also attribute an affective value to this subset. Sentence 7a, shows how the adjective 'poor' may (but does not have to) denote a subset of the 'artists', but also attributes its affective value to this set. Sentence 7b shows how the positive affective value of the adjective 'skilled' is used negatively because of the very small subset of the artists (denoted by the adjective 'scarce') having this quality.

7. (a) These poor artists should receive subsidy
- (b) Skilled artists are scarce
8. *Some* artists make nice things.

We did not formalise these difficult cases of quantification, but did try to formalise a special subset of *quantifying adjectives* like ‘some’, ‘many’, ‘none’, and ‘all’. These words are specifically used to denote a subset, and their meaning is more or less the same for all participants in communication. Sentence 8 is an example of a sentence containing such a quantifier. *Quantifiers* are defined as words having a value between zero and one representing the fraction denoted by these words. The affective value of expressions attributed to quantified targets are multiplied with this fraction value.

2.5 Verbs Implementing a Relation Between Two Objects

In subsection 2.1 we have showed how verbs, amongst others, implement the notion of *distance* between two objects. Verbs may also implement a *relation* property. With relation we mean the positive or negative propagation of attributed affect from one entity to the other. This is illustrated by the sentences in 9.

9. (a) The prince [*REL*+supports] the WWF.
- (b) The prince [*REL*+supported] the Nazi regime.
- (c) *Artists* [*REL*-detest] criminality.
- (d) *Artists* [*REL*-detest] morality.

The relation defined by the verb ‘to support’ in sentences 9a and 9b defines a *positive* relation between the subject and the direct object. Affective values keep their orientation when they are being propagated from one entity to the other. The verb ‘to detest’ in the sentences 9c and 9d, however, defines a *negative* relation. Positive becomes negative (9d), and negative becomes positive (9c).

When the relation property is combined with the notion of distance, the affective contribution of verbs to the propagation of affect within a sentence can be formalised.

3 The Experiment

The mechanisms discussed in the previous section were implemented simple grammar defining a subset of Dutch. In order to test out these mechanisms a pilot experiment was conducted. We wanted to know if language users would agree upon the affective meaning of words and sentences, and in the same time use their evaluations in order to test the performance of the “affective grammar” that was created from the implementations of the mechanisms. In this experiment we asked a group of human judges to evaluate words and sentences according to their affective meaning.

3.1 The Setup

A corpus containing 42 sentences was created in order to test the performance of the grammar. These sentences were taken from a discussion about subsidised art between readers of the *NRC Handelsblad*, a Dutch newspaper. The criteria for the selection of the sentences was

that they contained the word *kunstenaars* ‘artists’, leaving out the sentences that could not be parsed by affective grammar. Some sentences were simplified removing their auxiliary verbs.

Questionnaires were created containing the content words appearing in the 42 sentences of the corpus. We chose to organise these words according to their functionality, resulting in a list of words having an affective value, a list of intensifiers, and a list of quantifiers. These words had to be evaluated using a 7-point scale representing respectively the affective value (ranging from ‘very negative’ to ‘very positive’), the potency of the intensification (ranging from ‘very weak’ to ‘very strong’), and the fraction (ranging from ‘nothing’ to ‘all’). Another questionnaire consisted of the sentences that had to be evaluated according to the affective value they propagated towards the target ‘artists’, ranging from ‘very negative’ to ‘very positive’.

The questionnaires were distributed on the Internet. We also asked the participants for their age and sex, in order to possibly find any correlations between age, sex, and their evaluations.

3.2 The Results

Thirty-eight persons filled out the questionnaires. Forty-five percent of the the participants was male, 55% of them female, and their average age was 25. A multiple raters variation of Cohen’s kappa statistic was used in order to calculate the *interrater agreement* [14]. A value of $\kappa = 0.395$ was calculated for the raters’ agreement of the words being positive or negative, and this value is too low in order to say that the raters agreed upon the affective value of words. A value higher than 0.4 would mean a fair to good level of agreement beyond chance alone. This score was to be expected because of the fact that the existence of a set of words having an affective meaning that is more or less equal for all language users does *not* mean that there does not exist a set of words having a meaning that is context-dependent and different for all participants. We even expect that the lexicon contained words having an affective meaning that is not the same for all participants (words like ‘world’ and ‘fancy-fair’), but did not have the criteria to distinguish between these two sets of words. Based upon the assumption that disagreement between users would lead to a neutral mean affective value, the mean value of the evaluations of the words was used as the input to the affective grammar, and the mean value of the evaluations of the sentences as the expected output. Looking at the data, and drawing some histograms of their evaluations hinted a possibly normal distribution of the evaluations, and this would justify the choice to use the mean value of the words and sentences.

Different tests were conducted in order to test out the performance of the grammar. As a baseline we have chosen for the normalised bag-of-words method, stating that the affective value of a sentence is equal to the sum of the affective value of its words, and having a maximum value of 1, and a minimum value of -1 . Other methods testing out the contribution of the different mechanisms were created, leaving them out of the grammar.

The baseline method correctly predicted the semantic orientation of 71.4% percent of the sentences, while the complete grammar yielded a score of 76.4%. The other tests showed out that this increased performance can be attributed to the mechanism of negation. The affective grammar scored a mean standard deviation over all sentences of 0.149, while the baseline method scored a mean standard deviation of 0.307. The other tests suggest a positive contribution of all defined mechanisms to the overall performance.

4 Conclusions

Our work has shown how the propagation of affect towards an object may be formalised at lexical and grammatical level. This approach formalises affect in a more general way than then NLP approaches we reported about, abstracting general forms from the specific cases of affective constructions in sentences. We have not only given an empirical discussion of these mechanisms, but also tried to formalise and implement them in an affective grammar. The performance of this grammar was tested in a pilot-experiment, using the mean value of the evaluations of human judges of words and sentences. The grammar outperformed the normalised bag-of-words baseline method, and several other tests hinted a meaningful positive contribution of all defined mechanisms. We conclude that our lexical, grammatical approach is a step forward in the NLP research towards the analysis of affect in language. Its strongest feature is its formalisation of the *propagation* of the attitude (affective content) towards an object. In a future paper we will report about the exact formalisation and implementation of the several mechanisms.

References

1. Wiebe, J.: Tracking the point of view in narrative. *Computational Linguistics* 20:2, 233–287 (1991).
2. Wiebe, J., Wilson, T., Bruce, R., Bell, M., Martin, M.: Learning subjective language. Technical Report TR-02-100, Department of Computer Science, University of Pittsburgh (2002).
3. Osgood, C.: *The Measurement of Meaning*. University of Illinois Press, Urbana IL. (1957).
4. Kamps, J., Marx, M.: Words with attitude. CCSOM Working Paper 01-194 (2001).
5. Beckwith, Fellbaum, Gross, Miller: Introduction to wordnet: An online lexical database (1993).
6. Frijda, N., Swagerman, J.: Can computers feel? theory and design of an emotional system. *Cognition and Emotion* 1 (1987) 235–257.
7. Ortony, A., Clore, G., Collins, A.: *The cognitive structure of emotions*. New York: Cambridge University Press (1988).
8. Hovy, E.H.: *Generating Natural Language Under Pragmatic Constraints*. Lawrence Erlbaum, Hillsdale, New Jersey. (1988).
9. Turney, P., Littman, M.: Unsupervised learning of semantic orientation from a hundred-billion-word corpus. Technical report, National Research Council Canada (2002).
10. Hatzivassiloglou, V., McKeown, K.R.: Predicting the semantic orientation of adjectives. In: Proc. of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL, Madrid (1997) 174–181.
11. Liu, H., Lieberman, H., Selker, T.: A model of textual affect sensing using real-world knowledge. Technical report, MIT Media Laboratory, Cambridge, USA. (2003).
12. Mulder, M.: A lexical grammatical implementation of affect. Master's thesis, University of Twente (2003).
13. Broekhuis, H.: Adjectives and adjective phrases. Technical report, University of Tilburg (1999) *Modern Grammar of Dutch Working Paper* 2.
14. Green, A.: Kappa statistics for multiple raters using categorical classifications. In: Proceedings of the 22nd Annual SAS User Group International Conference. (1997).