The Detection and Analysis of Bi-polar Phrases and Polarity Conflicts

Klenner, Manfred ; Tron, Susanna ; Amsler, Michael ; Hollenstein, Nora

The Detection and Analysis of Bi-polar Phrases and Polarity Conflicts

Manfred Klenner, Susanna Tron, Michael Amsler, and Nora Hollenstein

Institute of Computational Linguistics
University of Zurich, Switzerland
{klenner|tron|mamsler|hollenstein}@cl.uzh.ch

Abstract. In fine-grained sentiment analysis one has to deal with the composition of bi-polar phrases such as e.g. *just punishment*. Moreover, the top down prediction of phrase polarity as imposed by certain verbs on their direct objects sometimes is violated by the bottom up composed phrase polarity (e.g. 'to approve war'). We introduce a fine-grained polarity lexicon built along the lines of the Appraisal Theory and we investigate the composition of bi-polar phrases - both, from a phrase internal point of view and from a verb-centered perspective. We have specified a multi-lingual polarity resource (French, English, German) and a system pipeline that carries out sentiment composition for these languages. We discuss examples with reference to each of these languages.

1 Introduction

Advanced sentiment analysis builds on sentiment composition where the polarity of words from a prior polarity lexicon are combined to give the polarity of phrases in order to incrementally approach the polarity of the whole clause. Simple approaches, where the polarity of a phrase or sentence is calculated from the ratio of positive and negatvie word level polarities and the number of negators are bound to fail in the presence of sophisiticated language, that is, language beyond product reviews - a text genre most desparately analysed in the field of sentiment analysis so far. We argue that more fine-grained distinctions are needed to analyse the polar load of texts than reference to the basic distinctions 'positive' and 'negative'. The Appraisal Theory (Martin and White (2005)) has introduced such a finer distinction: the one between judgement, emotion and attitude. A word or phrase is no longer just positive, but might be positive on a moral (humanity), emotional (love) or factual (victory) basis.

A particular problem arises if two or more words with different polarities are combined to form a phrase. What is the phrase-level polarity, then? Take *(this is) just punishment* - is this positive, negative or should we rather just stick with the term 'bi-polar'? If we follow the distinctions made by the Appraisal Theory, another option would be to just refer to different kinds of polarity. We could say that *just punishment* is positive at the moral level (since the punishment is described as being just) but from a factual perspective it is negative (since it mentally or physically injures a person).

Another source of polarity conflicts are verbs such as 'to admire', 'to prefer' etc. They seem to impose a polarity expectation on their arguments (e.g. the direct object). So 'admire' requires a positive or a neutral noun phrase as its direct object. But what if the noun phrase was negative? Does one admire negative objects (in the broadest sense). It depends on the kind of negativity, we argue. How can we explain why *He admires his sick friend* does not produce a conflict while *He admires his deceitful friend* does, although both noun phrases are negative (*sick friend* and *deceitful friend*). One hypothesis is that the distinction between morally negative and (just) factually negative might be the cause. Also, being sick is unintended (passive) while being deceitful is not (is active).

2 Fine-grained Polarity Lexicon

We aim at a compositional treatment of phrase- and sentence-level polarity. In order to assure high quality, we rely on a manually crafted polarity lexicon specifying the polarities of words (not word senses). Recently, fine-grained distinctions have been proposed that distinguish between various forms of positive and negative polarities, e.g. (Neviarouskaya et al. (2009)). For instance, the Appraisal Theory (Martin and White (2005)) suggests to distinguish between appreciation (*sick friend*), judgement (*deceitful friend*) and emotion (*angry friend*) .

We have adopted and slightly adapted the categories of the appraisal theory. Our French, German and English polarity lexicons comprise 15,700 single-word entries (nouns, adjectives, adverbs), manually annotated for positive and negative prior polarity where each class further specifies whether a word is factually, morally or emotionally polar. We also coded whether the word involves an active part of the related actor (where applicable) and whether it is weakly or strongly polar. Our ultimate goal is to combine this resource with our verb resources (described below) in order to predict the polarity of the verb arguments and to be able to deal with conflicts arising from violated polarity expectations of the verb.

Also part of our lexicon are shifters (inverting the polarity, e.g.,*a good idea* (positive) vs. *not a good idea* (negative)) and intensifiers and diminishers. Figure 1 shows the details of our (in the paper focussed) French lexicon. For the English version the following statistics holds: shifters 24, intensifiers 94, diminishers 20, positive nouns 583, negative nouns 1345, positive adjectives 1097, negative adjectives 1475. Figure 2 gives an overview of our labels together with same examples: The prefixes A, F and J denote appreciation (factual), affect and judgement, respectively. The German lexicon comprises 2103 negative nouns, 1249

word class	NEG	POS	DIM	INT	SHI	Total
Adjectives	1550	858	3	34	5	2450
Nouns	1332	508	1	10	5	1856
ALL	2917	1411	5	79	26	4438

Fig. 1. French polarity lexicon

Tag	Meaning	Examples
A POS	Appreciation Positive	optimisation, beautiful, productive
F POS	Affect Positive	sensitive, happiness, love
J POS	Judgment Positive	charity, fidelity, charming
A NEG	Appreciation Negative	illness, unstable, loss
F NEG	Affect Negative	hatred, mourn, afraid
J NEG	Judgment Negative	corrupted, dictator, torture
DIM	Diminisher	less, decreasing
INT	Intensifier	more, vast
SHI	Shifter	not, absence of

Fig. 2. Complete list of polarity labels with examples

positive nouns, 1482 positive adjectives, 1861 negative adjectives, 71 intensifiers, 15 diminishers and 14 shifters.

3 Phrasal Level Sentiment Composition

According to the principle of compositionality and along the line of other scholars (e.g. Moilanen and Pulman (2007)), after mapping polarity from the lexicon to the text, in the next step we calculate the polarity of nominal and prepositional phrases, i.e., based on the lexical marking and taking into account the syntactic analysis of a dependency parser, we conduct a composition of polarity for the phrases. In general, the polarities are propagated bottom-up to their respective heads of the NPs/PPs in composition with the other subordinates. To conduct this composition we convert the output of the dependency parser into a constraint grammar format and use the `vislcg3`-tools VISL-group (2013) which allows us to write the compositional rules in a concise manner.

Composition at the np level is straightforward, if the polarity labels do match (including the prefix). Independent of the prefix (A,F,J) is it save to induce that a positive adjective coupled with a positive noun yields a positive noun phrase (given that no shifters are around): the prefix is left unchanced, of course. So *a lucky* (A_POS) *donator* (A_POS) gives a positive np *lucky donator* (A_POS). The same is true for the other prefixed polarities (F_POS + F_POS = F_POS etc.)

The interesting cases are those noun phrases where the adjective has a different polarity from the noun. For instance, here is a couple of noun phrases taken from the (English) Gutenberg corpus where a F_NEG adjective is combined with a F_POS noun.

Here, the prefix stays, but the polarity needs to be fixed to either positive or negative. It seems clear from these cases that the negative polarity wins: *disappointed hope* is negative. The same is true for F_POS-F_NEG combiniation and the J prefix variants. So always NEG wins. However, an empirical analysis should confirm these hypotheses.

adjective	noun	adjective	noun
nervous	emotion	disappointed	affection
angry	passion	grief	joys
furious	passions	anxious	hopes
nervous	gratitude	angry	pleasure
disappointed	hopes	disappointed	love
angry	joy	sad	astonishment
unhappy	passions	sad	pleasure

Fig. 3. Examples of bi-polar phrases: F_NEG-F_POS combinations

But what if prefixes differ, e.g. a J_POS adjective collides with a F_NEG noun? Which prefix, which polarity should we keep? In order to make the problem behind them clear, we call these cases *bi-polar* noun phrases, although we believe that in most cases a decision can be taken: whether it is positive or negative. Even the prefix might be clear. See Figure 4 for some examples. We could

adjective	noun	adjective	noun
earnest	regret	sincerely	anxious
kindly	regret	wisest	sorrow
honest	concern	heroic	angers
noble	rage	honest	shame
lively	despair	decent	sorrow

Fig. 4. Examples of bi-polar phrases: J_POS-F_NEG combinations

just stipulate a composition rule but rather found it more appropriate to base such a decision on an empirical study (see section 5).

Before we turn to the other conflict scenario, where verbs and their polarity expectations are violated, we introduce the verb resource itself.

4 Polar Verbs: Effects and Expectations

In order to merge the polar information of the NPs/PPs on the sentence level one must include their combination via their governor which is normally the verb. Neviarouskaya et al. (2009) propose a system in which special rules for verb classes relying on their semantics are applied to attitude analysis on the phrase/clause-level. Reschke and Anand (2011) show that it is possible to set the evaluativity functors for verb classes to derive the contextual evaluativity, given the polarity of the arguments. Other scholars carrying out sentiment analysis on texts that bear multiple opinions toward the same target also argue that a more complex lexicon model is needed and especially a set of rules for verbs that define how the arguments of the subcategorization frame are affected - in this special case concerning the attitudes between them, see Maks and Vossen (2012).

Next to the evidence from the mentioned literature and the respective promising results, there is also a strong clue coming from error analysis concerning sentiment calculation in which verbs are treated in the same manner as the composition for polar adjectives and nouns described above. This shows up especially if one aims at a target specific (sentence-level) sentiment analysis: in a given sentence "*Attorney X accuses Bank Y of investor fraud.*" one can easily infer that *accuse* is a verb carrying a negative polarity. But in this example the direct object *Bank Y* is accused and should therefore receive a negative "effect" while the *Attorney X* – as the subject of the verb – is not negatively affected at all. Second, the PP *of investor fraud* is a modification of the accusation (giving a reason) and there is intuitively a tendency to expect a negative polarity of this PP - otherwise the accusation would be unjust (In the example given, the negative expectation matches with the composed polarity stemming from the lexically negative "fraud"). So it is clear that the grammatical function must be first determined (by a parser) in order to accurately calculate the effects and expectations that are connected to the lexical-semantic meaning of the verb.

Furthermore, the meaning of the verb (and therefore the polarity) can change according to the context (cf. "report a profit" (positive) vs. "report a loss" (negative) vs. "report an expected outcome"(neutral)). This leads to a conditional identification of the resulting verb polarity (or verbal phrase respectively) in such a manner that the polarity calculated for the head of the object triggers the polarity of the verb. In German, for instance, there are verbs that not only change their polarity in respect to syntactic frames (e.g. in reflexive form) but also in respect to the polarity of the connected arguments, too (see Fig. 5). We there-

German	English	Polarity
für die Kinder sorgen	to take care of the kids	positive
für Probleme[neg.] sorgen	to cause problems	negative
für Frieden[pos.] sorgen	to bring peace	positive
sich sorgen	to worry	negative

Fig. 5. Several examples for the use of the German verb "sorgen".

fore encode the impact of the verbs on polarity concerning three dimensions: effects, expectations and verb polarity. While effects should be understood as the outcome instantiated through the verb, expectations can be understood as anticipated polarities induced by the verb. The verb polarity as such is the evaluation of the whole verbal phrase. To sum up: in addition to verb polarity, we introduce effects and expectations to verb frames which are determined through the syntactic pattern found (including negation), the lexical meaning concerning polarity itself and/or the conditional polarity respective to the bottom-up calculated prevalent polarities. This results at the moment in over 120 classes of verb patterns with regard to combinations of syntactic pattern, given polarities in grammatical functions, resulting effects and expectations, and verb polarity.

As an example we take the verb class *fclass_subj_neg_obja_eff_verb_neg* which refers to the syntactic pattern (subject and direct object) and at the same time indicates which effects and/or expectations are triggered (here negative effect for the direct object). If the lemma of the verb is found and the syntactic pattern is matched in the linguistic analysis, then we apply the rule and assign the impacts to the related instances. However, the boundary of syntax is sometimes crossed in the sense that we also include lexical information if needed. For instance, if we specify the lemma of the concerning preposition in the PP as in *fclass_neg_subj_eff_reflobja_prepobj[um]_verb_neg* (in this case "um" (for); note the encoded reflexive direct object), we leave the pure syntax level.

As mentioned above, one of the goals is the combination of the resources (polarity lexicon and verb annotation). This combination provides us with new target specific sentiment calculations which were not possible in a compositional sentiment analysis purely relying on lexical resources and cannot be reliably inferred via a fuzzy criterion like nearness to other polar words. The effects and expectations of an instantiated syntactic verb pattern in combination with bottom-up propagated and composed polarity can therefore be used to approach the goal of sentence-level sentiment analysis based on a deep linguistic analysis. Furthermore our system offers a possibility to detect violations of expected polarities ("admire a deceitful friend"), i.e., if the bottom-up composed polarity and the effects or expectations coming from the verb frame have an opposite polarity. In our empirical study we wanted to find out whether verb expectations actually are violated in real texts and how reliably those cases could be identified. Our verb resources comprise 305 of these verbs for German, 210 for English and 320 for French.

5 Empirical Investigation

We carried out two different experiments: bi-polar noun phrase level polarity composition and verb-based polarity prediction. Our experiments are based on texts from LeSoir (5'800 articles) and articles from the news platform AgoraVox (4'300 articles), alltogether about 6 million words.

5.1 NP-level

Our hypothesis was: conflicting polarities in a noun phrase always result in a negative NP, especially we claim that

- a negative adjective reverses or negates the positive polarity of the noun that it modifies.
- a positive adjective functions as intensifier of the negative polarity of the noun that it modifies.

Figure 6 gives an overview on the most frequent combinations.

In order to verify this hypothesis, we randomly selected a sample of 20 cases for each of the 6 most frequent combination types from our result. The evaluation consisted of two steps: first, we evaluated whether the noun phrases are

Adj-Noun Combination	Frequency	Adj-Noun Combination	Frequency
1. A POS , A NEG	1041	4. A POS , J NEG	242
2. A NEG , A POS	691	5. A NEG , J POS	234
3. J POS , A NEG	349	6. J NEG , A POS	155

Fig. 6. Most frequent combinations

overall negative, positive or ambiguous. We attributed the value "yes", "no" or "ambiguous" to each noun phrase. The results are listed in Figure 7 below. According to our manual evaluation, 97 out of 120 selected conflict cases (which

Adj-Noun Combination	yes (negative)	no (positive)	ambiguous
A POS , A NEG	14/20 (70%)	4/20 (20%)	2/20 (10%)
A NEG , A POS	17/20 (85%)	2/20 (10%)	1/20 (5%)
J POS , A NEG	12/20 (60%)	4/20 (20%)	4/20 (20%)
A POS , J NEG	18/20 (90%)	0/20 (0%)	2/20 (10%)
A NEG , J POS	17/20 (85%)	2/20 (10%)	1/20 (5%)
J NEG , A POS	19/20 (95%)	0/20 (0%)	1/20 (5%)
Total	97/120 (81%)	12/120 (10%)	11/20 (9%)

Fig. 7. Composition results

corresponds to 81%) validate our hypothesis. Indeed, we can easily identify positive adjectives as intensifiers of their negative head nouns:

- *célèbre catastrophe* "famous catastrophe" (`A_POS adjective + A_NEG noun`)
- *glorieuse incertitude* "glorious uncertainty" (`J_POS adjective + A_NEG noun`)
- *violation délibérée* "deliberated violation" (`A_POS adjective + J_NEG noun`)

The following examples illustrate how negative adjectives act as negators or shifters of their positive head nouns:

- *goût amer* "bitter taste"(`A_NEG adjective + A_POS noun`)
- *fausses innocences* "false innocence"(`A_NEG adjective + J_POS noun`)
- *ambition cynique* "cynical ambition"(`J_NEG adjective + A_POS noun`)

We evaluated 19% of the selected cases as either ambiguous or as contradicting our hypothesis. A contradiction of the hypothesis means that the overall polarity of a noun phrase should have been computed as overall positive instead of overall negative, such as in the following cases:

- *sourire ravageur* "charming smile"(`A_NEG adjective + A_POS noun`)
- *bouleversante sincérité* "overwhelming sincerity"(`A_NEG adjective + J_POS noun`)
- *lutte antiterroriste* "antiterrorist fight"(`J_POS adjective + A_NEG noun`)

Based on our empirical study, we formulate one aggregation rule for each of the 6 analysed combination types:

- If an A POS adjective modifies an A NEG noun, the A POS adjective acts as an intensifier. The overall polarity of the NP is A NEG.
- If an A NEG adjective modifies an A POS noun, the A NEG adjective shifts the positive polarity of the noun. The overall polarity of the NP is A NEG.
- If an J POS adjective modifies an A NEG noun, the J POS adjective adds a further qualification to the noun. The overall polarity of the NP is A NEG.
- If a A POS adjective modifies an J NEG noun, the adjective acts as an intensifier of the noun. The overall polarity of the NP is J NEG.
- If a A NEG adjective modifies an J POS noun, the adjective shifts the noun. The overall polarity of the NP is reversed to J NEG.
- If an J NEG adjective modifies a A POS noun, the meaning and polarity type of the adjective overrule those of the noun. The overall polarity of the NP is J NEG.

Additional knowledge is needed in order to cope with the exceptions discussed above. These rules provide the best solution given our current resources.

5.2 Verb Level

We searched for verb expectation violations, since we believe that they form an interesting phenomenon. They might help identify parts of a text that represent controversial opinions or offending passages. Expectation violations are rare in our French corpus, given 450'000 sentences, only about 500 conflicts were found:

- 410 (81.18% A_NEG or A_POS conflict
- 72 (14.25% J_NEG or J_POS conflict
- 23 (4.55% F_NEG or F_POS conflict

For instance, if the verb has a positive expectation, a A_-F_- or J_NEG might cause a conflict. No clear conclusions can be drawn from our data concerning a general rule for these cases. However, we identified four different verb-specific conflict classes: the fine-grained polarity tags do determine whether a polarity conflict occurs or not, but they behave differently depending on the verbs. The categories and rules that we identified are the following:

1. `A_NEG`, `F_NEG` and `J_NEG` always cause conflicts, except if they are further labelled as `passive`.
   Verbs: alimenter, nourrir
2. `A_NEG` and `F_NEG` do not produce conflicts, except if they are further labelled as `strong`. `J_NEG` words and expressions always generate conflicts.
   Verbs: accepter, accueillir, accorder, adorer, aider, aimer, apprécier, encourager, défendre, permettre, privilégier, prôner, soutenir, suggérer
3. Nor `A_NEG`, `F_NEG` or `J_NEG` can cause conflicts. The verbs have a positive impact on the negative polarity: they diminish the negative polarity.
   Verbs: corriger, soulager

4. `A_NEG`, `F_NEG` and `J_NEG` always generate conflicts because
   (a) the verb intensifies negative expressions.
   Verbs: assurer, conforter, cultiver, favoriser
   (b) the proposition establishes a negative polarity effect or relation with regard to somebody or something.
   Verbs: mériter, désirer, offrir, profiter, promettre

For instance, the fourth category deals with verbs that always generate conflicts, regardless of the type of negative polarity they modify. We divided the verbs into two distinct subcategories (a and b). The verbs listed in the subcategory a), *assurer* "assure", *conforter* "strengthen", *cultiver* "cultivate", *favoriser* "favour", have, as far as our data shows, an intensifying effect when they modify negative expressions. The type of negativity is not relevant:

1. a. *favoriser le désespoir* "favour despair"
   b. *favoriser le terrorisme* "favour terrorism"
2. a. *cultiver le mensonge et la trahison* "cultivate lies and betrayals"
   b. *cultiver la haine* "cultivate hatred"

We have started to further explore verb expectation conflicts for the other languages. Especially for German, where we have a larger corpus (compared to French), namely the DeWaC corpus (see Baroni et al. (2009)) comprising 90 Million sentences. Hopefully, we can get a clearer picture given more data.

6 Related Work

No special attention is paid to bi-polar phrases in the literature. A simple composition rule is used, namely that positive and negative yields negative (e.g. Choi and Cardie (2008)). In this paper, we have had a closer look at these special noun phrases and tried to fix better tailored rules.

The role that verbs play in sentiment analysis is not so broadly acknowledged compared to adjectives and nouns. However, there are a few approaches that strive to clarify the impact of verbs, e.g. Chesley et al. (2006), Neviarouskaya et al. (2009) and Reschke and Anand (2011). Chesley et al. (2006) use verb classes and attach a prior polarity to each verb class. They show how these verb classes contribute as features to the accuracy of their approach. No attention is paid to the verb's arguments. This, however, is the primary focus of the work of Neviarouskaya et al. Neviarouskaya et al. (2009). In their approach, verb classes are used to specify effects on the grammatical roles of verbs (subject, ..).

An approach that focuses on the interplay between the polarity of the bearers of grammatical roles of a verb and the overall verb frame polarity is Reschke and Anand (2011). Again, verb classes are used, e.g. verbs of having, withholding, disliking and linking. Frame polarity depends on the polarity of the subject and the direct object. E.g. an instantiated verb frame of the verb "lack" is positive, if the subject is negative and the direct objects is positive ("your enemy lacks good luck" is positive). Nothing is derivable about the polarity preferences of these verbs, e.g. that "lack" has positive polarity preference for its direct object. This is what our approach reveals, so these two approaches are complementary.

7 Conclusion

We have introduced fine-grained lexical resources for French, German and English sentiment analysis. In order to properly compose word level polarity to phrase level polarity and finally clause level polarity the role of bipolar phrases needs to be clarified. We have carried out experiments with a French corpus in order to develop such composition rules.

We also have introduced novel verb resources, where verbs have effects and expectations on their arguments. Such a resource is useful in order to fix the contextual polarity of neutral noun phrases occuring as an argument of these verbs. They, so to speak, inherit the polarity expectation or effects. In this paper, we have, however, focussed on conflicts arising from a top down restriction that gets violated bottom up. Further work is needed in order to clarify how to deal with such violations. Our study suggests that it is verb-specific, but that it depends on our fine-grained polarity categories as well. This is future work. Detecting such violations might enable our sentiment analysis system to detect interesting text passages (e.g. controversial stance).

References

Baroni, M., Bernardini, S., Ferraresi, A., and Zanchetta, E. (2009). The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. In *Language Resources and Evaluation (LREC)*, pages 209–226.

Chesley, P., Vincent, B., Xu, L., and Srihari, R. K. (2006). Using verbs and adjectives to automatically classify blog sentiment. In *Proc. AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs*, pages 27–29.

Choi, Y. and Cardie, C. (2008). Learning with compositional semantics as structural inference for subsentential sentiment analysis. In *Proc. of EMNLP*.

Maks, I. and Vossen, P. (2012). A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems*, 53(4):680–688.

Martin, J. R. and White, P. R. R. (2005). *Appraisal in English*. Palgrave, London.

Moilanen, K. and Pulman, S. (2007). Sentiment composition. In *Proc. of RANLP-2007*, pages 378–382, Borovets, Bulgaria.

Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2009). Semantically distinct verb classes involved in sentiment analysis. In Weghorn, H. and Isaías, P. T., editors, *IADIS AC (1)*, pages 27–35. IADIS Press.

Reschke, K. and Anand, P. (2011). Extracting contextual evaluativity. In *Proc. of the Ninth Intern. Conf. on Computational Semantics*, pages 370–374.

VISL-group (2013). *VISL CG-3. http://beta.visl.sdu.dk/cg3.html*. Institute of Language and Communication (ISK), University of Southern Denmark.