

Guidelines for the Annotation of the Sentiment Corpus

Uladzimir Sidarenka

February 24, 2016

1 Overview

1.1 Introduction

In this assignment, your task is to annotate sentiments in a corpus of Twitter messages. We define *sentiments* as polar (either positive or negative) evaluative subjective opinions about some persons, subjects, or events. In this assignment, you have to annotate both – text spans denoting the opinions (*sentiments*) and text spans denoting the subjects and events being evaluated (*sentiment targets*). Additionally, you should also mark opinions’ holders (*sentiment sources*) and lexical elements which might significantly change the polarity or the intensity of a sentiment. These elements are:

- *emotional expressions*, which are words or phrases that unequivocally possess some evaluative lexical meaning by themselves (these are typically words like *hassen* (*hate*), *bewundern* (*admire*), *schön* (*nice*) etc.);
- *negations*, which are words or expressions that might completely flip the polarity of an emotional expression or a sentiment to the opposite (e.g. *nicht* gut (*not good*), *kein* Talent (*not a talent*) etc.);
- *intensifiers* and *diminishers* (or *downtoners*), which are words and expressions that might increase or decrease the evaluative sense of an emotional expression, respectively. Examples of intensifiers include words like *sehr* (*very*), *besonders* (*especially*), *insbesondere* (*particularly*) etc. Typical examples of diminishers are words like *ein wenig* (*a little*), *ein bisschen* (*a bit*), *gewissermaßen* (*to a certain degree*) etc.

After marking these elements, you should also specify the values of their attributes. A complete list of all elements along with the description of their possible attributes is given in Section 2. In Section 3 we also provide a short summary of the task. In Section 4, you then may find answers to some questions which caused particular difficulties during the previous runs of annotation. Finally, Section 5 gives a couple of complete annotation examples for some sentences from our corpus.

1.2 Terminology and Format

Terminology. Throughout this document, we use the term *markable* to denote an annotated span of text. The term *markable type* (or *markable tag*) refers to the tag assigned to that markable. Additional attributes associated with the annotated text spans are called *markable attributes*.

We do not make a distinction between the terms *opinions* and *sentiments* and use both words interchangeably throughout this text.

Format. In these guidelines, we rely on the following conventions regarding the text format.

We specify shell commands in gray boxes in **typesetting** font as shown in the example below:

```
echo 'Hello world!'
```

The **typesetting** font is also used for literal mentions of markable types, markable attributes, file names, directory paths, and executable commands.

Examples of words and phrases are given in *italics* and their respective English translation is provided in parentheses.

Examples of sentence annotations are shown in regular font. Text enclosed in markables is *emphasized* and surrounded by square brackets (e.g. [*markable text*]). The type of the markables is given as a subscript after the closing right bracket; optional markable attributes are specified after the type, separated from it by a colon, e.g.:

EXAMPLE 1.1. [[*Der neue Papst*]_{target} gilt als [*bescheidener*]_{emo-expression}, [*zurückgenommener*]_{emo-expression} Typ.]_{sentiment:polarity=positive}
 ([[*The new Pope*]_{target} is believed to be a [*sober*]_{emo-expression}, [*modest*]_{emo-expression} man.])_{sentiment:polarity=positive}

1.3 Annotation Tool

For annotating this corpus, you should use MMAX2 – a freely available annotation tool – which you can download under the following link:

http://sourceforge.net/projects/mmax2/files/mmax2/mmax2_1.13.003/MMAX2_1.13.003b.zip/download

After you have downloaded this file, you should unpack the received archive, change to the newly created directory 1.13.003/MMAX2 in your terminal shell and execute the following commands:

```
chmod u+x ./mmax2.sh
nohup ./mmax2.sh &
```

An MMAX2 window should then appear on your screen. If you have never used MMAX2 before, please read the document `mmax2quickstart.pdf` which you can find in the subdirectory MMAX2/Docs of the downloaded archive.

1.4 Corpus Files

You should also have received a copy of corpus files either as a tar-gzipped archive or via the version control system Git.

If you got a .tgz archive of the corpus, then unpack it using the command:

```
tar -xzf archive-name.tgz
```

After that, a directory called `sentiment` should appear in your current folder.

If you received access to the Git repository of the project, you should clone the project to your local computer using the command:

```
git clone ssh://hebe.ling.uni-potsdam.de/var/local/git/Depot/socmedia socmedia
```

A directory called `socmedia` should then appear in the current folder on your local computer. You can find your annotation files in the directory `lingsrc/corpus/sentiment/annotator-ANNOTATOR_ID` in the newly appeared `socmedia` folder (ANNOTATOR_ID is the ID number which was previously assigned to you by the author of these guidelines).

In order to load an annotation file into your MMAX2 program, please change to the MMAX2 window and click on the menu `File -> Load`. In the displayed pop-up window, select the path to the `sentiment/annotator-ANNOTATOR_ID` folder¹ and click on one of the *.mmax files found there. The chosen project should then be loaded into your MMAX2 editor.

If you have any difficulties with launching MMAX2 or loading project files into it, please contact the author of these guidelines via e-mail (`sidarenk@uni-potsdam.de`).

2 Tags and Attributes

In the following, we provide a short list of all markables and their possible attributes that will be used in this annotation:

¹Please, make sure that the path to the `sentiment` folder does not contain any white spaces. Otherwise, MMAX2 might fail to load the project.

- | | |
|--|--|
| 1. sentiment -markable with the attributes: | (a) polarity , |
| (a) polarity , | (b) intensity , |
| (b) intensity , | (c) sarcasm , |
| (c) sarcasm ; | (d) sentiment-ref ; |
| 2. target -markable with the attributes: | 5. intensifier -markable with the attributes: |
| (a) preferred , | (a) degree , |
| (b) anaph-ref , | (b) emo-expression-ref ; |
| (c) sentiment-ref ; | 6. diminisher -markable with the attributes: |
| 3. source -markable with the attributes: | (a) degree , |
| (a) anaph-ref , | (b) emo-expression-ref ; |
| (b) sentiment-ref ; | 7. and, finally, the negation -markable with |
| 4. emo-expression -markable with the at- | the attribute: |
| tributes: | (a) emo-expression-ref . |

A more detailed description of these markables and the values of their respective attributes is given in the next subsections.

2.1 sentiment

Definition. *Sentiments* are polar subjective evaluative opinions about people, subjects, or events.

According to our definition, a sentiment must always satisfy the following three conditions:

- it has to be **polar**, i.e. a sentiment should always reflect either positive or negative attitude to its respective target. Cases like *Ich glaube, er wird heute früher kommen* (*I think he will be earlier today*) should not be marked as **sentiment** because the attitude of the author is neither positive nor negative but neutral;
- it has to be **subjective**, i.e. you should not mark as **sentiments** mere statements of objective facts like, for example, *Beim Angriff wurden 14 Glasscheiben beschädigt* (*14 glass plates were broken during the attack*), even if you have your personal negative associations with the reported events. Sentiment instead should always unequivocally show *the personal opinion of the immediate author of an expression*;
- a sentiment has to be **evaluative**, which means that it should always refer to an explicit target and judge about its properties. You should not mark as **sentiments** cases like *Ich bin heute so glücklich* (*I am so happy today*) because these statements do not evaluate anything in particular but simply express emotions.

Example. Typical examples of sentiments are evaluative sentences like the one show in Example 2.1.

EXAMPLE 2.1. [*Ich mag den neuen James Bond Film nicht.*]_{sentiment}
 ([*I don't like the new James Bond movie.*]_{sentiment})

This sentence expresses a personal subjective opinion of the author, this opinion is polar and strictly negative, and it also has an explicit evaluation target – the *movie*. So, we put **sentiment** tags around this sentence.

We also consider contrastive comparisons to be a special type of evaluative opinions. But unlike other types of sentiments, comparisons typically express a relative subjective judgement, i.e. an object is regarded to be either better or worse than another compared object, but we usually do not know if the author actually likes or dislikes either of them. For this type of evaluations, we have introduced a special value of the **polarity** attribute – called **comparison** (cf. Table 1).

You should NOT mark as sentiments polar opinions for which its unknown if they are true or not. These typically are sentences like *Ich weiß nicht, ob ich meinen Bruder mag* (*I don't know if I like my brother*). In this example, neither we nor the author actually know if the author likes or dislikes her brother. Exceptions from this rule are cases like *Ich zweifle, dass er ein guter Mensch ist* (*I doubt that he is a good man*) or *Ich glaube nicht, dass er diesen Preis verdient hat* (*I don't think that he has deserved this award*) which express author's disagreement with some positive evaluations and, therefore, act themselves as negative judgements about the targets. Special care should be taken when dealing with questions and irrealis (cf. Questions 1, 2, 3, and 4 in FAQ Section of this document).

Sentiments can be ironic or sarcastic, i.e., the judgement expressed is not to be taken at face value. We mark this with a general attribute *sarcasm* (see Table 1) but do not distinguish between different kinds (irony versus sarcasm, etc.).

Boundaries. **sentiment** markables should encompass both the object being evaluated (the target) and the actual phrase phragment which expresses the evaluation (typically an emo-expression, if it exists). After determining these two elements, you should put the **sentiment** tags around the *minimal complete syntactic or discourse-level unit in which both target and evaluation expression appear together*.

In Example 2.2, for instance, the evaluated object is *Buch* (*book*), the evaluative expression is *langweiliges* (*boring*), and the minimal syntactic unit which simultaneously covers both of these elements is the noun phrase *ein langweiliges Buch* (*a boring book*). We therefore put **sentiment** tags around this noun phrase but do not put anything else inside them.

EXAMPLE 2.2. Auf dem Tisch lag [*ein langweiliges Buch*]_{sentiment}.
(There was [*a boring book*]_{sentiment} on the table.)

Sentiments are not restricted to just noun phrases, they can also be expressed by complete clauses or even multiple sentences (i.e. discourse-level units). The main point is that a **sentiment** span has to be *complete*, i.e. it should capture the common syntactic or discourse-level ancestor element of both evaluation and target and also include all other descendants of that common ancestor. Furthermore, a **sentiment** markable has to be *minimal*, i.e. it should only cover the lowest possible ancestor element of evaluation and target but should not include parents or siblings of this ancestor.

Example 2.3 shows how a sentiment relation can be expressed by a clause:

EXAMPLE 2.3. Wir akzeptieren das, weil [*wir alle ein bisschen in Petterson verliebt sind*]_{sentiment}.
(We accept this because [*we all are a little bit in love with Petterson*]_{sentiment}.)

In this sentence, the evaluative statement is made about *Petterson* who acts as sentiment's target. The author says that they all *in ihn verliebt sind* (*are in love with him*) which is her subjective evaluative opinion. Both target and evaluative expression appear together in one verb phrase with the head verb *sein* (*to be*). So, we mark this complete verb phrase including its grammatical subject *wir* (*we*) which is the syntactic descendant of the head verb.

From the boundary rules just described it follows that some elements that take part in the evaluation can be located outside the sentiment span. For example, this can happen with sources mentioned in a separate clause:

EXAMPLE 2.4. [*Klaus*]_{source} hat sich ein iphone gekauft und [*findet es ganz toll*]_{sentiment}.
([*Klaus*]_{source} has bought an iphone and [*likes it very much*]_{sentiment}.)

Attributes. After you have marked a **sentiment** span, you should next set the values of its attributes. Acceptable attributes with their meanings and values are given in Table 1.

2.2 target

Definition. *Targets* are objects or events that are being evaluated by a sentiment expression.

Because sentiments are required to be evaluative, there MUST always be at least one target for each sentiment relation. Conversely, if there is no annotated sentiment, we do not annotate targets.

Table 1: Attributes and values of **sentiments**.

Attribute	Value	Value's Meaning
polarity	<i>positive</i>	sentiment expresses positive attitude about its respective target, e.g. <i>Es war ein fantastischer Abend</i> (<i>It was a fantastic evening</i>);
	<i>negative</i> (default)	sentiment expresses negative attitude about its respective target, e.g. <i>Seine Schwester ist einfach unausstehlich</i> (<i>His sister is simply obnoxious</i>)
	<i>comparison</i>	sentiment expresses a comparison of two objects with preference given to one of them, e.g. <i>Mir gefällt das rote Kleid mehr als das blaue</i> (<i>I like the red dress more than the blue one</i>)
intensity	<i>weak</i>	sentiment expresses a weak evaluative opinion, e.g. <i>Der Auftritt war mehr oder weniger gut</i> (<i>The appearance was more or less good</i>)
	<i>medium</i> (default)	sentiment has a middle emotional expressivity, e.g. <i>Mir hat das neue Album gut gefallen</i> (<i>I enjoyed the new album</i>)
	<i>strong</i>	this sentiment expresses a very emotional polar statement, e.g. <i>Dieses Festival war einfach umwerfend!!!</i> (<i>This festival was simply terrific!!!</i>)
sarcasm	<i>true</i>	this polar attitude is derisive, i.e. its actual polarity is the opposite of its apparent form. (An apparent praise, for example, could be meant as a rebuke and vice versa. The actual sense, however, can only be inferred on the basis of world knowledge or reasoning.) An example of a sarcastic sentiment is the following passage: <i>Mein Jüngerer ist in der Prüfung durchgefallen. Klasse!</i> (<i>My youngest has failed his exam. Well done!</i>) In this case, you should set the polarity attribute of the sentiment to negative and the value of the sarcasm attribute to <i>true</i> .
	<i>false</i> (default)	no sarcasm is present – the polar attitude has its literal meaning; this is the default. setting

Occasionally, a target may be elided (*Ist gut.* / *Is good.*) and thus is not to be marked. However, for the reader of the tweet it must be possible to interpret the elision, i.e., to recover the missing constituent, in order for a sentiment to be present.

Example. An example of a sentiment target is given in Sentence 2.5:

EXAMPLE 2.5. Mein Bruder ist nicht begeistert von [*dem neuen Call of Duty*]_{target}.
(My brother is not impressed by [*the new Call of Duty*]_{target}.)

In this sentence, the author is telling us about the subjective opinion of her brother regarding the new version of a computer game. This new computer game is the object of the evaluation and we annotate it as **target**.

Boundaries. Similar to **sentiments**, you should put the **target** tags around the minimal complete syntactic or discourse-level units which denote the objects or events being evaluated. These are usually noun phrases (e.g. *Mir wird's schlecht, wenn ich [diese Werbung]_{target} im Fernsehen sehe* (*I feel sick when I see this [ad]_{target} on TV*)) or clauses (e.g. *Ich hasse wenn [Voldemort mein Shampoo benutzt]_{target}*. (*I hate when [Voldemort is using my shampoo]_{target}*)).

In case a target is an anaphoric expression (a pronoun) and its antecedent (a full NP) is present in the same tweet (be it within or outside of the sentiment span), both are being annotated and linked with an *anaph-ref* relation (see Table 2).

Occasionally, a Twitter hashtag or @-mention may serve as a target. Then they are marked as usual, without adding a specific attribute.

If a sentiment has multiple targets, you should mark each one of them separately (cf. Example 2.6).

EXAMPLE 2.6. Meiner Mutter haben [*Nelken*]_{target} und [*Dahlien*]_{target} immer gefallen.

(My mother has always liked $[carnations]_{\text{target}}$ and $[dahlias]_{\text{target}}$.)

Similar, in comparisons, you should also annotate each compared object separately. Additionally, for the object which is being dispreferred, you should also set the value of the **preferred** attribute to **false** (cf. Example 2.7).

EXAMPLE 2.7. Ich mag $[Domino-Eis]_{\text{target:preferred=true}}$ mehr als $[Magnum]_{\text{target:preferred=false}}$.
(I like $[Domino \text{ ice cream}]_{\text{target:preferred=true}}$ more than $[Magnum]_{\text{target:preferred=false}}$.)

Attributes. Further possible attributes of **targets** are shown in Table 2.

Table 2: Attributes and values of **targets**.

Attribute	Value	Value’s Meaning
preferred	<i>true</i> (default)	in comparisons, this value means that the respective target is being considered better than another compared object, e.g. <i>Die neue Frisur passt ihr garantiert besser als die alte</i> (The new hairstyle <i>suits her definitely better than the old one</i>);
	<i>false</i>	in comparisons, this value marks the target element which is being considered worse than its counterpart, e.g. <i>Die zweite Saison von Breaking Bad war viel spannender als die dritte</i> (The second season of <i>Breaking Bad</i> was much more exciting than the third one);
sentiment-ref	\longrightarrow (<i>directed edge</i>)	directed edge pointing from target to its respective sentiment . You need to draw this edge in two cases: <ul style="list-style-type: none"> • when the target is located at intersection of two different sentiments (in this case, you should draw an edge from target to sentiment, which this target actually belongs to), • when the target of an opinion is expressed outside the sentiment span;
anaph-ref	\longrightarrow (<i>directed edge</i>)	directed edge pointing from target expressed by a pronoun or pronominal adverb to its respective non-pronominal antecedent (in order to draw this edge, you also need to mark the antecedent as target)

2.3 source

Definition. Sentiment *sources* are immediate author(s) or holder(s) of evaluative opinions. These are typically writers of the messages or persons or institutions whose opinion is being cited.

If sentiment’s source of is not explicitly mentioned in the message, we assume it to be the author of the tweet. You need not annotate anything as **source** in this case.

In parallel with targets, sources are only annotated when a sentiment is present. (E.g., if a tweet contains a quotation with an explicit source, but there is no sentiment, than the source is not annotated.)

Example. An example of an explicit sentiment source is the pronoun *Sie* (*she*) in Example 2.8.

EXAMPLE 2.8. $[Sie]_{\text{source}}$ mag die neue Farbe nicht
($[She]_{\text{source}}$ doesn’t like the new color)

Note that in case of citations you should only mark the immediate person or the institution whose original opinion is being cited, but you should not mark the citing person as a **source** (cf. Example 2.9).

EXAMPLE 2.9. Laut Staatsanwalt soll die [*Angeklagte*]_{source} sich missbilligend über ihren Vorgesetzten geäußert haben.
 (According to the attorney, the [*defendant*]_{source} had made disapproving remarks about her boss.)

Boundaries. For determining the boundaries of **sources**, you should proceed in similar fashion as we did for **targets** and **sentiments** and only mark complete minimal syntactic units. Sources are most commonly expressed by noun phrases. And, similar to **targets**, if the source of a sentiment is expressed by multiple separate noun phrases, you should mark each of them separately (cf. Example 2.10).

EXAMPLE 2.10. [*Ihr*]_{source} und [*ihrer Mutter*]_{source} gefällt die neue Farbe nicht.
 (Neither [*she*]_{source} and [*her mother*]_{source} likes the new color)

Attributes. The attributes of the **source** tag are listed in Table 3. They are fully identical to the attributes of the **target** markables.

Table 3: Attributes and values of **sources**.

Attribute	Value	Value’s Meaning
sentiment-ref	→ (<i>directed edge</i>)	cf. Table 2
anaph-ref	→ (<i>directed edge</i>)	cf. Table 2

2.4 emo-expression

Definition. *Emo-expressions* are lexical items (i.e., words or idioms) that have a clearly distinguishable polar subjective meaning.

Example. An example of an emo-expression is the word *ekelhaft* (*disgusting*) in Example 2.11.

EXAMPLE 2.11. Beim Aufräumen des Zimmers haben wir einen [*ekelhaften*]_{emo-expression} Teller mit verschimmeltem Essen unter dem Bett gefunden.
 (When we cleaned the room, we found a [*disgusting*]_{emo-expression} plate with moldy food under the bed.)

This term is unequivocally polar, since it expresses a negative attitude towards its target, and subjective, since it stands for a personal perception of the author.

A notable difference between **emo-expressions** on the one hand and **sources** and **targets** on the other hand is that **emo-expressions** should always be annotated in text no matter whether a target-oriented sentiment exists in their vicinity or not. **sources** and **targets**, however, should only be marked along with a **sentiment** element they pertain to. Therefore, in a sentence like *Ausgezeichnet!* (*Excellent!*), you should mark the single present word as an **emo-expression** but should not annotate any **sentiments** in such cases (as the target of this expression is unclear).

Another distinguishing feature of **emo-expressions** is that they should only encompass *lexical* items (i.e., words or idioms), whereas **sources** and **targets** should typically comprise *syntactic* elements (i.e., noun and verb phrases). In order to decide whether a given element is a lexical item or not, you should think about whether this element could appear as a separate entry in a defining or idiomatic dictionary. In case it can, the language unit under scrutiny is a lexical element, otherwise it is a free noun or verb phrase which should not be regarded as an **emo-expression**.

An example of a lexical item is the word *cooler* (*cool*) in the sentence below:

EXAMPLE 2.12. Das war ein echt [*cooler*]_{emo-expression} Film.
 (This was a really [*cool*]_{emo-expression} movie.)

Since the lemma of this adjective might appear as a lexicon entry, and its meaning is definitely positive and subjective, this word should be annotated as an **emo-expression**. It is, however, unlikely that the complete noun phrase *ein cooler Film* (*a cool movie*) would be included into a defining dictionary, therefore, one should not consider this whole phrase as an **emo-expression** instance.

In rare cases, however, some ad hoc phrases might appear like lexical items but be too rare to be included into a lexicon. An example of such phrase is provided below:

EXAMPLE 2.13. Norwegen erinnert an die frühe Ursula von der Leyen.
(Norway reminds one of early Ursula von der Leyen.)

In this sentence, the phrase *die frühe Ursula von der Leyen* (*early Ursula von der Leyen*) is a metaphoric term which was created spontaneously for this particular sentence only. Nevertheless, we can replace this term with an almost equivalent lexicon synonym *Einzelgänger* (*loner*). In such cases, when we can provide a dictionary equivalent for an ad hoc term, we can also consider the ad hoc phrase as an **emo-expression**, but this should be regarded as an exception rather than regularity.

Another difficulty when annotating **emo-expressions** might arise from the fact that many words and idioms are ambiguous and can have several different lexical meanings, but only some of these meanings might be polar and subjective. In these cases, you should only mark a lexical item as an **emo-expression** if its actual sense in the given context is polar. If the active meaning of the lexeme is objective, you must not annotate it (cf. Example 2.14).

EXAMPLE 2.14. Dieser Wein ist ein echtes [*Juwel*]_{emo-expression} in meiner Kollektion.
(This wine is a real [*jewel*]_{emo-expression} in my collection.)
Koh-i-Noor ist das teuerste Juwel heutzutage.
(Koh-i-Noor is the most expensive jewel nowadays.)

In the above example, the meaning of the word *Juwel* (*jewel*) is metaphoric and subjective in the first sentence but literal and objective in the second. So you should only annotate this word as an **emo-expression** in the former case but not mark it as such in the latter.

Finally, it might be difficult to decide whether a polar term is subjective or not. Prominent examples of such borderline cases are the so-called subjective facts, e.g., *Tod* (*death*), *Krankheit* (*disease*), *Bombe* (*bomb*), *wählen* (*elect*) etc., which some people perceive as subjective expressions while others regard them as objective facts. For such cases, we have introduced a special attribute **subjective_fact**, which you can set to **true** when you encounter such expressions and think that they also belong to the **emo-expression** class.

To summarize, when annotating emotional expressions in corpus, you should proceed as follows:

1. determine an exact word or phrase which, in your opinion, might have a polar lexical meaning;
2. decide whether this word or idiom can appear as a separate entry in a monolingual dictionary or lexicon or not;
3. if it can, decide which lexical meanings in general this phrase might have and which of them are polar;
4. look whether the currently active meaning of this term in the given context is one of the polar senses;
5. if it is, annotate the given term as an **emo-expression** and decide whether it is subjective or not;
6. in the latter case, set the attribute **subjective_fact** of the newly created **emo-expression** to the value **true**.

Boundaries. **emo-expressions** are typically expressed by:

- nouns, e.g. *Held* (*hero*), *Ideal* (*ideal*), *Betrüger* (*fraudster*) etc.;

- adjectives or adverbs, e.g. *schön* (nice), *zuverlässig* (reliably), *hinterhältig* (devious), *heimtückisch* (insidiously) etc.;
- verbs, e.g. *lieben* (to love), *bewundern* (to admire), *hassen* (to hate) etc.;
- idioms, e.g. *auf die Nerven gehen* (to get on one's nerves) etc.;
- smileys, e.g. :), :-(, ☺, ☹ etc.

If an **emo-expression** is formed by an idiomatic phrase, you should always annotate the complete idiom. For verbs which take on an evaluative sense only if used with certain prepositions (e.g. *to go for sth.* in the sense of *to like*), you should annotate both the verb and the preposition as a single markable (please refer to the MMAX manual to see how to annotate discontinuous spans).

Occasionally, a Twitter hashtag may serve as an emo-expression (*#excellent*). They are marked as usual, without adding a specific attribute.

Attributes. When determining the value of the **polarity** attribute of an **emo-expression**, you should disregard any possible contextual modifiers like intensifiers or negations and set the value of this attribute to the lexical (or also called *prior*) polarity of the phrase (the one it would have without any negations and other modifiers) (cf. Example 2.15).

EXAMPLE 2.15. Es war keine [*gute*]_{emo-expression:polarity=positive} Idee.
(It was not a [*good*]_{emo-expression:polarity=positive} idea.)

Also, when determining the value of the **polarity** attribute of an **emo-expression**, you should analyze its polarity from the perspective of the subject or event which is being evaluated (in case when such subject is present in the context). This means that in cases like *Ich vermisse meine Freundin* (*I miss my girlfriend*), the polarity of the emo-expression *vermissen* (*to miss*) is still positive because the author evidently has a positive attitude to the girlfriend even if he experiences sadness because of her absence.

Further attributes of **emo-expressions** include **intensity**, **sarcasm**, **subjective_fact**, and **sentiment-ref**. Possible values and descriptions of these attributes are summarized in Table 4.

2.5 intensifier

Definition. *Intensifiers* are elements which increase the expressivity and the polar sense of an emotional expression.

Example. An example of an intensifier is the word *sehr* (*very*) in Example 2.16.

EXAMPLE 2.16. Wir suchen eine [*sehr*]_{intensifier} zuverlässige Polin als Haushaltshilfe.
(We are looking for a [*very*]_{intensifier} reliable Polish woman as domestic help.)

Boundaries. Intensifiers are usually expressed by adverbs or adjectives like *sehr* (*very*), *sicherlich* (*certainly*) etc., but other ways of expressing them are still possible (cf. Example 2.17). However, we only regard separate lexical items as intensifiers: iterated letters, as in *gooood* are not annotated, and neither are punctuation marks (such as exclamation marks).

EXAMPLE 2.17. Dieser Junge ist stark [*wie ein Pferd*]_{intensifier}.
(This boy is strong [*as a horse*]_{intensifier}.)

Attributes. An **intensifier** should always relate to some **emo-expression** and you should also always explicitly show that relation by drawing an edge attribute from **intensifier** to its respective **emo-expression** markable.

Further possible attributes of **intensifiers** are shown in Table 5.

Table 4: Attributes and values of **emo-expressions**.

Attribute	Value	Value’s Meaning
polarity	<i>positive</i>	emotional expression has a positive evaluative meaning, e.g. <i>gut</i> (<i>good</i>), <i>verhimmeln</i> (<i>to ensky</i>), <i>Prachtkerl</i> (<i>corker</i>) etc.
	<i>negative (default)</i>	emotional expression has a negative evaluative meaning towards its target, e.g. <i>versauen</i> (<i>to botch up</i>), <i>rotzig</i> (<i>snotty</i>), <i>Dreckskerl</i> (<i>scum</i>) etc.
intensity	<i>weak</i>	emo-expression has a weak evaluative sense, e.g. <i>solala</i> (<i>so-so</i>), <i>nullachtfünfzehn</i> (<i>vanilla</i>), <i>durchschnittlich</i> (<i>mediocre</i>) etc.
	<i>medium (default)</i>	emo-expression has middle stylistic expressivity, e.g. <i>gut</i> (<i>good</i>), <i>schlecht</i> (<i>bad</i>), <i>robust</i> (<i>tough</i>) etc.
	<i>strong</i>	emo-expression expresses a very strong positive or negative evaluation, e.g. <i>allerbeste</i> (<i>bettermost</i>), <i>zum Kotzen</i> (<i>to make one puke</i>), <i>Kacke</i> (<i>shit</i>) etc.
sarcasm	<i>true</i>	emo-expression is derisive, i.e. its actual polarity is the opposite of its apparent form. (This means that an apparent praise which appears in text is in fact meant as a rebuke and vice versa. The actual sense, however, can only be inferred on the basis of world knowledge or reasoning.)
	<i>false (default)</i>	no sarcasm is present – the polar attitude has its literal meaning; this is the default setting
subjective_fact	<i>true</i>	the emo-expression represents a possibly objective circumstance, which, however, might have a clear polar association or connotation, e.g., <i>Nordkorea</i> (<i>North Korea</i>), <i>Atomtest</i> (<i>atom test</i>), <i>Diabetes</i> (<i>diabetes</i>) etc.
	<i>false (default)</i>	the emo-expression expresses a purely subjective attitude
uncertain	<i>true</i>	the annotator is uncertain whether the given term forms an emotional expression or not, e.g. <i>angetrunken</i> (<i>half-drunk</i>), <i>einfältig</i> (<i>naïve</i>)
	<i>false (default)</i>	the given term unequivocally forms an emo-expression
sentiment-ref	\rightarrow (<i>directed edge</i>)	arrow pointing to the sentiment which this emo-expression belongs to. You should only draw this edge if an emo-expression is located in the overlapping of two sentiment spans or outside of the sentiment span which it belongs to

2.6 diminisher

Definition. *Diminishers* are words or phrases that decrease the polar lexical sense of an **emo-expression**.

Example. In Example 2.18, the diminisher is expressed by the adverb *weniger* (*less*).

EXAMPLE 2.18. [*Weniger*]_{diminisher} erfolgreiche Unternehmen verzichten auf externe Berater.

The [*less*]_{diminisher} successful companies do not use external consultants.

Attributes. Similar to intensifiers, diminishers should always relate to some emotional expression and you should also explicitly show this relation by drawing an edge attribute.

The attributes of diminishers mainly correspond to that of intensifiers. The only difference concerns the **degree** attribute which shows how strong an intensifier *increases* but a diminisher *decreases* the lexical sense of an emo-expression. A list of possible attributes for the **diminishers** is summarized in Table 6.

Table 5: Attributes and values of **intensifiers**.

degree	<i>medium (default)</i>	the intensifier moderately increases the polar sense of the emotional expression, e.g. <i>ziemlich (quite)</i> , <i>recht (fairly)</i> etc.
	<i>strong</i>	the intensifier strongly increases the polar sense and stylistic markedness of the emotional expression, e.g. <i>sehr (very)</i> , <i>super (super)</i> , <i>stark (strongly)</i> etc.
emo-expression-ref	\longrightarrow (<i>directed edge</i>)	a directed edge pointing from the intensifier to the emo-expression whose meaning is being intensified

2.7 negation

Definition. *Negations* are elements which turn the polarity of an **emo-expression** to the complete opposite.

Example. In Example 2.19, for instance, the negative article *kein (not)* makes the *contextual* polarity of the word *interessant (interesting)* to be negative, even though the prior polarity of this word is unequivocally positive.

EXAMPLE 2.19. Diese Geschichte war überhaupt nicht [*interessant*]_{negation}!
This story was [*not*]_{negation} interesting at all!

The role and the meaning of negations are closely related to that of diminishers. In order to help you better differentiate between these elements, we have listed the most obvious differences between the two classes:

- *Semantic differences.* While diminishers only decrease the lexical sense of an emo-expression, a part of this original sense still remains active (i.e. *a hardly understandable speech* is still understandable); negations, on the contrary, fully deny that meaning and turn it to the complete opposite (*a not understandable speech* is absolutely unintelligible);
- *Part-of-speech differences.* While diminishers are usually expressed by adjectives or adverbs, negations are typically represented by the negative article *kein (no)*, the negation particle *nicht (not)*, or adjectives or verbs, e.g. *Es ist sehr zweifelhaft, dass die neue Version von Windows besser wird (It is very doubtful that the new Windows version will be any better)*

Attributes. The only attribute of negations is the mandatory edge **emo-expression-ref**. You should draw this edge from the **negation** to the **emo-expression** being negated. Like intensifiers and diminishers, negations should always relate to at least one **emo-expression**.

3 Summary

To summarize, your task in this assignment is to find subjective evaluative opinions about some subjects or events. You should annotate these opinions with the **sentiment** tags and determine the polarity and the intensity of the expressed attitudes. After that, you should annotate subjects and

Table 6: Attributes and values of **diminishers**.

degree	<i>medium (default)</i>	diminisher moderately decreases the polar sense of its respective emo-expression , e.g. <i>wenig (few)</i> , <i>bisschen (little)</i> etc.
	<i>strong</i>	diminisher strongly decreases the polar sense of the emo-expression , e.g. <i>kaum (hardly)</i> etc.
emo-expression-ref	\longrightarrow (<i>directed edge</i>)	see Table 5

Table 7: Attributes and values of **negations**.

emo-expression-ref	\longrightarrow (<i>directed edge</i>)	an edge from negation to the emo-expression being negated
--------------------	--	---

events which are being evaluated and mark them as **targets**. The holders of the opinions should be annotated as **sources**. Both, **sources** and **targets** can only exist in the presence of a **sentiment**.

Another important task is to annotate words and phrases which convey a polar evaluative meaning by themselves. We call these words **emo-expressions** and you should always annotate them regardless of whether a sentiment relation is present or not. If an **emo-expression** is intensified, diminished, or negated by another word or phrase, you should also annotate this modifying element as **intensifier**, **diminisher**, or **negation**, respectively.

4 FAQ

This section provides some examples of difficult and controversial annotation cases and gives possible solutions to them. Please read them carefully before you start doing the annotation.

1. Q: Should I annotate sentiments in questions?

A: It primarily depends on the type of the question. You should typically distinguish two cases:

- If it is a *yes-no-question* or *wh-question* which asks whether a particular sentiment statement is true or not, then you should not annotate this sentence as **sentiment** because the validity state of this evaluation is unknown. In Example 4.1, for instance, we do not know whether the asked person actually likes or dislikes her new skirt, so we do not annotate sentiment in this case;

EXAMPLE 4.1. Gefällt dir der neue Rock?
(Do you like the new skirt?)

- If this is a *wh-question* which asks about the reasons or some extra aspects of a polar opinion but does not raise the truth of this opinion to question, then you should mark a sentiment relation in this case (cf. Example 4.2).

EXAMPLE 4.2. [*Warum hasst du deine Schwester?*]_{sentiment}
([*Why do you hate your sister?*]_{sentiment})

2. Q: Should I annotate sentiments in wishes?

A: Basically, yes. If someone expresses a wish to get or to do something, then this person typically also has a positive attitude to the desired object or activity (cf. Examples 4.3 and 4.4).

EXAMPLE 4.3. [*Habe sooooo Lust auf [einen Dattel / Bananen Milchshake]*]_{target .. :-)} *₋ *₋]_{sentiment:polarity=positive}
([*Am sooooo up for [a date / banana milk shake]*]_{target .. :-)} *₋ *₋]_{sentiment:polarity=positive})

EXAMPLE 4.4. [*Ich will [jetzt nach Hause gehen]*]_{target}]_{sentiment:polarity=positive}
([*I now want [to get home]*]_{target}]_{sentiment:polarity=positive})

This rule also applies to cases, when the author wants another person or object to get a particular property or to do something (cf. Example 4.5).

EXAMPLE 4.5. [*Ich möchte, dass [das neue Modell weniger Kraftstoff verbrauchen würde]*]_{target .}]_{sentiment:polarity=positive}
([*I want that [the new model consumed less fuel]*]_{target .}]_{sentiment:polarity=positive})

If you think that the sentence also expresses an evaluation of the object for which a particular property is wished (in the above case, it would be *the new model*), you are also free to annotate an additional sentiment with that object as a **target** (cf. Example 4.6).

EXAMPLE 4.6. [*Ich will, dass [das Auto]_{target} weniger Kraftstoff verbrauchen würde.*]_{sentiment:polarity=negative}
 ([*I want that [the new model consumes less fuel]_{target}.*]_{sentiment:polarity=negative})

It might however not always be the case that the object for which some action or property is desired is actually being evaluated (cf. *Bruder* (*brother*) in Example 4.7).

EXAMPLE 4.7. *Ich will, dass mein Bruder mit mir in den Zoo geht.*
 (I wish my brother would go with me to the zoo.)

You should also take special care when dealing with suggestions and recommendations. While recommendations might presuppose an appraisal in some cases (cf. Example 4.9), they can also be completely legitimate objective sentences as well (cf. Example 4.9).

EXAMPLE 4.8. [[*Die Regierung [soll]_{emo-expression:polarity=positive} mehr für die Umwelt tun*]_{target}]_{sentiment:polarity=positive}
 ([*The government [should]_{emo-expression:polarity=positive} do more for the environment*]_{target}]_{sentiment:polarity=positive})

EXAMPLE 4.9. *An der nächsten Kreuzung sollst du nach links abbiegen.*
 (You should turn left at the next crossing)

You should not consider as sentiments recommendations in advertising slogans, since these not necessarily express the real opinion of the authors:

EXAMPLE 4.10. *Kauft jetzt den neuen Staubsauger von Bosch.*
 (Buy the new vacuum cleaner from Bosch now)

3. Q: Should I annotate sentiments in conditional sentences?

If conditional sentence describes some *external* condition, under which the author would like or dislike a particular thing or event, then you should annotate the whole expression as a **sentiment** and the (dis-)liked thing as a **target** (cf. Examples 4.11 – 4.14).

EXAMPLE 4.11. [*Wenn es regnet, mag ich es immer, [vor dem Fenster zu sitzen]_{target}.*]_{sentiment:polarity=positive}
 ([*I always like [to sit in front of the window]_{target}, if it is rainy.*]_{sentiment:polarity=positive})

EXAMPLE 4.12. [*Wenn das Wetter besser wäre, würde ich gern [joggen gehen]_{target}.*]_{sentiment:polarity=positive}
 ([*If the weather was better, I would like [to go jogging]_{target}.*]_{sentiment:polarity=positive})

EXAMPLE 4.13. [*Selbst wenn es keine Alternative gäbe, würde mir [dieses Auto]_{target} nicht gefallen.*]_{sentiment:polarity=negative}
 ([*Even if there was no other alternative, I would not like [this car]_{target}.*]_{sentiment:polarity=negative})

EXAMPLE 4.14. [*Wenn ich gesund wäre, würde ich gern [mit euch campen]_{target}.*]_{sentiment:polarity=positive}
 ([*If I wasn't sick, I would gladly [go camping with you]_{target}.*]_{sentiment:polarity=positive})

If, on the other hand, the condition describes some *internal* change of the object or an action which the object inherently should do, in order that the author liked or disliked it, then you should only annotate that condition as a **target** and the whole expression as a **sentiment** (cf. Examples 4.16 – 4.18)².

²In Example 4.17, we have annotated all predicates of the target sentence with one **target** tag because only the joint action of the Pope is a sufficient condition for the author to love him. In other words, the author has positive attitude not to the each potential deed of the Pope but to all of these deeds as a whole. This is the only possible exception from our rule that we mark each conjoined target separately.

EXAMPLE 4.15. [*Wenn [dieses Auto weniger Kraftstoff verbrauchen würde]*_{target}, *würde ich es gerne kaufen.*]_{sentiment:polarity=positive}

(*[If [this car consumed less fuel]*_{target}, *I would definitely buy it.*]_{sentiment:polarity=positive})

EXAMPLE 4.16. [*Wenn [dieses Auto weniger Kraftstoff verbrauchen würde]*_{target}, *würde ich es gerne kaufen.*]_{sentiment:polarity=positive}

(*[If [this car consumed less fuel]*_{target}, *I would definitely buy it.*]_{sentiment:polarity=positive})

EXAMPLE 4.17. [*RT @VanessaLeii: Wenn [er jetzt raus kommt, die Arme hebt und "Don't cry for me Argentina" singt]*_{target}, *mag ich ihn.*]_{sentiment:polarity=positive} #Papst

(*[RT @VanessaLeii: If [he now comes out, raises his hands, and starts singing "Don't cry for me Argentina"]*_{target}, *I will love him.*]_{sentiment:polarity=positive} #Pope)

EXAMPLE 4.18. [*Wenn das Camping nicht so viel Aufwand machen würde, würde ich es gerne machen.*]_{sentiment:polarity=positive}

(*[If [camping wouldn't mean so much work]*_{target}, *I would love it.*]_{sentiment:polarity=positive})

Please notice the difference between the Examples 4.14 and 4.18. In both sentences, the judged object is *camping*. But in the former case, the author would love the camping if the author's state would change (in this case, the author's state is an external object with regard to camping). In the latter case, the author would love camping, if the camping's properties would change (which in that case is an internal change of the evaluated object).

We admit, however, that not all cases of conditionals can be covered by the above rule of thumb. So, if you see other evaluations in conditional sentences, you can also annotate additional sentiments or also annotate completely different than the way we have suggested. We admit that multiple interpretations are possible in these cases and do not enforce you to agree with our view of these phenomena.

4. Q: Should I annotate sentiments in irrealis sentences?

A: Most irrealis sentences which might express sentiment fall in two major categories: 1) implicit wishes and 2) conditions.

Typical examples of implicit wishes are cases like *Es wäre schön, wenn* (*It would be nice, if ...*) or *Es wäre schrecklich, wenn ...* (*It would be terrible, if ...*). In both of these cases, you should proceed similarly as we did for explicit wishes (cf. Question 2) and mark the whole wish expression as a **sentiment**, then annotate the wished property or the event as a **target**, and set the **polarity** value of that **sentiment** to **positive**, if the property or event is desirable, and to **negative** otherwise.

For conditional sentences, please refer to Question 3 for instructions.

5. Q: Should I mark sentiments in insults?

A: If you can locate the target, then yes. For example, in sentence *Du bist ein Idiot!* (*You are an idiot!*), *Du* (*You*) is the target of a negative evaluation. On the other hand, curses like *Idiot!* (*Idiot!*) do not have any explicit target and, therefore, should not be annotated as a **sentiment** according to our definition.

6. Q: Should I annotate sentiments in defenses?

A: Usually not. If a soldier defends his position or a PhD student defends her thesis, it does not necessarily imply that he or she likes it. The same is true in cases when someone defends another person in a dispute.

7. Q: Should I annotate as sentiments sentences which do not have any explicit emotions except for the smiley at the end, e.g. "kam bis heute nichts an :("?

A: If the smiley shows author's attitude to the object or event described in the tweet, then yes, you should annotate such cases as sentiments. If, on the contrary, the emoticon only serves politeness or phatic purposes, then you should not annotate it. We should, however, note that many examples are boundary cases and it will often depend on your interpretation. As a possible help for making decisions on such tweets, we suggest you to look at the type of the emoticon in use, because certain types are more often associated with judgements. Negative smileys like “:(“ or “☹”, for example, usually tend to appear with negative sentiments (cf. example in question); positive smileys, on the contrary, are much more ambiguous and typically only express an evaluative judgement if they show satisfaction or dissatisfaction of the writer with the facts stated in tweets; the winking smiley (e.g. ;-)) is by far the most ambiguous emoticon and it is only rarely involved in a sentiment relation.

8. **Q:** Should I annotate as sentiments cases like *etw. zustimmen (to agree with sth.)*, *etw. unterstützen (to support sth.)*, *sich für etw. entscheiden (to opt for sth.)*, and *j-m etw. vorwerfen (to accuse so. of sth.)*?

A: These cases are a little bit tricky because subjective and objective information are mixed here. But we would rather say “yes” unless the context strongly suggests that the expressed information is purely objective. For example, if an attorney accused a defendant of a crime in the court, she would basically do her job and it would not necessarily be true that she had any personal attitude to the defendant. On the contrary, if I accused someone of mean behavior, it would usually be my subjective judgement and, therefore, a sentiment. The same is true for support feelings: if a person supports someone's opinion, she is usually judging positively about it. This, however, may not always be the case.

9. **Q:** How would you annotate the following cases of comparisons?

- *Seehofer hat die Grünen ausgeschlossen , aber die Linke nicht (Seehofer has excluded the Greens, but not the Left)*;

A: Without any further context, I cannot see any sentiment relation here. So, I would probably not annotate anything.

- *Lieber starke Mitte statt linker Rand (Better strong middle than left edge)*;

A: This is a comparison with *starke Mitte (strong middle)* as the preferred target, and *linker Rand (left edge)* as the dispreferred one;

- *Die #spd wird lieber mit den rechten von #cdu, #csu koalieren als mit der #linke (The #spd will better form a coalition with rightists from the #cdu, #csu than with the #linke)*;

A: Here again is a comparison with the *#spd* as a source, the *#cdu*, *#csu* as the preferred targets, and the *#linke* as the dispreferred target;

- *Die #AfD + vereinigt mehr ökonomische Kompetenz als alle etabl. Parteien + Bunde... (The #AfD + combines more economic expertise than all established parties + federal...)*;

A: Again, a comparison with *#AfD* as the preferred target and *established parties* and *federal* as the dispreferred ones;

- *Freiheit statt Bevormundung (Freedom instead of paternalism)*;

A: Comparison, with *freedom* as the preferred target and *paternalism* as the dispreferred one;

- *Fettarme Milch hat mittlerweile mehr Prozent wie die FDP (lowfat milk has meanwhile more percents than the FDP)*;

A: I would rather say that this is a sarcasm about the FDP. Because we usually cannot compare a bottle of milk with a political party. If we do so, then usually in order to kid about this party;

- *Was ist der Unterschied zwischen einem Smart und der FDP ? Der Smart hat wenigstens 2 Sitze :) (What is the difference between a Smart and the FDP? The Smart has at least two seats);*

A: The same as the previous question – sarcasm about the FDP;

10. **Q: How should I determine the intensity of a comparison?**

A: As for the other type of sentiments, you should estimate the stylistic expressivity of the sentence. If a sentence expresses a strong emotional evaluation, then you should set the **intensity** attribute of that **sentiment** to **high**. If, on the other hand, the sentence rather makes an objective statement of facts, then you should mark the **intensity** of such **sentiment** (if it ever should be extracted) to **medium** or **weak**.

For example, in the sentence *this lousy Telekom is waaaaaay less reliable than O2*, the strength and the stylistic expressiveness of the sentence are much higher than in the sentence *Telekom has a less reliable connection than O2*. Consequently, we should set the value of the **intensity** attribute in the former case to **strong** and in the latter case to **medium**.

11. **Q: It is said that we should disregard negations when determining the polarity of an emo-expression. What about sentiments, shall we take into account negations there when determining their polarity?**

A: Yes. The polarity of an **emo-expression** represents the polar sense of that single lexical item. The polarity of a **sentiment**, on the contrary, shows the joint meaning of the whole phrase, so negations should be taken into account if they affect this polarity.

12. **Q: Is it possible that sources and targets are expressed by other means than the ones described in these guidelines?**

A: Yes. These guidelines are in no way exhaustive, they should only give you a better intuition of how sources or targets might typically look like.

13. **Q: What is target in the example *a really nice weekend*. The whole phrase?**

A: No, it is only the word *weekend*. *really* is an intensifier and *nice* is an emo-expression. The whole noun phrase *a really nice weekend* forms a sentiment with positive polarity.

14. **Q: What is target in cases like *Now, we will begin with hair coloring. Cooool!!!*. The whole sentence or only “hair coloring”?**

A: In this case, both the whole sentence and the noun phrase could be considered as targets, because the verb phrase is in fact semantically tantamount to the noun phrase. If one is happy about *hair coloring*, then she is also happy about the beginning of the hair coloring. The same is true, for example, about a postal package and the arrival of that package. Since information in the verb phrase is usually more elaborate and specific than in the noun phrase, we would recommend you to annotate the whole verb phrases, i.e. clauses, in such cases.

15. **Q: How should I judge whether a word is an emo-expression?**

A: Emotional expressions are usually abstract concepts that are strongly associated with some subjective polar feelings. These expressions will serve as our primary anchors for automatically finding sentiments in texts. Thus, if you can imagine the a given abstract word can be used in some context where it would positively or negatively characterize something, or if you have a strong polar subjective feeling associated with that word or phrase then you should mark it as an emo-expression.

For example, we usually associate negative emotions with words like *Betrug* (fraud), *Schuld* (guilt), or *vorwerfen* (accuse of). Moreover, one can say that someone *begeht einen Betrug* (commits a fraud), *hat die Schuld für den Eklat* (is at fault for the scandal), or *wird Unehrlichkeit vorgeworfen* (is accused of dishonesty), and it would negatively characterize the target person. Therefore, we should regard these words as emo-expressions. On the other hand, concrete terms

like *Stein* (*stone*) or *Krebs* (*cancer*) should not be considered as emo-expressions since their primary meaning is concrete and objective. Expressions which are associated with emotions but do not have a distinct polarity, like exclamation marks, for example, should not be marked as emo-expressions either.

16. **Q: A sentiment is always related to some target, does the same apply to emo-expressions?**

A: No. While it is surely true that a sentiment always requires a target according to our definition, emotional expressions can be any words or phrases that have a positive or negative connotation. So, for example, words like *Erfolg* (*success*), *Missgeschick* (*misfortune*), *verärger* (*upset*) all can be considered as emo-expressions if you think that there are positive or negative feelings associated with them.

17. **Q: What should we do when target and its emo-expression are represented by one nominal compound, e.g. *Ausländerhass* (*hatred of foreigners*)?**

In such cases, the compound words should be split and you should annotate their constituents separately. But, unfortunately, it is difficult in MMAX to split words, but still possible. Because you and other annotators will use the same word splitting files and your markables will also be bound to that specific splittings, the word-splitting operation should be performed simultaneously for all annotators. Therefore, it would be best practice if you would send an e-mail to the author of these guidelines with the name of the file, the tweet, and the word which in your opinion is a compound which requires splitting. I then would centrally split these words and ask all other annotators to get a new version of the files.

18. **Q: How should I annotate chains of intensifiers/diminishers – each separately or the whole chain with one tag?**

A: Each element should be tagged separately, e.g. *Du bist die [aller]_{intensifier} [aller]_{intensifier} 'Beste!* (*You are the [very]_{intensifier} [very]_{intensifier} best!*)

EXAMPLE 4.19. @DenisQuadt : Wer meint , die #piraten hätten noch nichts umgesetzt, möge das bitte lesen :)

(@DenisQuadt : Wer meint , die #piraten hätten noch nichts umgesetzt, möge das bitte lesen :)

5 Examples

In this section, we provide a couple of real-world examples of complete annotations with explanations of our decisions. In cases when we do not specify an attribute for a markable, it is assumed that this attribute takes its default value.

EXAMPLE 5.1. WAS HABEN ALLE MIT [IHREN [VERF*CKTEN]_{emo-expression:polarity=negative,intensity=strong,sarcasm=false} [GRÜNEN AUGEN]_{target}]_{sentiment:polarity=negative,intensity=strong,sarcasm=false} (WHAT DO THEY ALL HAVE WITH [THEIR [F*CKED]_{emo-expression:polarity=negative,intensity=strong,sarcasm=false} [GREEN EYES]_{target}]_{sentiment:polarity=negative,intensity=strong,sarcasm=false})

Explanation: In this case, we have only an evaluative opinion about the eyes. It is not clear whether the author has any emotions towards the people with green eyes, only that she finds these eyes *verf*ckt* (*f*cked*). *verf*ckt* is an intense abusive word with clear negative meaning, so we set the polarity of this emo-expression and the sentiment it pertains to to **negative** and set the intensity of both elements to **strong**.

EXAMPLE 5.2. [Wo ist der [#Jubel]_{emo-expression:polarity=positive,intensity=strong,sarcasm=true} von [#CDU]_{target} [#CSU]_{target} & [#FDP]_{target} über den Tod der Mieterin nach #Zwangsräumung?]_{sentiment:polarity=negative,intensity=medium,sarcasm=true} ([Where is the [#exultation]_{emo-expression:polarity=positive,intensity=strong,sarcasm=true} of [#CDU]_{target} [#CSU]_{target} & [#FDP]_{target} about the death of the renter after forced #eviction?]_{sentiment:polarity=negative,intensity=medium,sarcasm=true})

Explanation: First of all, we do not mark *Jubel von #CDU, ... über den Tod von ...* (*the exultation of the #CDU ... about the death of ...*) as sentiment, because the existence of this feeling is raised to question (cf. explanation for Question 1 in Section 4). On the other hand, the mere hypothesis about the presence of such glee feeling from the side of a political party, which should presumably care about its potential voters, is already a sarcasm. The **emo-expression** which shows us that it is a sarcasm is the word *#Jubel* (*#exultation*). The primary sense of this word has a positive polarity; the strength of the expressed emotion is higher than that of the word *Freude* (*joy*) and we therefore set the value of the **intensity** attribute to **high**. Moreover, this word by itself is already meant sarcastically in the given context, so, we accordingly set the value of its **sarcasm** attribute to **true**. But for all that the word *Jubel* (*exultation*) has a strong intensity, the overall way of expressing sentiment is rather subtle and does not show high exaggeration of the author. So, the total intensity of that sentiment is set to **medium** rather than **high**.

A slightly more difficult case is represented by the following example, which we will process step by step:

EXAMPLE 5.3. RT @JochenFlasbarth : Guter #Spiegel-Titel , wie Welzer , Sloterdijk und andere Promi #Nichtwähler die Demokratie verspielen : Träge , frustriert
(RT @JochenFlasbarth : A good #Spiegel title , how Welzer , Sloterdijk, and other celebrity non-voters squander the democracy : Sluggish , frustrated)

Explanation: First of all, we should look for words which have an unambiguous lexical polarity, i.e. the emo-expressions, as they are our primary cues when detecting a sentiment. There is the word *guter* (*good*) with an obvious positive polarity, and there are words *verspielen* (*to squander*), *träge* (*sluggish*), *frustriert* (*frustrated*), whose polarity is unequivocally negative. Since we have two sets of emo-expressions with contradicting polarities, it is most likely that we also have two types of sentiments – one with a positive evaluation and one with a negative. The positive evaluation is made about the suggested #Spiegel title *wie Welzer , Sloterdijk und andere Promi #Nichtwähler die Demokratie verspielen: Träge , frustriert* (*how Welzer , Sloterdijk, and other celebrity non-voters squander the democracy: Sluggish , frustrated*). The author thinks this title is good and therefore has a positive attitude to the title as such. The annotation for that type of sentiment might look as follows:

EXAMPLE. [RT[@JochenFlasbarth]_{source:sentiment_ref=1} : [Guter]_{emo-expression:polarity=positive,intensity=medium,sarcasm=false, sentiment_ref=1} #Spiegel-Titel , [wie Welzer , Sloterdijk und andere Promi #Nichtwähler die Demokratie verspielen : Träge , frustriert]_{target:sentiment_ref=1}]_{sentiment:polarity=positive,intensity=medium,sarcasm=false,id=1}
([RT[@JochenFlasbarth]_{source:sentiment_ref=1} : A [good]_{emo-expression:polarity=positive,intensity=medium,sarcasm=false, sentiment_ref=1} #Spiegel title , [how Welzer , Sloterdijk, and other celebrity non-voters squander the democracy : Sluggish , frustrated]_{target:sentiment_ref=1}]_{sentiment:polarity=positive,intensity=medium,sarcasm=false,id=1})

Another sentiment is instantiated by the set of emo-expressions with negative polarity. All the words *verspielen* (*to squander*), *träge* (*sluggish*), and *frustriert* (*frustrated*) seemingly relate to the celebrity non-voters, including *Welzer* and *Sloterdijk*, and author's attitude to these people is obviously negative. So, we mark this phrase as follows:

EXAMPLE. [RT[@JochenFlasbarth]_{source:sentiment_ref=2} : Guter #Spiegel-Titel , wie [Welzer]_{target:sentiment_ref=2} , [Sloterdijk]_{target:sentiment_ref=2} und [andere Promi #Nichtwähler]_{target:sentiment_ref=2} die Demokratie [verspielen]_{emo-expression:polarity=negative,intensity=medium,sarcasm=false,sentiment_ref=2} : [Träge]_{emo-expression:polarity=negative,intensity=medium,sarcasm=false,sentiment_ref=2} , [frustriert]_{emo-expression:polarity=negative,intensity=medium,sarcasm=false,sentiment_ref=2}]_{sentiment:polarity=negative,intensity=medium,sarcasm=false,id=2}
([RT[@JochenFlasbarth]_{source:sentiment_ref=2} : A good #Spiegel title , how [Welzer]_{target:sentiment_ref=2} , [Sloterdijk]_{target:sentiment_ref=2} , and [other celebrity non-voters]_{target:sentiment_ref=2} [squander]_{emo-expression:polarity=negative,intensity=medium,sarcasm=false,sentiment_ref=2} the democracy : [Sluggishly]_{emo-expression:polarity=negative,intensity=medium,sarcasm=false, sentiment_ref=2} , [frustrated]_{emo-expression:polarity=negative,intensity=medium,sarcasm=false, sentiment_ref=2}]_{sentiment:polarity=negative,intensity=medium,sarcasm=false,id=2})

In both cases, @JochenFlasbarth is the original author of cited opinion, so we mark it as **source**. But since we have two sentiment relations, we mark this word as **source** twice and draw an edge (in our example denoted by attribute **sentiment_ref**) to the respective **sentiment** markable in each case.