ANNOTATION GUIDELINES FOR RELATIONAL EMOTION ANNOTATION IN FICTION (REMAN)

1. Project Background and Description

REMAN (Relational Emotion Annotation in fiction) is a project aimed at semantic annotation of emotions and relations between entities in fiction texts. The project is funded by Center for Reflected Text Analytics (https://www.creta.uni-stuttgart.de/), University of Stuttgart.

The goal of the project is to create a corpus of the text excerpts from the literary fiction that are annotated for the directed emotions.

REMAN annotation will include:

1. Annotation of entities occurring in given texts
2. Annotation of emotions that describe relationships between these entities
3. Annotation of semantic relationship between those entities

A simple example of such an annotation is shown below and is referred to as *Running example 1*. Note: the shown examples are the fully annotated sample texts. We will guide you through the entire annotation process later in the document.

Running example 1
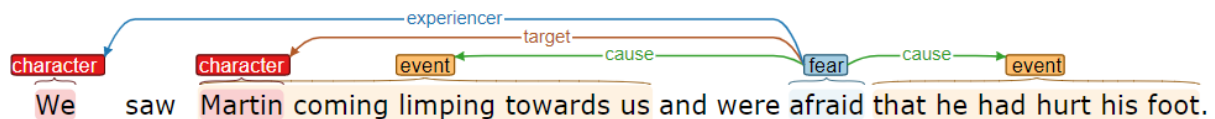
The rest of the document is structured as follows. Sections 2 gives detailed information about the concepts of the project with their definitions. Section 3 outlines the general annotation procedure, and Section 4 is a quick guide to the annotation process.

2. Basic Concepts and Definitions

The annotation starts with the annotation of phrases followed by the annotation of relations.

2.1 Phrase Annotation

Each annotation unit includes 3 consecutive sentences.

The second sentence in the annotation unit is the **target** for the annotation with respect to emotion and entities, and sentences number 1 and 3 are given for context. That means that *target sentence requires complete annotation, while context sentences require only selective*

annotation of information relevant for target sentence (e.g., entities related to the emotion in the second sentence).

The labels appearing above the annotated text are called tags.

2.1.1 Emotion

Definition In this project, we conceptualize an emotion as one's experience that falls under one or several categories from the following list:

1. **Joy** *(e.g.,* glad, happy, *etc.)*
2. **Sadness** *(e.g.,* sad, broken-hearted, *etc.)*
3. **Anger** *(e.g.,* furious, angry with, *etc.)*
4. **Fear** *(e.g.,* afraid of, frightened, *etc.)*
5. **Surprise** *(e.g.,* astonished, surprised, *etc.)*
6. **Anticipation** *(e.g.,* looking forward, expecting, *etc.)*
7. **Trust** *(e.g.,* trust)
8. **Disgust** *(e.g.,* hate, abominate, cannot stand, *etc.)*
9. **Other** (any emotion not listed above or of unclear nature)

Emotions from 1 to 8 are based on Plutchik's basic emotion classification [1], emotion 9 is our addition needed to account for the emotions that are not classified as basic in Plutchik's model.

> An emotion is a central concept of the REMAN project. It serves a foundation for the other related concepts and, if present, should always be annotated <u>in the target sentence</u> of the annotation unit.

If the emotion is not clear and does not fall into one of the 8 specific emotions, use "other-emotion" tag. This also applies to such emotions as *boredom, interest, distraction.*

Emotions can be expressed using various words and expressions. Typical linguistic realizations of emotion expressions are adjectival phrases (e.g., "John is happy"), verb phrases (e.g., "John fears Bill"), noun phrases (e.g., "My fear is grounded on…"), prepositional phrases ("…through the mistake I was greatly saddened"), *etc.*

Emotion expressions should be annotated as short as possible. In most cases, regardless of the type of the phrase, you will select a single token (e.g., *fear, angry, happy*). However, in some cases it is not possible to select just one token, especially in the case of implicit emotions. For example, "the corners of her mouth went down" is a description of sadness

---

[1] *Plutchik, R. (2001). The Nature of Emotions Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice.* American scientist, 89*(4), 344-350.*

and there is no single word that could be selected to mark it. Therefore, the whole phrase should be selected for the annotation.

The emotions can be modified in three separate ways: they can be intensified (strong), weakened (weak), or negated. While the first two modifier types define the degree of how intense an emotion is, the third type applies to situations when an emotion is used in a negated context. The presence of modifiers in vicinity to an emotion word should in most cases be the only clue to whether the emotion is strong, weak, or negated. In some cases, it is possible to deduce the emotion intensity from the context. However, the annotators should try to mark the intensity solely based on the presence of modifiers before or after the emotion word or if the suffixes serve as modifiers. The intensity markers are not annotated and **the modifier can only be applied to the emotion word.**

2.1.2 Entity

Definition An entity is something that has clear identity of an object, concept, state, or event. Entities do not necessarily have material manifestation. Entities are only annotated if their constituent parts are targets, causes, or experiencers (see Section 2.2) of an emotion.

Entities are limited to the following types:

- Character: any entity that is conceptualized by the author as a character. For example, *he, I, mister, John, we, people, dog,* etc.
  Characters can only be noun phrases (e.g., "the sister of Apollo") or adjectival phrases (e.g., "the pretty girl") that may also contain prepositional phrases (e.g., "the man with an umbrella"). Characters should be annotated as short as possible but without losing characteristic information. Consider the example below.
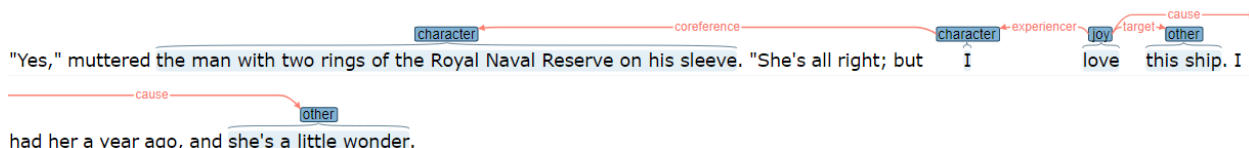
  Annotating the character in the first sentence as only "the man" would result in an incomplete annotation, because the document may mention some other man who has three rings of the Royal Naval Reserve, not two.

  Characters are only annotated if they are targets, experiencers, or causes of an emotion (see Section 2.2).

- Event: The occasion or happening that has an impact on the emotion experienced by the Experiencer. Events can be expressed in many ways. Events are normally limited to the following categories (examples are given after a colon):
  a. **Verbal phrases**: *...she didn't say a single word*
  b. **Adverbial phrases**: *...the children were playing loudly*

c. **Prepositional phrases**: ...*he screwed it up*
d. **Clauses**:
    i. *... John said [that he steals money from his mother]*
    ii. *The man [whose car I hit] hates me*
e. **Tense phrases**: *She received a letter with bad news...*
f. **Noun phrases**: *a wedding, a birthday party, the war, etc.*

If possible, events should ideally span the entire phrase and in case of clauses should include the complementizers or determiners ("that", "which", "the", etc.).

Note that events of type *f* are noun phrases that are semantically events, i.e. refer to things that we know as **events**.

- Other: umbrella concept for everything else that is neither a character, nor event. The Other category is limited to the following types of phrases (examples are given after a colon):
    a. **Adjectival phrases**: ...*pretty face*
    b. **Noun phrases**: ...*this house*
    c. **Tense phrases or sentences**: *She was the only treasure on the face of the Earth that my heart coveted.*

If possible, events should ideally span the entire phrase and in case of noun or adjectival phrases should include the determiners ("the", "this", etc.).

Note: Events and Other are only annotated if they are targets or causes of an emotion (see Section 2.2)

Important! *As far as both Event and Other annotations can take tense phrases, i.e. have verbs in their content, it is necessary to distinguish between what is an event and what is not. The main difference between Event and Other is that Event relates to something that happens/happened/will happen. In this sense, Event is a <u>dynamic</u> concept. At the same time, Other is more related to states, something more long-lasting than an event. In this sense, Other is a <u>static</u> concept.*

2.2 Relation Annotation

The relationship between Entities and Emotions is defined in terms of Cause-Experiencer-Target relations. The relationship can be established between an Emotion and an affected Entity.

Additionally, there is a Coreference relation that is established between a coreference expression and an entity to which it refers.
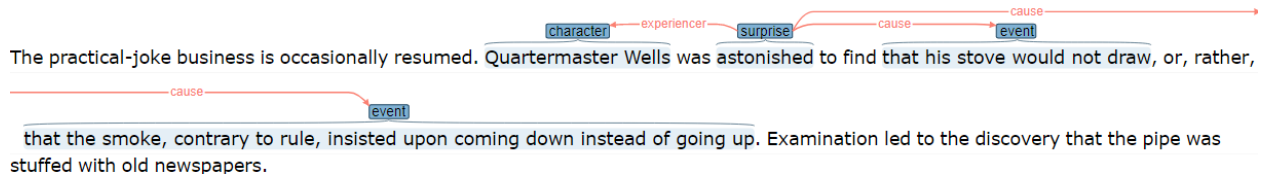
The relationship property has 4 values: Experiencer, Cause, Target, and Coreference:

1. **Experiencer**: <u>Only entity of type Character</u> that experiences the expressed emotion
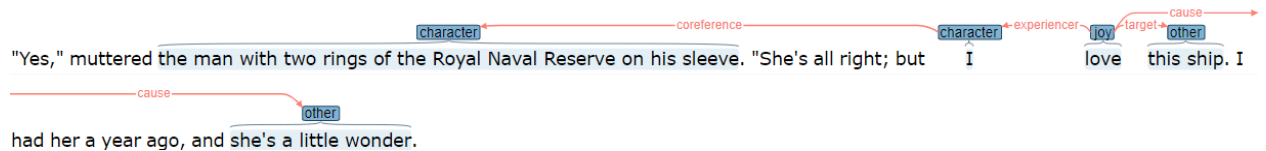
2. **Cause**: Entity of any type, which serves as a stimulus, something that evokes the emotional response in the Experiencer.

    If there are multiple causes for the emotion, or if the document contains explanation or additional information for why certain stimulus evokes the emotion, then these parts of texts should also be included in the annotation. See three examples below.
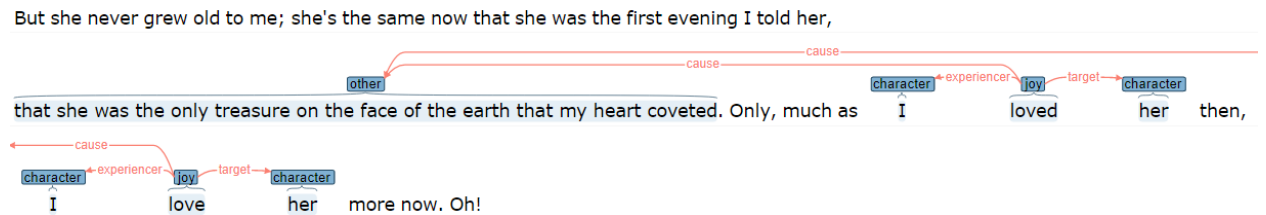
    Example 1:

    The practical-joke business is occasionally resumed. Quartermaster Wells was astonished to find that his stove would not draw, or, rather, that the smoke, contrary to rule, insisted upon coming down instead of going up. Examination led to the discovery that the pipe was stuffed with old newspapers.

    Example 2:

    "Yes," muttered the man with two rings of the Royal Naval Reserve on his sleeve. "She's all right; but I love this ship. I had her a year ago, and she's a little wonder.

    Example 3:

    But she never grew old to me; she's the same now that she was the first evening I told her, that she was the only treasure on the face of the earth that my heart coveted. Only, much as I loved her then, I love her more now. Oh!

3. **Target**: Entity of any type towards which the emotion experienced by the Experiencer is directed (e.g. "I hate you" – *hate* is an emotion, *I* is an experiencer, *you* is a target).

4. **Coreference**: **Coreference is only resolved, if the expression that refers to some referent, is of type Character AND is an experiencer, target or cause of an emotion (see Section 2.2).** If the coreference can only be resolved by crossing the sentence boundaries (e.g., the target sentence contains a coreference whose referent is found in any other sentence), then the coreference arc should cross the sentence boundaries.

    Note: the coreference should only be resolved if the referent has a higher level of specificity than the coreference (i.e., it is a proper noun or a noun denoting a group or class of objects). **Do not link two pronouns with coreference arcs (e.g., he →**

**him). Do not link two proper names with coreference arcs (e.g., Sarah →
Sarah).**

In other words, if the experiencer, target, or cause is a pronoun (personal or
possessive) and there is a name to which the pronoun is linked in the text, resolve
the coreference. Personal names (e.g., Sarah) have higher priority than class of
objects (e.g., "princess"): If there is both "Sarah" and "princess" in the text, resolve
coreference to Sarah. If there is only "princess", resolve coreference to "princess".

Recommended Annotation Procedure

It is recommended that you do the annotations in two major steps as shown below:

1) **Phrase annotation**
   a) Annotate emotions (if any)
      i) annotate intensity if applicable
   b) Annotate entities (if any)
      i) Choose the character that is the closest to the emotion word in terms of tokens
         (to the left or to the right). If the closest character is expressed via pronoun,
         select it and resolve coreference if possible. This rule <u>has no exceptions</u>: Select
         the closest one, even if it is a possessive pronoun, but there is a name just few
         tokens further than a pronoun.
   c) Annotate coreference (if any)
2) **Relation annotation**
   a) Annotate relations
      i) choose between cause-experiencer-target-(coreference)

You will learn how to perform these steps in the following section.

3. Annotation Tutorial

In this tutorial, you will be introduced to the annotation environment and learn how to follow the annotation procedure presented in the Section 3.

> To access the annotation environment, you must be logged in on a computer inside the institute or establish a VPN connection.

4.1 Access the annotation environment

If the requirement specified above is satisfied, do the following:

1. In your browser, open the following link [blinded for the review]
2. Login with the credentials you received from the administrator of the project. Click *Annotation*. The following screen appears:
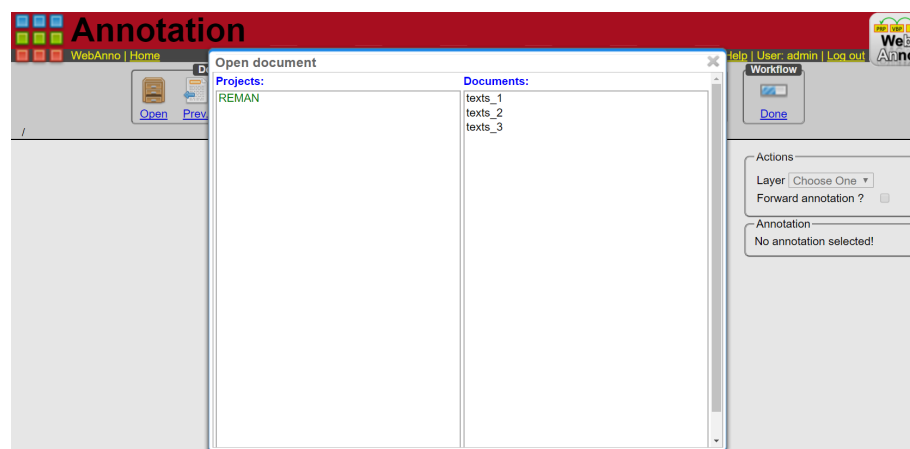

Figure 1: Selecting the text for the annotation

3. Select REMANde from Projects, then in the Documents section double-click the first document. You will enter the annotation mode that looks as shown below. Immediately, you see the first annotation unit.

Figure 2: Annotation environment

We will now annotate this very first example.

As recommended in Section 3, we will perform the annotations in two major steps, from the phrase annotation to the relation annotation. Phrase annotation is done in the EmotionSpan layer[2], while relation annotation is done in the EmotionRelation layer.

3.1 Quick guide to the annotation process

1. There are two panels on the right side of the screen (Figure 3). In the drop-down menu "Layer" under Actions (Figure 3), you see the EmotionSpan layer. It is the only visible layer with which you will work. Disregard Forward annotation tick box everywhere in this project.
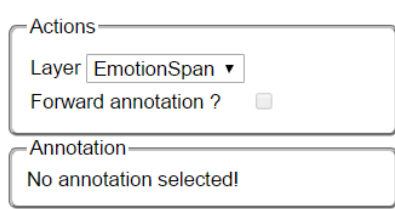

Figure 3: Actions and Annotation panels

2. Annotate emotions. There is only one emotion expression found in the target sentence (sentence 2: "angry"). Select token "angry" with a mouse or double-click it.
3. As soon as you do it, the Annotation panel located below Actions panel will display the properties of the EmotionSpan layer, as shown on Figure 4 below.
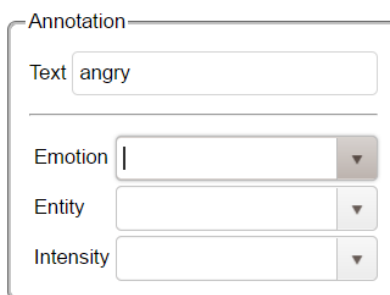

Figure 4: Activated annotation panel

4. Click on the Emotion drop-down menu and select *anger* from the list of available emotion tags.
5. Leave the fields Entity and Intensity blank.
6. Proceed to the next step.

8

[2] *Layer is an annotation scheme that constitutes the backbone of the annotation. You can think of it as of specific kind of annotation. For example, EmotionSpan layer is for annotating emotions, entities, and events, while EmotionRelation layer is for annotating relations between these emotions, entities, and events.*

7. Annotate entities. Double-click the token "Madam". The annotation panel gets activated (Figure 4). Click Entity drop-down menu and select *character*. Repeat this procedure for the tokens "us all". **As far as "us all" is a direct target of the emotion and refers to "the pale girl" and "the pretty one" in the first sentence, these two phrases should also be tagged as "character" such that a Coreference relation can be established later**.

8. With mouse, select tokens "if you open that door". The annotation panel gets activated (Figure 4). Click Entity drop-down menu and select *event*.

9. At this step, your annotations should look as follows:

The pale girl started to do so, but the pretty one held her back. "No, Nellie; Madam will be angry with us all if you open that door."

10. Annotate coreference relation. Click on "us all" and draw an arc to the "the pale girl". The annotation panel gets activated as shown on Fig.6. Select *coreference* from the Relation drop-down menu. Repeat the same for "the pretty one".

11. Annotate emotion relations. Click on emotion tag (here, *anger*) with a left-mouse button and holding it down draw an arc to a *character* tag above "Madam"[3].

```
┌─Annotation──────────────────┐
│                             │
│  Text  [angry] - [Madam]    │
│  ─────────────────────────  │
│                             │
│  Relation  [          ▼]    │
│                             │
└─────────────────────────────┘
```

Figure 5: Annotation panel for EmotionRelation layer

12. Click on Relation drop-down menu and select *experiencer* tag.

13. Draw an arc from the emotion tag to the *character* tag above "us all". Select *target* from Relation drop-down menu.

14. Draw an arc from the emotion tag to the *event* tag above "if you open that door". Select *cause* from Relation drop-down menu.

15. At this step, your annotations should look as follows:

The pale girl started to do so, but the pretty one held her back. "No, Nellie; Madam will be angry with us all if you open that door."

[3] *Remember this simple rule: Arcs denoting relationships always originate from the emotion expression.*

16. Congratulations! You annotated the first annotation unit. Proceed with the annotation by clicking the Next button on the project panel Page in the upper center of the screen.

> Do not click Finish button unless you finished annotation in the entire document. Clicking the button prematurely will lock the project and you will need to contact the administrator to unlock it. Just click Next on the Page panel, if you completed the annotation for the text on the screen.

3.2 Other considerations, rules, and exceptions

3.2.1 **Other emotions** Use "other-emotion" emotion tag **only if you doubt** the nature of the expressed emotion.

3.2.2 **Character entities** It is not required that all characters appearing in the annotation unit are tagged as *character*. *Only the characters involved in the emotional expressions are subject to annotation*.

3.2.3 **Annotation spans** You may be challenged to decide the actual annotation spans (i.e., how many tokens should an annotation include). Therefore, follow this simple rule when doing the annotations:

> Everything should be annotated as short as possible but not shorter

That means that the shorter annotations (spanning maximum few tokens) are generally preferred but such brevity should not harm the completeness of relevant information. Make sure that the phrases or words that you annotate are **valid English phrases**. For example, the word *box* is not a valid way of using this word outside of dictionary, it should be either *a box* or *the box*. So, it makes sense to include determiners in the annotation if they present. The same applies to German.

3.2.4 **Nothing to annotate** It can happen that there is nothing to annotate in the annotation unit with respect to emotion. If this is the case, you may skip the annotation unit.

A simple rule that may help you to decide whether to annotate the parts of the text or not can be formulated as follows:

If the <u>text can be rephrased</u> as:

`CHARACTER X feels EMOTION Y (due to SOMETHING or SOMEBODY)`

then it should never be omitted from the annotation. Note: Parentheses indicate optional condition.

3.2.5 Relationship(s) crossing sentence boundaries

Sometimes the cause, the experiencer, and the target of the emotion are not found in the target sentence (2nd sentence). If this is the case, and the experiencer, target, or cause are mentioned in the context sentences, the relationship arcs originating from an emotion expression must cross the sentence boundaries.

3.2.6 Multiple experiencers

It may often happen that there is more than one experiencer of the same emotion of multiple emotions in the same text. In such cases, all experiencers should be linked to relevant emotions.

3.2.7 Multiple mentions of the same experiencer

It often happens that the same person, who is an experiencer of an emotion, is mentioned several times in the text (e.g. "I went… I felt… <emotion word> … I …"). In this case, annotate only one mention of the character selecting the one that is closer (in terms of tokens) to the emotion phrase.

3.2.8 Negated emotions

Only emotions words with a negation modifier (*not, nor, neither, don't, etc.*) or with a clear context can be negated. You should never annotate one emotion as being an inverse of the opposite emotion. For example, *sadness* should never be annotated as *negated joy* and vice versa.

3.2.9 Multiple emotions, same experiencer

If there are multiple emotions in the text that has the same experiencer, target, or cause, do the following: Mark emotion keyword, select the character that is closest to this emotion word (see Recommended Annotation Procedure above). If coreference can be resolved, do it. Repeat for every emotion keyword. In case of multiple emotions this can result in multiple coreference arcs, it's okay.