Task Description for Source/Target Belief and Sentiment Evaluation (BeSt) at TAC 2016

V.2.4

0. Changes to This Document

- This revision is oriented towards the general public.
- The notion of “provenance” is operationalized.
- A detailed description of the evaluation metric is added.
- V.2.3: Added clarification on “none” sentiment polarity and on missing sources; removed “Proposed” from title of this document.
- V.2.4: Added clarification about micro- vs. macro-averaging (we use micro-averaging for the evaluation).

1. Introduction

This document summarizes the plans for the 2016 BeSt Evaluation, which is an evaluation of sentiment and belief detection with source and target, where sources are named entities and targets are named entities or events or relations.

The evaluation has the following characteristics:

- It is interested in sources, attitudes, and targets: who has what mental attitude towards what?

- The evaluation is not interested in trigger words or linguistic markers of the detected attitude, only in detection of the attitude itself.
- The evaluation includes belief and sentiment.
- The source is an entity of type Person, Geo-Political Entity (GPE), or Organization. The target can be any relation, or any event. In addition, for sentiment only, the target can also be any entity.
- There are two conditions for evaluation:
    - We provide gold entities, relations, and events (EREs).
    - We provide predicted EREs. The predictions will come from a combination of systems.
- For both conditions, participants will have access to files specifying EREs of interest; this includes in-document co-reference of entity mentions and event mentions. The tasks of finding entities, relations, and/or events, and related tasks such as co-reference, are **not** part of this evaluation.
- The evaluation will be on English texts, but will also contain smaller Chinese and Spanish tracks.

2. Conceptual Description of Task

This section provides a **conceptual** description of the task. The actual implementation of the task in terms of input and output files (including file formats) is detailed below.

The following questions illustrate what the evaluation is getting at. We use the term “private state” to refer to either belief or sentiment.

1. Does JohnFromTulsa like Obama?
2. Who has (or is claimed to have) negative sentiment towards Obama?
3. Who is self-reporting a belief about Obama, and what is it?
4. What private states does BigGuyAtlanta express (or do others report he expresses) about the the annexation of Crimea?
5. What private states of others is BigGuyAtlanta reporting?
6. What is Hillary Clinton’s sentiment towards the Benghazi hearings?
7. Does BigGuyAtlanta have a belief about Obama?
8. Does BigGuyAtlanta believe that Obama was born in Kenya?

The systems we will evaluate and determine the sentiment and/or belief from a holder (source) towards a target, which is an entity, a relation, or an event.

The basis of the evaluation are private state tuples (PSTs), which are 4-tuples of the following form:

(**source-entity**, **target-object**, **value**, **provenance-list**)

The meaning of the components of the PST are as follows.

1. The 4-tuples express the belief or sentiment of the **source-entity** towards the **target-object** (which can be an entity, a relation, or an event).
2. The **value** is:
    a. A sentiment value (*positive*, *negative*), or
    b. A belief value (*CB*, *NCB*, *ROB*) where:
        i. *CB* = committed belief, meaning that the source is convinced the target is true. Note that this does not mean it "happened" in the past, a source can hold a committed belief about an event in the future.
        ii. *NCB* = non-committed belief, meaning that the source thinks it is possible or probable that the target is true, but is not certain.
        iii. *ROB* = reported belief. Sometimes, a writer reports on a different source's belief, without letting the reader know what his or her belief state is.

3. The **provenance-list** is a list of pointers to the text passages which support the identified claim about belief or sentiment. The **provenance-list** contains an entry for every single piece of textual evidence that supports the specific private state claim expressed by the PST. We consider an instance of provenance to be the target mention ID (as defined in the rich-ere.xml file, see below), along with the file name.

All the private states expressed in a document collection can be expressed as a collection of PSTs. The same (**source-entity, target-object**) pair can occur several times with different values. There are two reasons for this:

1. A source can have several different private states with respect to the same target. For example, the writer can have positive sentiment towards the election of Clinton, and also have a non-committed belief towards it. A source can even have conflicting private states, for example both positive and negative sentiment. This happens when someone changes his or her mind, or when they react to different aspects of the target. In this evaluation, all private states should be found; there is no aggregation or temporal analysis of conflicting private states.
2. Because the provided ERE files only record in-document coreference, it is possible that what is in fact the same source and target and the same private state get recorded multiple times (if they are expressed in multiple documents).

The task for the evaluation is as follows:

Input: a source text file and an ere.xml file which lists entity mentions, relation mentions, and event mentions, as well as intra-document coreference among them.

Output: a best.xml file which refers to the input ere.xml file and which lists the belief and sentiment relations from entity mentions to entity mentions, relation mentions, and event mentions. All mentions will be mentions introduced in the ere.xml file.

The file formats are discussed in Sections 3 and 4, which also explains how these file formats correspond to the conceptual presentation of the task that this section is about.

Note that performers participating in the evaluation do not do entity, relation, event recognition, or coreference resolution. These tasks will already have been performed.

For the numerical evaluation results, the parameters of the evaluation are as follows:

1. There are two annotation conditions:
    a. The entity mentions, relation mentions, and event mentions in the ere.xml file are gold annotations.
    b. The entity mentions, relation mentions, and event mentions in the ere.xml file are the output of an automatic system, as described in Section 5.
2. There are two evaluation conditions:
    a. The **provenance-list** is scored fully in the evaluation (all provenance instances need to be found).
    b. The **provenance-list** is scored as one-is-enough (a single instance of a correct provenance counts as full score).

This gives us 4 evaluation settings. We use recall, precision, and F-measure as measures.

Training data of both types (gold and predicted) will be available, so that performers can choose to have two systems optimized for the two annotation conditions.

3. The Input Files

There are two input files:

- The text file.
- The ERE annotation file, gold or predicted; either way, with *.rich_ere.xml name.

We discuss the formats in some detail.

3.1. Text File Format for Source Files

The training data is largely discussion forums, but will also include a bit of newswire.

The forum source files in the training data are not xml and there is no dtd, but there are some xml-like elements to mark the dialog structure.

- **Post** is a single post by an author (with possibly multiple sentences). It has attributes author, datetime, and id. Note that the author in the post open tag is typically annotated as an entity span! Posts are not nested.
- **Quote** is element corresponding to a use of the quote facility provided by the discussion forum, which results in a visually separated box in the original forum. Quotes have their own author. Quotes appear within posts, and can be nested in other posts.
  IMPORTANT: all material in a quote element is excluded from the evaluation!

The newswire source data will be in xml format, with the simple addition of a doc tag. The package will include a dtd file.

All source files in the evaluation data will be released as a well-formed xml file.

In both the training and evaluation data, all character offsets are counted in the same way: including all tags in the source file.

3.2. File Format for Entities, Relations, Events: RICH_ERE.XML

This is a well defined xml file. The dtd can be found here:
http://volta.ccls.columbia.edu/~rambow/best-eval-2016/ .

The ERE file contains list of all mentions of annotated entities, relations, and events. While all person, GPE, and organization mentions are annotated, not all relations or events are annotated, only those of certain types. Please refer to the ERE annotation manual for details (the guidelines are in the package that participants receive from the LDC).

In general, we distinguish between objects and object mentions. For example, the person Barack Obama is an entity, and he may be mentioned several times in a text. The mentions are textual occurrences and are represented using a pointer to a specific text file and character offsets in that file. The entity himself is not a textual occurrence and does not have an offset. Instead, Obama is an actual person.

One detail to watch out for: while we have events and event mentions, and relations and relation mentions, the terminology is different for events: we have “hoppers” and event mentions. A hopper is a (metaphorical) container in which several related events are grouped; this is done

because the notion of `event' is more complex than that of an entity, with less of a clear single referent in the real world, and thus with less clarity about which event mentions actually refer to the same event. For the sake of this evaluation, you can think of the term "hopper" as actually meaning "event", as the ontological malaise caused by the problem of event coreference is not directly relevant to the belief and sentiment task. We will, however, continue to use the "hopper" terminology.

A detailed description of the file format for rich_ere.xml files can be found below in Appendix A.

4. Specific Output Format: best.xml format

4.1. File Format for Belief and Sentiment Annotation: BEST.XML

The required output for belief and sentiment prediction is the same as that in which the LDC is providing source-and-target belief and sentiment annotation (with some simplifications listed below). This is a well defined xml file. The dtd and the relax NG schema can be found here: http://volta.ccls.columbia.edu/~rambow/best-eval-2016/ . The annotation guidelines will be made available with the data.

The best.xml file contains a list of all expressed beliefs and sentiments, with source mentions and target mentions. They are organized first by type of private state (belief or sentiment), then by targets mentions, with information about the source mention added for each target mention. Note that even though the file annotates source and target *mentions*, conceptually we think of beliefs and sentiments as having source and target objects, not mentions: a person in the real world has, for example, a sentiment towards an event; see Section 2.

Note that sometimes, there is no source for sentiment and/or belief in the gold annotation, for one of two possible reasons:

- There in fact is no source mentioned ("Obama is not liked").
- The source is the author but the author is unnamed. This is the case in all of the newswire files, and none of the discussion forums. This affects mainly belief, as the author of the newswire is in fact expressing many beliefs.

At this point we cannot distinguish these two cases.

A detailed description of the file format for best.xml files can be found below in Appendix B.

4.2. Belief Representation in BEST.XML

A belief has the following attributes:

Attribute	Possible values	Meaning
Type	CB, NCB, ROB, NA	Committed belief (CB), non-committed belief (NCB), reported belief (ROB), not a belief (NA)
Polarity	pos, neg	Polarity of the proposition about which there is a belief
Sarcasm	yes, no	Was sarcasm used to express the belief?

In the 2016 BeSt evaluation, the NA value for belief (which covers cases like “I wish John would arrive tomorrow” which do not actually express a belief but, typically, a sentiment) will be taken to be equivalent with the absence of belief in the evaluation.

The polarity attribute for beliefs is very different from the polarity attribute of the sentiment: it is the polarity of the proposition which is the target of the belief. The sarcasm annotation reflects the judgment of the annotator, given the evidence in the text, as to the belief or sentiment of the source towards the target. The annotation is done on mentions in the text, such that these mentions are the ones which participate in the expression of the belief or sentiment. **The sarcasm and belief polarity attributes need not be predicted in the evaluation.**

Note that if the target of a belief is an event, then there is an additional annotation on the event arguments. Consider a case in which an author writes *Hillary may win the presidential election in 2016*. The author is *not* expressing non-committed belief about the event of winning happening (the speaker presumably is certain that the election will happen, and that someone will win it), but only about the *identity* of the winner. Thus, the winning event would be annotated as a committed belief, while the winner argument as non-committed belief. The fine-grained annotation will not be evaluated in the 2016 BeSt evaluation.

4.3. Sentiment Representation in BEST.XML

A sentiment has the following attributes:

Attribute	Possible values	Meaning
Polarity	pos, neg	Polarity of the sentiment
	none	There is no sentiment at all

Sarcasm	yes, no	Was sarcasm used to express the belief?

Sometimes, sentiment has a value of “none” rather than “pos” or “neg”. This is NOT neutral sentiment (as distinct from no sentiment), this is THE SAME THING as no sentiment. In the 2016 BeSt evaluation, “none” as the value for the polarity of a sentiment is treated like no sentiment for calculating recall and precision. Systems need to predict “none” sentiments in their best.xml files, these predictions are simply ignored.

The polarity attribute for sentiments is very different from the polarity attribute of the belief. It is an integral part of the private state (positive/negative sentiment), as opposed to the proposition which is the target of the belief. The sarcasm annotation reflects the judgment of the annotator, given the evidence in the text, as to the belief or sentiment of the source towards the target. The annotation is done on mentions in the text, such that these mentions are the ones which participate in the expression of the belief or sentiment. **The sarcasm attribute need not be predicted in the evaluation.**

4.4. Relation of BEST.XML File to Conceptual Representation of Source and Target Belief and Sentiment

In Section 2 above, we introduced private state tuples (PSTs), which are 4-tuples of the following form:

- (source-entity,target-object,value,provenance-list)

We explained the task conceptually using these tuples. How do these tuples relate to the best.xml file format? The main issue is that the best.xml files do not list entities, relations, or events, but rather entity mentions, relation mentions, and event mentions. We discuss each element of this 4-tuple in turn to show how it can be derived from the best.xml file.

1. Source entity: this can be obtained from the rich-ere.xml file, which maps each mention to its corresponding object. If there is no source (in the predicted or gold files), a special token “NONE” is inserted into the 4-tuple, and treated like any other source.
2. Target object: this can be obtained from the rich-ere.xml file, which maps each mention to its corresponding object.
3. Value: this is explicitly listed in best.xml file.
4. Provenance list: for the sake of this evaluation, the list of all target mentions will count as the provenance list. This is exactly what the best.xml file provides.

Recall that a polarity of “none” for a sentiment means that there is no sentiment; therefore, if a sentiment annotation in a best.xml file has polarity “none”, no 4-tuple is created at all.

This mapping is a conceptual mapping and need not be performed for the evaluation. The required submission is simply a best.xml file.

5. Data

The corpus will be the newswire and discussion forum data. The data for the evaluation should have gold Belief, Sentiment, and Rich ERE annotation.

Details for the three languages. Counts are in words.

ENGLISH

- 1st increment is available
- Total planned training (includes dev): up to 350k words
- Test: 200 documents (about 100k-150k words; hidden until evaluation)

CHINESE

- 1st increment is available
- Total planned training (includes dev): up to 250k words
- Test: 200 documents (about 100k-150k words; hidden until evaluation)

SPANISH:

- Total planned training (includes dev): up to 125k words
- Test: 200 documents (about 100k-150k words; hidden until evaluation)

The remaining training data will be released incrementally in April and May.

6. Scoring

The scoring is based on the 4-tuples. We perform a recall-precision analysis on the predicted 4-tuples against the gold 4-tuples described in Section 2 above. However, as the 4-tuples contain lots of information, we assign partial credit. When assigning partial credit, we always require that the target is correct.

1. Partial credit is given if the target is correct, but not the source.

2. Partial credit is given if the type of attitude is correct (i.e., belief or sentiment), but not the value (pos or neg for sentiment, CB, NCB, ROB for belief). No partial credit is given if belief is predicted when there is a sentiment and vice versa.
3. Partial credit is given for the provenance list (i.e., pointers to documents and specific text passages that support the claimed attitude from source to target). There are two conditions. In the full-provenance condition, partial credit is given based on recall-precision analysis of the provenance list. In the single-provenance condition, full credit is given if at least one correct provenance is detected.

Here is a detailed description.

1. Given a predicted best.xml and a gold best.xml file, both are first converted to the 4-tuple notation of Section 2. This happens as follows:

    a. For each (source-mention, target-mention) pair in the best.xml file, the corresponding source and target objects are retrieved from the rich-ere.xml file.
    b. The value (a belief or sentiment value) is retrieved from the best.xml file.
    c. If there is no 4-tuple with the source, target, and value in the first three positions, a new 4-tuple is created. The provenance list of this new tuple is set to be a list containing the target-mention .
    d. If there already is a 4-tuple with the source, target, and value in the first three positions, the target-mention is added to the provenance list.

    This gives us two sets of 4-tuples that express the same content as the gold and predicted best.xml files (in light of the shared rich-ere.xml file), respectively.

2. We then perform an initial analysis on the first three fields of the predicted 4-tuples against the gold 4-tuples. We sort all predicted 4-tuples by the type of match against the gold 4-tuples. We then process all tuples of this match type before moving to tuples of the next match type. Whenever a gold tuple is part of a successful match, it is removed from the pool of possible matches for subsequent predicted tuples.

    The match types are as follows. They are processed in the order given.

    a. If a predicted 4-tuple does not match any gold tuple on target and attitude type (belief or sentiment), it is a false positive.
    b. If the source, target, and value of a candidate tuple match a gold 4-tuple, then the tuple counts as a true positive with a true positive matching score of 1.
    c. If the source and target and attitude type match a gold 4-tuple, then the tuple counts as a true positive with a true positive matching score of ⅔.
    d. If the value and target of a candidate tuple match a gold 4-tuple, then the tuple counts as a true positive with a true positive matching score of ⅔.
    e. If only the target and attitude type of a candidate tuple match a gold 4-tuple, then the tuple counts as a true positive with a true positive matching score of ⅓.

f. Any gold 4-tuple that is not matched at least partially by a predicted 4-tuple under rules (b), (c), (d), or (e) counts as a false negative.

3. If a predicted tuple counts as a true positive under (2), we check the provenance list. Recall that an instance of provenance is the target mention, so two instances of a provenance match if they are the same target mention. For the provenance list, there are two conditions:

    a. In the **full-provenance condition**, a recall-precision matching of the predicted provenance list against the gold provenance list is performed. The resulting f-measure is used to scale the true positive matching score obtained in step (2). Note that this can be 0, if no correct instances of the provenance are identified.
    b. In the **single-provenance condition**, we check if *any* predicted provenance is correct; if yes, the tuple remains a true positive and the matching score from (2) is retained; if no, the tuple is counted as a false positive.

The resulting sums of true positives matching scores, the count of false positives, and the count of false negatives are used in a standard recall-precision calculation, with the f-measure as the final result.

The following information which is found in the gold best.xml files will not be evaluated against:

- Polarity for belief
- Sarcasm
- Belief towards event arguments
- Instances of the *NA* value for beliefs

Validation and scoring scripts will be distributed in early June for use in development. Note that the micro-average reported by the script is relevant and will be used in the evaluation. For the macro-average, the evaluation script calculates the recall, precision, and f-measure for each file, averages the recall and precision across files, and then calculates the f-measure. The macro-averaged results are affected by some files being outliers with no data points to be found, which results in a recall of 1. In contrast, for micro-average, the calculation script merges all files into one large data set and then calculates recall and precision on this merged data set. The fact that some files have no data points does not affect the overall evaluation results.

7. External Resource Restrictions and Sharing

There is no restriction on using external linguistic resources for training the belief and sentiment modules. Specifically, external lexical resources may be used. However, it is not allowed to actually answer the queries using external resources (that may record source-and-target belief or sentiment from sources other than the test set), be it at training or at run time.

8. Websites and Contact etc.

Information on registering for the Shared Task and on signing up for the email list can be found here:

http://www.nist.gov/tac/2016/KBP/registration.html

A. Appendix: Details on RICH-ERE.XML File Format

The structure of the rich-ere.xml file is as follows. The nesting in this list reflects the nesting of the elements. We do not discuss all attributes, only those relevant to this task. Certain attributes (such as type) are explained in the ERE annotation manual.

- Element **deft_ere** has an attribute docid, the file name.
    - The **entities**, **relations**, and **hoppers** are then listed separately as elements inside of **deft_ere**.
        - Within **entities**, we find one or more elements of type **entity**.
            - An **entity** has an ID and a type.
                - Each **entity** contains elements of type **entity-mention**. These have IDs (which are distinct from the **entity** IDs), source (the filename), offset, and length. The latter three attributes together allow us to extract the actual mention as a string.
                    - For clarity, the text that is the mention is repeated as an element **mention_text**.
        - Within **relations**, we find one or more elements of type **relation**.
            - A **relation** has an ID, a type, and a subtype.
                - Each **relation** contains one or more elements of type **relation-mention**. These have IDs (which are distinct from the **relation** IDs), and a realis attribute.
                    - Relations typically have two arguments which are listed as elements **rel_arg1** and **rel_arg2**. Each of these elements has attributes entity_id (an ID of an

entity defined elsewhere in the same xml file), entity_mention_ID (an ID of a mention of the entity, also defined elsewhere in the same xml file), and a role (which is specific to the relation type). The content is the text span corresponding to the entity mention.

- The **trigger** is the text span which expresses the relation. It has attributes source (a file name), offset, and length; these three attributes allow you to find the text, which is repeated as the content of the **trigger** element. Note that a **relation** need not have a **trigger**. In such a case, there is no link between a relation and text in the file other than through the arguments.

- Within **hoppers**, we find one or more elements of type **hopper**.
  - A **hopper** has an ID.
    - Each **hopper** contains elements of type **event-mention**. These have IDs (which are distinct from the **relation** IDs), and type, subtype, and realis attributes.
      - An **event_mention** typically has one or more arguments which are listed as (repeated) elements **em_arg**. Each of these elements is either an entity or a "filler". A filler is an argument of an event which is not annotated as an entity. Each entity argument has attributes entity_id (an ID of an entity defined elsewhere in the same xml file), entity_mention_ID (an ID of a mention of the entity, also defined elsewhere in the same xml file), and a role (which is specific to the event type). Each filler argument has types filler_id (which is not an entity, relation, or event), and a role. For both entity arguments and filler arguments, content is the text span corresponding to the entity mention.
      - The **trigger** is the text span which expresses the event (often a verb). It has attributes source (a file name), offset, and length; these three attributes allow you to find the text, which is repeated as the content of the **trigger** element.

Note that there are also elements called **fillers**. These can be ignored for the sake of this evaluation.

Note that coreference is not explicitly marked; instead, two mentions corefer if they are listed inside the same object (entity, relation, hopper).

Note that the exact same string (i.e, same file, same offset, same length) can be two distinct event mentions which are mentions for two different events! (Recall that events are called “hoppers” in the ere file.) The reason is that a sentence such as *What if I killed a woman on birth control?* can refer to two events, one of type conflict-attack, the other of type life-die. So while each event mention belongs to only one event (i.e., hopper), a *string* can be two different event mentions at the very same time. (This can’t happen for entities, and probably not for relations either.)

Here is an example. Note that we extracted one entity, one relation, and one hopper from an actual file, so that the relation and event have arguments which are not mentioned below (but of course are in the original file).

```
<?xml version="1.0" encoding="UTF-8"?>
<deft_ere kit_id="55490d3300000000000000019" doc_id="0a421343005f3241376fa01e1cb3c6fb"
source_type="multi_post">
  <entities>
    <entity id="ent-8" type="PER" specificity="specific">
      <entity_mention id="m-48" noun_type="NAM"
source="0a421343005f3241376fa01e1cb3c6fb" offset="14" length="5">
        <mention_text>izzoh</mention_text>
      </entity_mention>
      <entity_mention id="m-54" noun_type="PRO"
source="0a421343005f3241376fa01e1cb3c6fb" offset="70" length="1">
        <mention_text>I</mention_text>
      </entity_mention>
    </entity>
  </entities>
  <relations>
    <relation id="r-235" type="personalsocial" subtype="family">
      <relation_mention id="relm-1062" realis="true">
        <rel_arg1 entity_id="ent-18" entity_mention_id="m-618" role="per">my
parents</rel_arg1>
        <rel_arg2 entity_id="ent-8" entity_mention_id="m-60" role="per">my</rel_arg2>
        <trigger source="0a421343005f3241376fa01e1cb3c6fb" offset="80"
length="7">parents</trigger>
      </relation_mention>
    </relation>
  </relations>
  <hoppers>
    <hopper id="h-16">
      <event_mention id="em-1168" type="transaction" subtype="transfermoney"
realis="actual" ways="voluntary">
        <trigger source="0a421343005f3241376fa01e1cb3c6fb" offset="1573"
length="4">loan</trigger>
```

```
        <em_arg entity_id="ent-8" entity_mention_id="m-264" role="recipient"
realis="true">I</em_arg>
        <em_arg filler_id="f-1550" role="time" realis="true">about a year ago</em_arg>
      </event_mention>
    </hopper>
  </hoppers>
</deft_ere>
```

B. Appendix: Details on BEST.XML File Format

The structure of the best.xml file is as follows. The nesting in this list reflects the nesting of the elements. We do not discuss all attributes, only those relevant to this task.

- Element **committed_belief_doc** has an attribute Id (unclear what this is).
    - The **belief_annotations** and **sentiment_annotations** are then listed separately as elements inside of **committed_belief_doc**.
        - Within **belief_annotations**, we find elements for **relations** and **events**.
            - Within **relations**, we find a series of **relation** elements. These are the targets of belief. They have attribute ere_id, which is the ID of a **relation-mention** which can be found in the corresponding *.rich-ere.xml file.
                - Each **relation** element contains one **beliefs** element.
                    - Each **beliefs** element contains one or more **belief** element. It has attribute type, polarity, and sarcasm (explained above).
                        - An element **source** inside a **belief** contributes the source. It has attribute ere_id, which is the ID of an **entity-mention** which can be found in the corresponding *.rich-ere.xml file. Furthermore, there are attributes offset and length, and a content string corresponding to the mention; these last three elements are determined by the ere_id, and are provided for convenience.
                - A **relation** element also contains a **trigger** element which replicates information about the relation mention from the ere.xml file. Note that not all relations have triggers.
            - Within **events**, we find a series of **event** elements. These are the targets of belief. They have attribute ere_id, which is the ID of a

relation-mention which can be found in the corresponding *.rich-ere.xml file.

- Each **event** element contains one **beliefs** element.
  - Each **beliefs** element contains one or more **belief** element. It has attribute type, polarity, and sarcasm (explained above).
  - An element **source** inside a **belief** contributes the source. It has attribute ere_id, which is the ID of an **entity-mention** which can be found in the corresponding *.rich-ere.xml file. Furthermore, there are attributes offset and length, and a content string corresponding to the mention; these last three elements are determined by the ere_id, and are provided for convenience.
- An **event** element also contains a **trigger** element which replicates information about the relation mention from the ere.xml file. Note that not all relations have triggers.
- Each **event** element also contains an **arguments** element, which can be ignored for the sake of the 2016 BeSt evaluation, as discussed above.

Note that some sources are in the same sentence as the target, while many sources are not. This happens in particular when the source is the author of the sentence who is not manifested in a first person pronoun (in a discussion forum, this is the poster or the quoted author). For example, in a sentence such as *Mary will arrive tomorrow*, there is a committed belief by the author towards the arrival event, but the author has not lexical manifestation in the sentence. In these cases, the name of the author as it appears in the standard location for the author name (i.e., in discussion forums the location in the text file for the poster name), rather than the closest mention.

Here is an example. Note that we extracted one belief (towards a relation) and one sentiment (towards a relation) from an actual file (0a421343005f3241376fa01e1cb3c6fb.best.xml), which is in fact much larger.

```
<?xml version="1.0" encoding="UTF-8"?>
<committed_belief_doc id="tree-56acee960000000000000012">
  <belief_annotations>
    <relations>
      <relation ere_id="relm-1062">
        <trigger offset="80" length="7">parents</trigger>
        <beliefs>
          <belief type="cb" polarity="pos" sarcasm="no">
            <source ere_id="m-48" offset="14" length="5">izzoh</source>
          </belief>
        </beliefs>
```

```
      </relation>
    </relations>
    <events>
      <event ere_id="em-1168">
        <trigger offset="1573" length="4">loan</trigger>
        <beliefs>
          <belief type="cb" polarity="pos" sarcasm="no">
            <source ere_id="m-48" offset="14" length="5">izzoh</source>
          </belief>
        </beliefs>
        <arguments>
          <arg ere_id="m-264" offset="1558" length="1">
            <text>I</text>
            <beliefs>
              <belief type="cb" polarity="pos" sarcasm="no">
                <source ere_id="m-48" offset="14" length="5">izzoh</source>
              </belief>
            </beliefs>
          </arg>
          <arg ere_id="f-1550" offset="1578" length="16">
            <text>about a year ago</text>
            <beliefs>
              <belief type="cb" polarity="pos" sarcasm="no">
                <source ere_id="m-48" offset="14" length="5">izzoh</source>
              </belief>
            </beliefs>
          </arg>
        </arguments>
      </event>
    </events>
  </belief_annotations>
  <sentiment_annotations>
    <entities>
     <entity ere_id="m-618" offset="77" length="10">
        <text>my parents</text>
        <sentiments>
          <sentiment polarity="pos" sarcasm="no">
            <source ere_id="m-48" offset="14" length="5">izzoh</source>
          </sentiment>
        </sentiments>
      </entity>
    </entities>
  </sentiment_annotations>
</committed_belief_doc>
```