1. General overview

Our goal is to evaluate for **Open Information Extraction (OIE)** systems. OIE aims to extract relations and their arguments from unstructured text in unsupervised manner. In its simplest form, an OIE system extracts triples (or n-ary tuples) consisting of subject, relation, and object from a given sentence. For example, from the sentence:

>*"Bell is a telecommunication company which is based in Los Angeles."*

an OIE system may extract the facts:

>*("Bell", "is", "telecommunication company")*

>*("Bell", "is based in", "Los Angeles")*

In this study, we focus on **clause-based OIE**. A **clause** is a part of a sentence that expresses some coherent piece of information. For instance, the sentence above consists of the two clauses:

>*"Bell is a telecommunication company"*

>*"Bell is based in Los Angeles"*

The extractions are generated out of individual clauses, which means that we **ignore the information in any other clauses** (details later).

2. Labeling

Each extraction is labeled with two labels (correct / incorrect):

1. **Fact label:** correctness of the extracted fact itself (together with its factuality and quantity values), and

2. **Attribution label:** correctness of the extracted attribution.

We discuss these two labels in what follows.

2.1 Fact label

The **fact label** indicates whether the whole extraction (ignoring an attribution, if any) is correct.

Rule F1: An extraction is considered **correct** if it contains all necessary information from the clause from which it has been extracted. In general, we label extractions as **correct** if they are entailed by their clause. For example, from the sentence:

>*"Bell is a telecommunication company which is based in Los Angeles."*

both of the following extractions should be labeled as **correct**:

>*("Bell", "is", "telecommunication company")*

> *("Bell", "is based in", "Los Angeles")*

In contrast, the triple:

> *("Bell", "is based in", "telecommunication company")*

is considered **incorrect**.

Rule F2: Information present in all clauses other the one from which the extraction has been taken must be ignored. This includes subordinate clauses (such as conditionals, clauses connected with "while", and so on). For example, in the sentence:

> ***"If it rains, the grass gets wet."***

all of the following extractions should be labeled as correct:

> *("grass", "gets", "wet")*
>
> *("grass", "gets wet if", "it rains")*
>
> *("it", "rains")*

Rule F3: Further information present in the clause can be ignored if this information does not change the meaning of the extraction. For example, from the sentence:

> ***"Albert Einstein was born in 1879 in Ulm."***

then both of the following extractions should be considered **correct**:

> *("Albert Einstein", "was born in 1879 in", "Ulm")*
>
> *("Albert Einstein", "was born in", "Ulm")*

If, however, the lack of a constituent of a clause is not sufficient, this should be labeled as **incorrect**:

> *("Albert Einstein", "was")*

Rule F4: In cases when relations/arguments lack words which contain crucial information (i.e. the lacking of word(s) is changing the meaning of the triple), the **factual label** should be **incorrect**.

For example, from the sentence:

> ***"Jack likes data mining."***

extractions should be labeled as follows:

> *("Jack", "likes", "data mining")* as **correct**
>
> *("Jack", "likes", "mining")* as **incorrect**

because mining and data mining have different meaning.

Rule F5: Some OIE systems output some implicit extractions, which may contain words that are not present in the input sentence. Nevertheless, they should also be labeled as **correct** if they correctly represent the information given in their clause.

For example, from the sentence:

"Mr. Mike Johnson lives in Berlin, Germany."

the following extractions should be labeled as **correct**:

("Mike Johnson", "is", "male")

("Berlin", "is in", "Germany")

and the following extraction as **incorrect**:

("Berlin", "is", "Germany")

We refer to the combination of *polarity* and *modality* as *factuality* (more on these later). For example, if the *polarity* is **NEGATIVE** and *modality* is **POSSIBILITY**, then the *factuality* is **NEGATIVE POSSIBILITY**.

Polarity distinguishes between *positive (+)* and *negative (–)* instantiations of the triples, that is, it conveys the distinction between affirmative and negative contexts.

Rule F6: The extraction should only be labeled *correct* if it expresses the same polarity (positive or negative) as its clause.

For example, suppose we have the sentence:

"John did not need the training."

then the *factual label* on both of the following extractions should be **correct**:

1. ("John", "did need", "training")   2. ("John", "did not need", "training")

Polarity: NEGATIVE   ⇔   Polarity: POSITIVE

Rule F7: If the polarity value is contained in another clause, then you should ignore the negative context in the extraction.

For example, if you have the sentence:

"It's not true that John Smith lives in Italy."

then, we have two clauses:

1) John Smith lives in Italy.

2) It's not true that John Smith lives in Italy.

Clause 2) contains clause 1). However, the negation is within clause 2) and not part of clause 1), which means for the *factuality label* the following extraction should be labeled as *correct*:

("John Smith", "lives in", "Italy")

POLARITY: POSITIVE; MODALITY: CERTAINTY (+, CT)

The *modality* is the part of the factual annotation that gives us information of whether an extraction is a *certainty* or a *possibility* within a clause.

Rule F8: The extraction should only be labeled correct if it expresses the same modality (certain or possible) as its clause.

Suppose we have the following sentences:

> *"Dewayne Robertson expects to meet with the Jets"*
>
> *"Dewayne Robertson probably meets with the Jets"*
>
> *"Dewayne Robertson will meet with the Jets"*

then, the **factual label** of the both of the following extractions should be **correct**:

1. ("D. Robertson", "meet with", "Jets")   2. ("D. Robertson", "will meet with", "Jets")

Polarity: POSITIVE   ⇔   *Polarity: POSITIVE*

Modality: POSSIBILITY   *Modality: CERTAINTY*

because in the first sentence, Dewayne Robertson "expects to" meet with the Jets, which is merely a possibility, not a certainty. In the second one, it is a future tense ("will" meet ...), which is also not a certainty, but a possibility.

Rule F9: If the quantity(ies) within the extraction contain the proper phrase for indicating a phrase which expresses some sort of quantity (given that all the other rules F1 to F8 are also correct), the **factual label** is **correct**.

For example, if you have the sentence:

> *"At least two e-mails were marked as confidential."*

then the **factual label** for the following extraction is considered as **correct**:

> *("QUANT_S_1 e-mails", "were marked as", "confidential")*
>
> *Factuality: (+, CT)*
>
> *QUANT_S_1 = At least two*

Rule F10: If at least one quantity placeholder is not represented correctly in the extraction (i.e. it lacks crucial information), then the **factual label** is **incorrect.**

For example, considering the same sentence:

> *"At least two e-mails were marked as confidential."*

The following extraction's **factual label** is considered as **incorrect:**

> *("QUANT_S_1 e-mails", "were marked as", "confidential")*
> *Factuality: (+, CT)*
>
> *QUANT_S_1 = At least*

Rule F11: Ignore correference resolution parts. For example, if we have the sentence:

"John was home and he opened the door."

then both of the following extractions are considered to be equivalent and *correct*:

1. ("John", "opened", "door") ⇔ *2. ("He", "opened", "door")*

Rule F12: Ignore wrong form of a word as long as the lemmas are correct.

For example, in the sentence:

"John loves his wife."

the following two extractions are considered to be equivalent and *correct*:

1. ("he", "has", "wife") ⇔ *2. ("his", "has", "wife")*

2.2 Attribution

The *attribution* of a triple is the supplier of the information for the triple. The attributions themselves contain annotations for factuality, which are different from the annotation of factuality for the triple itself.

Rule A1: If attribution is captured within the extraction implicitly (it is within the extraction) or explicitly (it is annotated as an attribution), then the attribution should be labeled as *correct*.

For example, for the sentence:

"The State Department does not believe that more than 3 million Americans live outside of the U.S."

both of the following extractions are equivalent and should be labeled as *correct*:

("QUANT_S_1 Americans" "live outside of" "U.S.")

Attribution: (T. S. D., NEGATIVE, POSSIBILITY)

Modality: CERTAINTY

Polarity: POSITIVE

Quantities: [QUANT_S_1 = more than 3 million]

⇔

("T. S. D.", "does not believe that", "3 m. Americans live outside of U.S.")

Attribution: no attribution detected

Modality: CERTAINTY

Polarity: POSITIVE

The attribution's modality is **POSSIBILITY** because the attributer *believes* the statement. If the predicate was not *believe*, but it was a predicate expressing a certainty, like *knows*, then the modality would have been *CERTAINTY*.

Rule A2: If the attribution phrase contains words which are not part of the attribution phrase (e.g. instead of *The State Department* to have *The State Department does*) or lacks some words which are

essential to the meaning of the attribution phrase (e.g. instead of **The State Department** to have just **Department**) then the attribution should be labeled as incorrect.

Rule A3: If one of the attribution's factuality values is wrong (e.g. instead of **NEGATIVE Polarity** we have **POSITIVE Polarity** or instead of **POSSIBILITY Modality** we have **CERTAINTY Modality**), then the whole **attribution label** should be **incorrect**.