OHDSI NLP working group

Implementation of term_exists in NLP tools
Definition of term_exists

From the OMOP CDM v5.2 NOTE_NLP table definition:

**Term_exists** is defined as a flag that indicates if the patient actually has or had the condition. Any of the following modifiers would make **Term_exists false**:

- Negation = true
- Subject = [anything other than the patient]
- Conditional = true
- Rule_out = true
- Uncertain = very low certainty or any lower certainties

A complete lack of modifiers would make **Term_exists true**. For the modifiers that are there, they would have to have these values:

- Negation = false
- Subject = patient
- Conditional = false
- Rule_out = false
- Uncertain = true or high or moderate or even low (could argue about low)
In CLAMP, `term_exists` could be interpreted as follows:

- **Term_exists = false** if:
  - Concept is negated (Assertion = absent); or
  - Concept has condition modifier (Attribute CON); or
  - Concept has subject modifier and the subject is not patient (Attribute SUB $\neq$ patient or pt)

- **Term_exists = true** otherwise,

**NB:** uncertainty not assessed

https://clamp.uth.edu
MetaMap implementation of term_exists

In general **term_exists** could be interpreted as follows:

- **Term_exists = false** if CUI tagged as negated:
  - Negation Flag = 1 in MMI output; or
  - CUI in NegConcepts/NegConCUI in XML or JSON outputs
- **Term_exists = true** otherwise,

**NB**: subject, conditional, and uncertainty not assessed!

MetaMap implementation of term_exists

In **XML** output:

<table>
<thead>
<tr>
<th>Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>&lt;MMOs&gt;</code></td>
<td>All the XML output generated for an entire input record or citation, consisting of</td>
</tr>
<tr>
<td><code>&lt;MMO&gt;</code></td>
<td>- <code>&lt;CmdLine&gt;</code>: the command used to start MetaMap,</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;AA&gt;</code>: any acronyms/abbreviation(s) found in the text,</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;Negation&gt;</code>: any negation(s) found in the text, and</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;Utterances&gt;</code>: the utterance(s) found in the text</td>
</tr>
<tr>
<td><code>&lt;Negations Count=&quot;N&quot;&gt;</code></td>
<td>All the data generated for a negation, including</td>
</tr>
<tr>
<td><code>&lt;Negation&gt;</code></td>
<td>- <code>&lt;NegType&gt;</code>: the negation type,</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;NegTrigger&gt;</code>: the negation trigger,</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;NegTriggerPI&gt;</code>: the negation trigger's positional information,</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;NegConcepts&gt;</code>: the negated concept(s), and</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;NegConcPIs&gt;</code>: the negated concept's StartPos/Length positional information</td>
</tr>
<tr>
<td><code>&lt;NegConcCUI&gt;</code></td>
<td>The CUI associated with the negated concept</td>
</tr>
<tr>
<td><code>&lt;NegConcepts Count=&quot;N&quot;&gt;</code></td>
<td>The negated concept(s), consisting of</td>
</tr>
<tr>
<td><code>&lt;NegConcept&gt;</code></td>
<td>- <code>&lt;NegConcCUI&gt;</code>: the negated concept's CUI, and</td>
</tr>
<tr>
<td></td>
<td>- <code>&lt;NegConcMatched&gt;</code>: the negated concept's name</td>
</tr>
<tr>
<td><code>&lt;NegConcMatched&gt;</code></td>
<td>The name of the negated concept</td>
</tr>
<tr>
<td><code>&lt;NegConcPIs Count=&quot;N&quot;&gt;</code></td>
<td>The StartPos/Length positional information of the negated concept</td>
</tr>
<tr>
<td><code>&lt;NegConcPI&gt;</code></td>
<td>The StartPos/Length positional information of the negation trigger</td>
</tr>
<tr>
<td><code>&lt;NegTrigger&gt;</code></td>
<td>The negation trigger</td>
</tr>
<tr>
<td><code>&lt;NegTriggerPIs Count=&quot;N&quot;&gt;</code></td>
<td>The StartPos/Length positional information of the negation trigger</td>
</tr>
<tr>
<td><code>&lt;NegType&gt;</code></td>
<td>The negation type</td>
</tr>
</tbody>
</table>
MetaMap implementation of term_exists

In JSON output (Negations):

2.3 The Negations Pair

The Negations pair spans lines 39–58 in Appendix A, and represents the negations that MetaMap detected in the text. MetaMap’s negation detection is based on Wendy Chapman’s NegEx, and includes some significant extensions, documented here.

The only negated concept in the text is heart attack, which is negated by the trigger no; the data describing this AA are rendered in JSON as

```json
{
    "NegType": "nega",
    "NegTrigger": "no",
    "NegTriggerPIs": [
    {
        "StartPos": "0",
        "Length": "2"
    }],
    "NegConcepts": [
    {
        "NegConcCUI": "C0027051",
        "NegConcMatched": "-- Heart Attack"
    }],
    "NegConcPIs": [
    {
        "StartPos": "3",
        "Length": "12"
    }
}
```

Type of negation
Negation trigger
Negation trigger character offsets
Negation trigger starting offset
Negation trigger character length
Negated concept(s)
Negated concept’s CUI
Negated concept’s string
Negated concept character offsets
Negated concept starting offset
Negated concept character length
In cTAKES 4.0, term_exists could be interpreted as follows:

- Term_exists = false if EventMention/EntityMention attribute(s)
  - polarity=-1 OR
  - subject="other" OR
  - uncertainty=1
- Term_exists = true otherwise,

**NB:** conditional not assessed (rule_out maybe not assessed by Assertion module)

http://ctakes.apache.org
https://cwiki.apache.org/confluence/display/CTAKES/cTAKES
Local context analysis options in cTAKES:

- **Assertion** component contains the PolarityCleartkAnalysisEngine, which is machine learning based.

- **NE Contexts** component contains the original negation annotator, which is implemented using the ContextAnnotator, rule-based.

- **NegexAnnotator** component in YTEX-UIMA, a drop-in replacement for the original cTAKES negation annotator.
cTAKES implementation of term_exists

cTAKES Assertion component assesses:

• **Negation** attribute annotator: polarity (1 or -1, -1 = negated) org.apache.ctakes.typesystem.type.constants.CONST.NE_POLARITY_NEGATION_ABSENT and CONST.NE_POLARITY_NEGATION_PRESENT.

• **Uncertainty** attribute annotator: uncertainty (0 or 1, 1 = uncertain) CONST.NE_UNCERTAINTY_PRESENT and CONST.NE_UNCERTAINTY_ABSENT.

• **Generic** attribute annotator: generic (true or false). CONST.NE GENERIC_TRUE and CONST.NE_GENERIC_FALSE.

• **Subject** attribute annotator: subject ("patient", "other", and other values). Constants with the prefix ATTR_SUBJECT_* within class org.apache.ctakes.typesystem.type.constants.CONST.
In NOBLE, term_exists could be interpreted as follows:

- **Term_exists = false** if concept attribute(s) are:
  - Polarity=Negative_Polarity OR
  - Experiencer≠Patient_Experiencer OR
  - Certainty=DefiniteNegatedExistence_Certainty

- **Term_exists = true** otherwise,

**NB:** conditional not assessed (rule_out maybe assessed with polarity)

http://noble-tools.dbmi.pitt.edu
Summary of implementation of term_exists

\[ \text{term_exists} = \text{false} \] if:

<table>
<thead>
<tr>
<th></th>
<th>OMOP CDM NOTE_NLP</th>
<th>CLAMP</th>
<th>MetaMap</th>
<th>cTAKES</th>
<th>NOBLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negation</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>Subject</td>
<td>Not patient</td>
<td>Not patient</td>
<td>-</td>
<td>Not patient</td>
<td>Not patient</td>
</tr>
<tr>
<td>Conditional</td>
<td>true</td>
<td>true</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rule-out</td>
<td>true</td>
<td>(negation true?)</td>
<td>negation =true</td>
<td>negation =true</td>
<td>(negation true?)</td>
</tr>
<tr>
<td>Uncertain</td>
<td>true (low certainty)</td>
<td>-</td>
<td>-</td>
<td>true</td>
<td>true (negated)</td>
</tr>
</tbody>
</table>
## Summary of implementation of term_exists

**term_exists = true if:** 

Caveat: Not all attributes assessed, and term_exists=true not certain

<table>
<thead>
<tr>
<th></th>
<th>OMOP CDM NOTE_NLP</th>
<th>CLAMP</th>
<th>MetaMap</th>
<th>cTAKES</th>
<th>NOBLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negation</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
<td>false</td>
</tr>
<tr>
<td>Subject</td>
<td>patient</td>
<td>patient</td>
<td>-</td>
<td>patient</td>
<td>patient</td>
</tr>
<tr>
<td>Conditional</td>
<td>false</td>
<td>false</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rule-out</td>
<td>false</td>
<td>(negation false?)</td>
<td>negation =false</td>
<td>negation =false</td>
<td>(negation false?)</td>
</tr>
<tr>
<td>Uncertain</td>
<td>false (high, moderate certainty)</td>
<td>-</td>
<td>-</td>
<td>false</td>
<td>false (definite, probable)</td>
</tr>
</tbody>
</table>