

Documentation of Select Nouns from LMF lexicon Web Service¹

Author: Muntsa Padró. Barcelona, 2012.

Contact: muntsa.padro@upf.edu

1 Overview

Given a LMF file with nouns classified with a score, a threshold for the members of the class and a **threshold for the non members of the class, separate those elements that are over the threshold from** those that are not. There are three cases:

1. class nouns (score > 0) over the class threshold: convert them to "class=yes".
2. no-class nouns (score < 0) over the non-class threshold (in absolute value): convert them to "class=no".
3. nouns under the threshold (in absolute value): convert them to "class=unknown"

2 Inputs, outputs and formats

2.1 Inputs

- *input*: LMF file with noun classification with scores. In UTF-8. For example:

```
<Lexicon>
  <LexicalEntry id="le_1">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="accident"/>
    </Lemma>
    <Sense>
      <feat att="corpusID" val="lab"/>
      <feat att="eventive" val="1"/>
    </Sense>
  </LexicalEntry>
  <LexicalEntry id="le_2">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="assembly"/>
    </Lemma>
    <Sense>
      <feat att="corpusID" val="lab"/>
      <feat att="eventive" val="0.6"/>
    </Sense>
  </LexicalEntry>
```

¹ This documented is licensed under a Creative Commons Attribution 3.0 Spain License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/es/>.

```

<LexicalEntry id="le_3">
  <feat att="partOfSpeech" val="noun"/>
  <Lemma>
    <feat att="writtenForm" val="table"/>
  </Lemma>
  <Sense>
    <feat att="corpusID" val="lab"/>
    <feat att="eventive" val="-0.9"/>
  </Sense>
</LexicalEntry>
</Lexicon>

```

- *class_threshold*: Threshold for members of the class. Only nouns classified with a highest score than this threshold will be considered as class members.
- *non_class_threshold*: Threshold for non-members of the class. Only nouns classified with a negative score that is highest (in absolute value) than this threshold will be considered as non-class member.

Optional parameters:

- *class_name*: class to be filtered (in case there are several classes in input LMF file). If empty, the same filter will be applied to all classes.
- *corpusLabel*: corpusLabel to be filtered (in case there are several corpusLabels in input LMF file). If empty, the same filter will be applied to all corpusLabels.

2.2 Outputs

- *filtered_LMF*: LMF file with filtered information. The belonging/not belonging to the class is now stated with yes/no/unknown. For example, with the above input example and using a threshold for both members and non-members of the class of 0.7 to, we will get:

```

<Lexicon>
  <LexicalEntry id="le_1">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="accident"/>
    </Lemma>
    <Sense>
      <feat att="corpusID" val="lab"/>
      <feat att="eventive" val="yes"/>
    </Sense>
  </LexicalEntry>
  <LexicalEntry id="le_2">
    <feat att="partOfSpeech" val="noun"/>
    <Lemma>
      <feat att="writtenForm" val="assembly"/>

```

```
</Lemma>
<Sense>
  <feat att="corpusID" val="lab"/>
  <feat att="eventive" val="unknown"/>
</Sense>
</LexicalEntry>
<LexicalEntry id="le_3">
  <feat att="partOfSpeech" val="noun"/>
  <Lemma>
    <feat att="writtenForm" val="table"/>
  </Lemma>
  <Sense>
    <feat att="corpusID" val="lab"/>
    <feat att="eventive" val="no"/>
  </Sense>
</LexicalEntry>
</Lexicon>
```

Related Web Service:

Select Nouns from LMF lexicon Web Service: <http://lod.iula.upf.edu/resources/246>