Urgency Memo 3

Henrique Sposito

2024-03-22

Types/categories/dimensions of urgency (and components)

Frequency:

We code definite and indefinite adverbs, though their current scores might match their "frequency" all that well. For instance, is "hourly" the same as "always" and "yearly" the same as "hardly ever". Perhaps we should consider this in relation to either other words in sentence or as bigrams (e.g. "every hour" and "later than"). This could also be done taking advantage of the NLP tags and their ordering.

```
annotate_text("Every hour that goes by it is getting worse. Our first goal is to adress unemployment. W
```

```
## successfully initialized (spaCy Version: 3.7.4, language model: en_core_web_sm)
```

```
##    doc_id sentence_id token_id        token        lemma   pos  tag    entity
## 1   text1           1        1        Every        every   DET   DT
## 2   text1           1        2         hour         hour  NOUN   NN
## 3   text1           1        3         that         that  PRON  WDT
## 4   text1           1        4         goes           go  VERB  VBZ
## 5   text1           1        5           by           by   ADP   IN
## 6   text1           1        6           it           it  PRON  PRP
## 7   text1           1        7           is           be   AUX  VBZ
## 8   text1           1        8      getting          get  VERB  VBG
## 9   text1           1        9        worse          bad   ADJ  JJR
## 10  text1           1       10            .            . PUNCT    .
## 11  text1           2        1          Our          our  PRON PRP$
## 12  text1           2        2        first        first   ADJ   JJ ORDINAL_B
## 13  text1           2        3         goal         goal  NOUN   NN
## 14  text1           2        4           is           be   AUX  VBZ
## 15  text1           2        5           to           to  PART   TO
## 16  text1           2        6       adress       adress  VERB   VB
## 17  text1           2        7 unemployment unemployment  NOUN   NN
## 18  text1           2        8            .            . PUNCT    .
## 19  text1           3        1           We           we  PRON  PRP
## 20  text1           3        2       should       should   AUX   MD
## 21  text1           3        3       tackle       tackle  VERB   VB
## 22  text1           3        4         this         this   DET   DT
## 23  text1           3        5        issue        issue  NOUN   NN
## 24  text1           3        6           no           no   ADV   RB    DATE_B
## 25  text1           3        7        later        later   ADV   RB    DATE_I
## 26  text1           3        8         than         than   ADP   IN    DATE_I
## 27  text1           3        9         next         next   ADJ   JJ    DATE_I
## 28  text1           3       10         year         year  NOUN   NN    DATE_I
## 29  text1           3       11            .            . PUNCT    .
```

Timing:

We code relational and relative time words but fail to code them in relation or relative to others... We also miss some ordinal terms (e.g. "second", "third", "fourth") when rankink these.

Degree:

There are 4 grades of degree adverbs according to (Yoo, Kim, and Kwon 2011), they are: maximizer, boosters, compromizers, and dividers. Each is multiplied by a different value (i.e. 2, 1.6, 1.2., and 0.5 respectively). I wonder if we should adopt this convention instead of the "important" and "unimportant" words?

Commitment:

In theory this component relates to the "intensity of the promise". That is, it "we should" is less of a commitment than "we must". This has just been added. Is this appropriate?

What about promises?

Overall, this is developing well but we still have issues when it comes to words, subjects, and promises.

```
"Every hour (frequency) that goes by it is getting worse (degree)."
```

```
## [1] "Every hour (frequency) that goes by it is getting worse (degree)."
```

```
# neither hour nor worse are scored
"Our first (timing) goal is to adress unemployment (topic)."
```

```
## [1] "Our first (timing) goal is to adress unemployment (topic)."
```

```
# the topic is not identified automatically
"We should (promise/commit level) tackle this issue no later than next year (timing)."
```

```
## [1] "We should (promise/commit level) tackle this issue no later than next year (timing)."
```

```
# this is the only "promise" retained when we extract promises but ...
# extract_promises("Every hour (frequency) that goes by it is getting worse (degree). Our first (timing
```

Some possible solutions:

- merge all neighboring promises and/or nearby promises with the same topic
- code lemmas instead of words (for urgency)
- code subjects based on common nouns and entities (for topic)
- take into consideration NLP tags when scoring and coding words/lemmas related to urgency (e.g. only score frequency for adverbs) - How to better use the NLP tags?
- How to normalize scores? We currently try and score each word coded on a 0 to 1 scale, then, we add everything across and divide by the median. Is there a better approach?

What else?

The obvious question would be weather to include emotional words (i.e. emotion as a dimension/component of urgency see Yoo, Kim, and Kwon 2011). Besides that, many works classify adverbs (see Toboada et al. 2011), adjectives (see Josef Ruppenhofer, Wiegand, and Brandes 2014; Melo and bansal 2013; Paradis 1997), and nouns (see Ruppenhofer, Brandes, and Steiner). I wonder if we should take this convention instead. This way we can use and refer to a broader literature (and dictionaries, see below)...

Further resources:

Additional Dictionares

- SO-CAL dictionary

NLP

- Subjects
- Modal verbs
- An NLP-based novel approach for assessing national influence in clause dissemination across bilateral investment treaties

References for articles/books on word identification/categories/scales/weights (to be adapted)

- Ordering adverbs by their scaling effect on adjective intensity
- The Effect of Negators, Modals, and Degree Adverbs on Sentiment Composition
- Emotional index measurement method for context-aware service
- Scales and Scores An evaluation of methods to determine the intensity of subjective expressions
- Comparing methods for deriving intensity scores for adjectives
- Learning Scalar Adjective Intensity from Paraphrases
- Lexicon-Based Methods for Sentiment Analysis
- Good, Great, Excellent: Global Inference of Semantic Intensities
- Degree modifiers of adjectives in spoken British English
- Not All Words are Created Equal: Extracting Semantic Orientation as a Function of Adjective Relevance
- Analyzing Appraisal Automatically