

Urgency Memo 3

Henrique Sposito

2024-03-24

Dimensions of urgency

Frequency:

Currently we code definite and indefinite adverbs, though their current scores sometimes might not match their “frequency” For instance, is “hourly” the same as “always” and “yearly” the same as “hardly ever”? Perhaps we should better account for frequency 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. Additionally, the number of words coded needs to be expanded and we might want to make clear in a codebook what rules we use to score them (e.g. all/most/some frequency adverbs finishing in “ly” get a score of 0.5).

```
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 ranking these. Timing is an important dimension of political urgency that we need to get a better hold on. If our definition of urgency implies that urgent preferences will be addressed sooner, we should think about weighting “timing” differently.

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?

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 when word is an adverb) and when joining promises - that is, 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 whether to include emotional words (i.e. emotion as a dimension/component of urgency see Yoo, Kim, and Kwon 2011). What do you think?

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 so that we can use and refer to a broader literature in linguistics (and their dictionaries), what do you think?

Further resources:

Additional Dictionares

- SO-CAL dictionary

NLP

- Extracting subjects
- On 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

The week's work...

- Updated documentation and made functions more flexible/consistent
- Improved function's weighting of terms
- Added more urgency words and components

And some more trying ...

```
# SO-CAL dictionaries
sample_text <- read_csv("sample_text.csv")

## New names:
## Rows: 20 Columns: 4
## -- Column specification
## ----- Delimiter: "," chr
## (3): title, text, country dbl (1): ...1
## i Use 'spec()' to retrieve the full column specification for this data. i
## Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## * ' -> '...1'

# annotate text at the sentence level
sentences <- annotate_text(sample_text$text, level = "sentences")

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

# get future promises
promises <- extract_promises(sentences)

# extract subjects
```

```

subjects <- extract_subjects(promises)
# get related terms to subjects
related_terms <- extract_related_terms(promises$sentence, subjects) # needs updating
# code urgency
urgency <- get_urgency(promises, related_terms)
head(urgency$sentence, n = 10)

```

```

## [1] "But | realize that for every success story , there are other stories , of men and women who \n
## [2] "To lift the standards of our public schools , we achieved historic education reform which \n m
## [3] "https://obamawhitehouse.archives.gov/the-press-office/2016/10/05/remarks-president-paris-agreer
## [4] "| know that everyone here will soon be marching over to the Capitol \n building to peacefully a
## [5] "As \n the first order of business this year , | urge the Senate to do the same , and | know th
## [6] "Only when we know where a \n country has come from , could we possibly understand why the coun
## [7] "With a new national \n commitment , our scientists and engineers will overcome obstacles to tal
## [8] "We should offer proper guidance to our workers so that they will improve their moral and educat
## [9] "But given the situation we are facing , not passing a bill now would cost these Americans \n m
## [10] "This should \n absolutely occur in every single contested state before the election is certifi

```

```

# code adjectives and adverbs
.assign_adv <- function(promises) {
  adverbs <- read.delim("adv_dictionary1.11.txt", header = FALSE)
  out <- data.frame(sentence = 1:(length(promises[["sentence"]]]))
  for (i in 1:length(adverbs[,1])) {
    out[[adverbs[,1][i]]] <- stringr::str_count(promises[["sentence"]], adverbs[,1][i])*
      adverbs[,2][i]
  }
  rowSums(out[-1])
}
.assign_adj <- function(promises) {
  adjectives <- read.delim("adj_dictionary1.11.txt", header = FALSE)
  out <- data.frame(sentence = 1:(length(promises[["sentence"]]]))
  for (i in 1:length(adjectives[,1])) {
    out[[adjectives[,1][i]]] <- stringr::str_count(promises[["sentence"]], adjectives[,1][i])*
      adjectives[,2][i]
  }
  rowSums(out[-1])
}
alternative_urgency <- function(promises) {
  promises |>
    dplyr::mutate(adjectives = .assign_adv(promises),
                  adverbs = .assign_adj(promises),
                  urgency = (adjectives + adverbs)/stats::median(ntoken)) |>
    dplyr::arrange(-urgency)
}
# code alternative urgency
alt_urgency <- alternative_urgency(promises)
head(alt_urgency$sentence, n = 10)

```

```

## [1] "We should seize the day , \n remain confident in our history , exhibit greater historical init.
## [2] "In the upcoming new year , we will unswervingly uphold \n the great banner of socialism with C
## [3] "We must strengthen the great \n unity of the Chinese people of all ethnic groups and the great
## [4] "Fellow deputies , \n\n To do a good job of governing the country , the Party should do a good \
## [5] "We should realize effective disarmament and arms \n control in a fair , rational , comprehensi
## [6] "As we respond to great \n humanitarian needs , we must actively respond to the other great cha

```

```
## [7] "We should actively promote and protect \n human rights , make universal education available , a
## [8] "This means that we should , with a \n keen sense of responsibility and mission , adapt to new c
## [9] "The extensive \n and profound schools of thought they established covered many important ideas
## [10] "We are convinced that the Chinese people and the Chinese nation , with a proud civilization sp

# What do you think about this alternative?
```