Final Draft

LexAntonyms – Antonyms in the SPECIALIST Lexicon

Table of Contents

1. Introduction

1.1 Definition - what are antonyms?

Over the years, many different definitions of antonym have been proposed by linguists. Below are the definition and key description of antonyms we used in the SPECIALIST Lexicon:

- Antonyms are words that have **opposite or contrasting** meanings. [1997 Kempson] [1986 Cruse] [2011 Paradis & Willners] [2015 Paradis]
- Antonyms in the SPECIALIST Lexicon are called LexAntonyms.
- Antonyms are "similar, but different". Antonyms are not words with meanings that are as different as possible, but words with meanings that are very similar, but different in some respect, such as referring to opposite ends of the same scale for measuring the same property. For example, what is the antonym of “king”? Is it “queen” or “peasant” or “president”? There are various dimensions of meaning which separate “king” from other concepts — gender, social status. Choosing each of these dimensions (domains) defines its own set of antonyms.
- Follow the rule of ‘oppositeness is logical incompatibility’ – that is, if a thing can be described by one of the members of an antonym pair, it is impossible for it to be described by the other. For example, if a person is alive, they are not also dead. Kempson (1977) defines opposites as word-pairs with a “binary incompatible relation”, such that the presence of one meaning entails the absence of the other. In this sense, giant and dwarf are good opposites, while giant and person are not.
- Manually-created lexicons focus on opposites, such as “hot” and “cold”. They also do not explicitly list word pairs that are not opposites but yet have some degree of contrast in meaning, such as “warm” and “cold” or “tropical” and “freezing”.
- Contrast hypothesis: if a pair of words, A and B, are contrasting, then there is a pair of opposites, C and D, such that A and C are strongly related and B and D are strongly related. For example, there exists the pair of opposites “hot” and “cold” such that “tropical” is related to “hot”, and “freezing” is related to “cold”. “Tropical” and “freezing” are contrasting words. This is called the contrast hypothesis [2013 Mohammad].

1.2 Types of antonyms

There are many ways to categorize antonyms. Here are categorizations used in this project.

- Basic English:

Source: What are the different types of Antonyms in English Language [2016 Prateek Agarwal], https://medium.com/@hdi.prateek/what-are-the-different-types-of-antonyms-in-english-language-3a19db18504a

  - Complementary:
    Complementary antonyms have no middle ground. Such as: [boy-girl], [off-on], [night-day], [entrance-exit], [exterior-interior], [true-false], [dead-alive], [push-pull], [pass-fail], etc.
  - Relational:
    Relational antonyms are similar to complementary antonyms, except that both must exist for them to be antonyms of each other. Such as: [above-below], [doctor-patient], [husband-wife],

[servant-master], [borrow-lend], [give-receive], [predator-prey], [buy-sell], [instructor-pupil], etc.

- Graded (scalar - polarity), [1997, 2001, Paradis]:
  These antonyms deal with levels of comparison and they can be two words on a scale. Many are relative terms, which can be interpreted differently by different people. Such as [young-elderly], [hard-easy], [happy-wistful], [wise-foolish], [fat-slim], [warm-cool], [early-late], [fast-slow], [dark-pale], [long-short], etc.

- Good or bad antonyms (canonicity) [2003 Murphy]:
  - Canonical antonyms:
    - Pairs are contrasted, clearly opposable pairs. The members of those pairs express opposite properties on the basic dimensions (domains) that are central to human life and way of living across times and cultures. Such as color, space, temperature, wealth, etc.
    - Canonical antonym pairs are forming part of an antonym canon that is learned through experience of the language. Focus on the facts of meaning and usage that support antonym canonicity, rather than the contribution of formal properties like morphology, orthography, etc. [2003 Murphy]
    - Example: black-white-color, hot-cold-temperature, dead-alive-existence, short-long-length, slow-fast-speed, slow-quick-time, etc.
  - Non-canonical antonyms:
    - They are less clearly opposable pairs.
    - The dimensions that non-canonical antonyms relate to are more specific nominal meaning domains. For instance, the fact that the adjective “dark” is an antonym of “white” is not an immediate given, unless we are talking about, say, chocolate, but not when speaking about coffee.
    - Example: white-dark, hot-iced, dry-fleshy.

- Negated antonyms - unbounded and bounded antonymous adjectives [2006, 2012 Paradis and Willners]:
  There are three patterns of meaning for antonym pairs in relations to negation [2012 Jones, P:100]:
  - **Bounded: X = not Y, Y = not X**
    - X is synonymous with their negated (positive) antonyms (Y)
    - Y is synonymous with their negated (negative) antonyms (X)
    - Bounded adjectives are absolute and divide some conceptual domain into two distinct parts (mostly not scalar).
    - Some lexical items might be considered scalar, but are limited to the very end of the bounded extreme of the scale, such as (absolutely) terrific and (totally) disgusting.
    - Example: dead-alive, false-true, closed-open, wrong-right, empty-full
    - **Negated antonyms can be used in query expansion by substituting the synonymous antonyms for better performance.**
  - **Unbounded: X ≠ not Y, Y ≠ not X**
    - Unbounded antonymous meanings typically combine with scalar degree modifiers, such as very/fairly. They occupy the opposite poles of a scale. They are unbounded in the sense

that extreme values of something like 'long' and 'short' tend towards the extreme but actually never reach an end-point.
    - Example: narrow-wide, light-dark, low-high, sad-happy, rich-poor, long-short
    - **This should not be used for query expansion. However, they can be used for general NLP applications (see section 1.4).**
  - **X = not Y, Y ≠ not X, where X is the negated antonyms (asymmetric bounded)**
    - Example: wrong = not right, but, right ≠ not wrong
    - A negative orientation is synonymous with a negated positive adjective
      - ⇨ [the food was bad] = [the food was not good]
      - ⇨ [It is impossible] = [It is not possible]
    - However, a direct positive term does not have the same meaning as the negated negative term.
      - ⇨ [the food was good] ≠ [the food was not bad]
      - ⇨ [It is possible] ≠ [It is not impossible]
    - **Negative antonyms need to be identified and included in the LexAntonym list for NLP applications on query expansions.**

1.3 Negation detection and Antonyms:

- Negation status identification for findings or diagnoses is an important medical/clinical data mining problem. Negative qualifiers assigned to a medical condition may indicate the absence of the condition, so the ability to reliably identify the negation status of medical concepts affects the quality of results produced by the indexing and search tools.
- Negation detection includes [2017 Enger]:
  - cue detection: detects words that signal negation, such as *no*, *not*, *without*, and *unfortunate*
  - scope detection: identifies the span of the sentence that is affected by the negation
- Antonyms and negation detection:
  - Query expansion can be utilized when antonyms with negation are detected in NLP applications to increase performance (precision and recall) by avoiding additional processes of word sense disambiguation (WSD) and scope detection.
  - For example: [Pharmacological treatment was <u>not successful</u>, but surgical intervention provided an acceptable result.] [PMID: 11092366]
    - not successful: not (negation cue word) is detected as a negation, successful has antonym of unsuccessful (bounded antonym pair)
    - With query expansion on antonym with negation, the text is converted to [Pharmacological treatment was <u>unsuccessful</u>, …]
    - [not successful] is converted to [unsuccessful] and then precisely mapped to a single concept C1272705 for better performance in concept mapping.
    - Unsuccessful should be detected as negation.
    - Without query expansion on antonyms with negation [not successful], the following UMLS concepts are mapped:
      - ⇨ not|C1417836|C1518422
      - ⇨ successful|C1272703|C0597535
    - It results in 4 combinations of concept mapping and requires further NLP processes for WSD, while with substituting negated antonyms, results in only 1 concept mapping.

- Example: [Even though hormone-refractory prostate cancer is still incurable, it is not untreatable.]
  - ⇨ It is not untreatable = It is treatable
  - ⇨ This is not a negation. It is a false positive negation if only 'not' is used for negation detection.
- PMID: 17432562 [this hormonal therapy is not without side effects.]
  - ⇨ not|C1417836|C1518422 (negation)
  - ⇨ without|C0332288|
  - ⇨ side effects| C0879626|C0877248| C0001688
  - ⇨ Two negations, the conclusion is with some side-effect
  - ⇨ not without = without
  - ⇨ Not a negation (false positive)
- PMID: 15326468 [Self-inflicted spinal cord injury with a small needle is difficult, but not impossible.]
  - ⇨ not impossible = possible
  - ⇨ Not a negation (false positive)
  - ⇨ not|C1417836|C1518422 (negation)
  - ⇨ impossible|C4743675
  - ⇨ possible|C0332149

- Other examples:
  - "renal failure may ensue if medical intervention is not successful" (=unsuccessful)
  - "In acute suppuration of the knee, excision is never successful." (=unsuccessful)

1.4 Antonyms and negation detection in Natural Language Processing Applications

Antonyms and negation are essential for many NLP applications:

- Negation detection
  - Negation detection has generated special interest in extra-propositional aspects of meaning in text in practical NLP applications. Especially in systems processing medical and clinical text, such as outpatient notes, patient records, etc.
  - NLP applications have been developed to extract clinical information from medical records. The most common types of information extracted are diagnoses or findings. Identifying the negation status of a finding is as important as identifying the finding itself [2011 Goryachev]. For example, a finding occurring in a negated context may indicate the absence of some medical condition. Search tools looking for documents containing a particular finding may return irrelevant results if they do not take the negation into account (misleading).
  - Extra-propositional aspects of meaning in text exploring different aspects of meaning such as factivity (Saurı and Pustejovsky, 2009), uncertainty/hedging (Farkas et al., 2010), committed belief (Prabhakaran et al., 2010), and modalities (Prabhakaran et al., 2012a). Among these, negation detection has generated special interest because of demonstrated needs for negation detection capability in practical applications such as information retrieval (Averbuch et al., 2004), information extraction (Meystre et al., 2008), sentiment analysis (Wiegand et al., 2010; Councill et al., 2010), and relation detection (Chowdhury and Lavelli, 2013). Accurately detecting negations is especially important in systems processing medical and clinical text.
  - In fact, most phrases indicating negation are stop words in information retrieval systems and are not even used for indexing. Is negative indexing helpful and needed?

- Concept mapping: Example: [Mild hyperinflation without focal pneumonia]: “without” is important from this patient’s clinical record. It indicates the absence of focal pneumonia in the patient. Not capturing this extra-propositional aspect of negation concerning focal pneumonia will lead to wrong and harmful inferences in downstream processing, e.g. by a clinical decision support system.
- In NLP, cTake, CLAMP or Metamap, negation is detected to denote whether a given concept is absent or present [2017 Manimaran, 2001 Chapman]. But, no further mapping.

- Antonyms
  - Antonym detection has applications in tasks of understanding language, such as paraphrase detection and generation, or contradiction detecting.
    - detecting and generating paraphrases:
      - ⇨ [The dementors caught Sirius Black / Black could not escape the dementors]
    - detecting contradictions:
      - ⇨ [Kyoto has a predominantly wet climate / It is mostly dry in Kyoto]
  - Some antonym pairs include negative terms, which can be used for negation detection.
  - **Paraphrase** or **Paraphrasing** in computational linguistics is the natural language processing task of detecting and generating paraphrases. Applications of paraphrasing are varied including: question answering, machine translation, sentiment analysis (SA) and information retrieval (Roth and Schulte im Walde, 2014; Mohammad et al., 2013) and textual inference.
    - Sentiment Analysis: in Sentiment Analysis the correct discrimination of antonyms (e.g. good from bad) is extremely important to identify the positive or negative polarity of a text. [2015 Enrico Santus]. For example, words of the same and opposing polarity need to be distinguished [2014 Roth]
    - Textual entailment: need to identify hypernymy because of directional inference requirements. [2014 Roth]
    - information retrieval, [2004 Averbuch], [2013 Mohammad]
    - question answering:
    - text summarization:
    - plagiarism detection:
    - Machine translation : [2013 Mohammad]
    - Dialogue systems: [2013 Mohammad]
    - Information extraction: [2008 Meystre]
    - Relation detection: [2010 Weigand, 2010 Councill]
    - Clinical question answering [2006 Lee]
    - Clinical decision support [2009 Demner-Fushman]
    - Medical information extraction [2010 Uzner]
    - Patient history tracking [2012 Tymoshenko]

2. Project Objective & Requirements

2.1 Objective

To systematically generate antonymous terms in the SPECIALIST Lexicon to support NLP applications for enhanced performance

2.2 Contribution

No comprehensive manual verified antonym list exists for NLP applications.

2.3 Requirements: scope & criteria (definition) of LexAntonyms:

- Must be in the Lexicon
- Must be canonical antonyms
- Must have same POS
- Source-1: Affixal negation (negation from derivation), [Lehrer 1985, Ljung 1974]
  - It results from a process of affixation that creates a new form through derivation (and negation thereby).
  - Prefix-D
  - Suffix-D
- Source-2: antonyms in the MEDLINE (co-occurred)
  - Tag bounded (B), unbounded (UB) and asymmetric-bounded (AB) antonyms (should only apply to adj)
    - Bounded antonyms (X = not Y, Y = not X)
    - Unbounded antonyms (X ≠ not Y, Y ≠ not X)
    - Asymmetric-bounded antonyms (X = not Y, Y ≠ not X, where X is negation)
      - ⇨ Please note that only B and AB can be used for negative antonym substitution.
  - Tag negative antonyms (useful for asymmetric-bounded antonyms)
- Lexical items wise:
  Use antonyms to refer to pair-wise relation of lexical items (antonym-pairs) in the contexts that are understood to be semantically opposite.
  - Include spelling variants of both antonyms
  - Extend to synonyms of both antonyms
- Tags and format for Antonym candidates:
-

Antonym-1 (positive)	Antonym-2 (negative)	POS	Sources	Canonical Antonym	Type	Negation	Domain
successful	unsuccessful	adj	PD	Y	UB	N	quality
careful	careless	adj	SD	Y	UB	BN	quality
useful	useless	adj	SD	Y	UB	N	quality
centered	centerless	adj	SD	N	NA	O	location / quality
wasteful	wasteless	adj	SD	N	NA	O	quality
asleep	awake	adj	ML	Y	B	O	physical property /TBA
long	short	adj	ML	Y	UB	O	size
good	bad	adj			AB	BN	quality

with	without	prep			B	N	existence

- Tags:
  - Sources: [PD]: prefixD, [SD]: suffixD, [ML}: MEDLINE
  - Canonical Antonym: [Y], [N]
  - Type: [B]: bounded, [UB]: unbounded, [AB]: asymmetric bounded, [NA]; not applicable
  - Negation: [N]: strict negative, [BN]: braod negative, [O]: otherwise, not negative
    - ⇨ This tag is used to generate negation detection cue word.

- Output format and examples:

Antonym-1 (positive)	**Antonym-2 (negative)**	**POS**	**Type**	**Negation**	**Domain**
successful	unsuccessful	adj	UB	N	quality
careful	careless	adj	UB	BN	quality
useful	Useless	adj	UB	N	quality
asleep	awake	adj	B	O	physical property/TBA
long	short	adj	UB	O	size
good	bad	adj	AB	BN	quality
with	without	prep	B	N	existence

- POS of Antonyms
  - Must have same POS
    - ⇨ SuffixD also apply to the same POS
    - ⇨ care|noun|careless|adj
    - ⇨ care|noun|careful|adj
    - ⇨ We could derive above two suffixD pairs to generate careful|careless|adj
  - careful|noncare|adj is not a LexAntonym:
    - ⇨ They have same POS (adj) and opposing meanings. The domain of "careful" and “noncare” are “quality” and “existence”, respectively. Accordingly, they have different domains and therefore are not LexAntonyms.
  - Focus POS on adj first, then verb, noun, adv
  - Negation tag is used to generate negation detection key words [N], [BN], [O]
    - ⇨ [N]: true negation
    - ⇨ [BN]: broad negation
    - ⇨ [O]: otherwise, not negation
  - Prep: LexAntonyms from preposition are reviewed and listed as below:

Antonym-1	Antonym-2	POS	Type (bounded)	Negation	Domain
with	without	prep	B	N	existence
like	unlike	prep	AB	BN	quality
in	out	prep	B	O	location
up	down	prep	NB	O	location

to	from	prep	B	O	location

- Type field:
  - PD: prefix derivation (with negation)
  - SD: suffix derivation (with negation)
  - B: bounded
  - NB: non-bounded
  - AB: asymmetric-bounded
- Negation field:
  - N: negative (this indicates the field of antonym-2 is the negative antonym)
    - ⇨ This is used for asymmetric-bounded antonym in NLP query expansion
    - ⇨ It can be used for negation detection (Antonym-2 with Type B and negation (N) are used for negation detection word, such as without)
  - O: otherwise (no negation in the antonym pair)
- Tags of type (B and AB) and negation (N) can be used in query expansion to enhance concept mapping. All antonyms can be used for antonym related applications in NLP applications.
- Included domain/dimensions: to start, the most high-frequency domains are tagged, which would be: **size** (big/small)**, location** (above/below)**, time** (before/after)**, quality** (good/bad)**, intensity** (heavy/light)**, speed** (fast/slow), **age** (old/young), **color** (light/dark), **temperature** (hot/cold), **existence** (is/isn't), **physical property** (soft/hard), and TBA for other antonyms that do not fit clearly into one of the previous categories.
- Also, co-occurring concept with frequency, such as hot summer and cold winter.
  - ⇨ [hot-cold] to concept [summer-winter] to [temperature]

2.4 Negation Detection Cue Words

- Generate negation detection cue word list and identify negative patterns (for query expansion):
  - This negation detection cue word list and negative antonyms provide a more comprehensive data set for negation detection. Common patterns are:
    - not|[not + adj]
    - no|[no + noun]
    - without|[without + noun]:
  - Negation detection cue word generation can be generated in the following steps:
    - from the Lexicon for POSs of adv, pron, aux, modals, prep, det, conj
      - adverbs:
        - ⇨ true negative/strict negation (negative): never, no, not,
        - ⇨ broadly negative (broad_negative): hardly, seldom, rarely
      - pronoun:
        - ⇨ type=indef(neg): none, nobody, nothing, no-one, etc.
      - Auxiliaries (negative):
        - ⇨ variant=isn't;pres(thr_sing):negative
        - ⇨ variant=aren't;pres(fst_plur,second,thr_plur):negative
        - ⇨ variant=don't;pres(fst_sing,fst_plur,second,thr_plur):negative
        - ⇨ variant=haven't;pres(fst_sing,fst_plur,second,thr_plur):negative
        - ⇨ …

- Modals (negative):
  - ⇨ variant=mayn't;pres:negative
  - ⇨ variant=mightn't;past:negative
  - ⇨ variant=mustn't;pres:negative
  - ⇨ variant=couldn't;past:negative
  - ⇨ variant=cannot;pres:negative
  - ⇨ variant=can't;pres:negative
- preps:
  - ⇨ true negative/strict negation (negative): without
  - ⇨ broadly negative (broad_negative): unlikely (not used as negation cue words)
- det:
  - ⇨ true negative/strict negation (negative): no
- conj
  - ⇨ true negative/strict negation (negative): neither, nor

- From LexAntonyms:
  - ⇨ Valid canonical antonyms
  - ⇨ Negative tag is (N), exclude (BN and O)
  - ⇨ From above example, negation cue words include: unsuccessful, useless, careless (?), without (prep), etc..
- No negation detection cue word for POS of compl (that)

3 Approaches

3.1 Previous Work

- Manually-created lexicons have limited coverage and do not include most semantically contrasting word pairs (are not comprehensive).
- Lexical-semantic related words (combinations of surface form and word sense) are often co-occurring. Lexical-semantic relation applies to two lexical units—combinations of both lexical and semantic properties.
- Co-occurrence hypothesis: [Charles and Miller (1989)] proposed that antonyms occur together in a sentence more often than chance.
- Co-occurring text include synonyms, near-synonyms, antonyms, other functional related words (pen and paper), [2015 Tesfaye]
- Vector Space Models (VSMs) approach is used to find words that are closely related. Words are represented by vectors. Cosine similarity and other ML/DL clustering algorithms were used to find related words.
- Negation detection (simplified methods):
  - Cue detection
    - Prefix: dis-, im-, in-, ir-, un-
      - ⇨ acceptable|inacceptable|adj|N
      - ⇨ accurate|inaccurate|adj|N
      - ⇨ curved|incurved|adj|O
      - ⇨ patient|inpatient|noun|O
      - ⇨ Lexicon:

- B: a-, an-, de-, dys-, in-, under-
- N: anti-, contra-, counter-, dis-, il-, im-, ir-, mis-, non-, un-

- Suffix: -less
  - ⇨ useful|useless|adj|N
  - ⇨ centered|centerless|adj|O
  - ⇨ wasteful|wasteless|adj|O
  - ⇨
- Single word: no, without, fail, nor, neither

- Scope detection

3.2 Proposed Methods

- Affixal negation - generate LexSynonym pairs from affixal negation
  - Prefix derivations with negation (PD)
    - Generated LexAntonyms if the prefix is tagged as negation [N]
    - Same POS
    - single word
    - automatically put the negated antonyms (the longer ones) on the 2nd field (antonym-2)
    - generate candidate list
    - tag: all prefixD with negation tag are canonical antonym pair, no need for manual tag (maybe a quick review when we generate the data)
    - Are they always belong to bounded, unbounded, or asymmetric-bounded antonyms? If not, we need to tag them [B] [AB] [UB] as well? It seems [possible|impossible|adj] is [AB]; [funded|underfunded|adj] is [UB]; [successful|unsuccessful|adj] is [B]? Should we wait to see the real data to decide?
    - Example of process:
      - ⇨ Derivation: [impossible|adj|E0033808|possible|adj|E0049058|N|P|im]
      - ⇨ Candidate: [possible|impossible|adj]
      - ⇨ Tagged: [possible|impossible|adj|**Y**]
      - ⇨ LexSynonym: [possible|impossible|adj|PD|N|TBA]
  - Suffix derivation with negation (SD)
    - Generated LexAntonym if the suffix is tagged as negation [N]
    - Same POS
    - single word
    - automatically put the negated antonym (the longer one) on the 2nd field
    - generate candidate list
    - tag, is it possible that suffixD with negation tag are not antonym pair?
      - ⇨ Are they all bounded, unbounded, or asymmetric-bounded antonyms? Do we need to tag them? It seems [careful|careless|adj] is [UB]; [windowed|windowless|adj] is [B]; [affectional|affectionless|adj} is [AB] or [UB]? Should we wait to see the real data to decide?
    - Example:
      - ⇨ Derivations:

        - [care|noun|E0015334|careless|adj|E0015344|N|]
        - [care|noun|E0015334|careful|adj|E0015340|O|]
        - [care|noun|E0015334|noncare|adj|E0604159|N|]
    - Candidates:
        - [careful|careless|adj]
        - [careful|noncare|adj]
    - Tagged:
        - [careful|careless|adj|Y]
        - [careful|noncare|adj|**N**]

- LexAntonyms with bounded tags:
    - Find co-occurring pairs in MEDLINE 5-grams
        - **co-occurrence hypothesis**: Antonyms (and synonyms) usually co-occur [Charles and Miller, 1989], [Paradis et al. 2009], [Mohammad 2013], [Mohammad and Hirst 2006, Tesfaye 2015]
        - A Task-oriented evaluation shows that antonyms tend to co-occur in a five-word window. [Mohammad and Hirst 2006]
    - **Develop pattern of the negation detection key-word, such as "and", "vs", etc.**
        - Use existing antonyms (from ref.) to find the patterns:
            - and: [the short and long term]
            - or: [the short or long term]
            - as well as: [short-term as well as long-term]
            - either xxx or yyy: [either small or large]
            - than: [in small than in large], [levels in men and women]
            - from xxx to yyy: [from small to large]
            - to: [the small to the large], [small to large size effect]
            - XXX YYY or XXX ZZZ: [too small or too large]
            - Compare with: [women compared with men in]
            - vs: [the presence vs. absence of]
            - …
    - Must be in the Lexicon (that is why they are called **Lex**Antonyms)
    - Single word only
    - Exclude synonyms
        - Co-occurrence words include synonyms and antonyms
    - Exclude antonyms from affixal negations (PD & SD), already done in the above sections
    - Above a threshold
        - At the beginning of the project, just focus on high frequency if the list is too long
        - Previous studies pointed out that frequency is not a factor to determine antonymous relationship.
    - Auto-tag CUIs (UMLS concept since the application is for concept mapping)
        - Only used to help tagging bounded and asymmetric-bounded LexAntonyms
    - Auto-tag variant=inv & cat = adj
        - Only used to help tagging bounded and asymmetric-bounded LexAntonyms
            - non-scalar (unbounded): no comparative or superlative adjectives
            - Is there any other Lexicon codes (in addition to variant=inv) we should consider??

- Generate candidate list for tagging
  - ⇨ Format: [antonym-1 |CUI-1|antonym-2|CUI-2|POS|inv|tag-1|tag-2|tag-3]
- Tags
  - Tag-1: tag canonical antonyms:
    - ⇨ [Y]: canonical
    - ⇨ [N]: non-canonical
  - Tag-2: tag antonym type: [B|UB|AB]
    - ⇨ Should have CUIs
    - ⇨ Should be non-scalar
    - ⇨ [B]: if X = not Y, Y = not X
    - ⇨ [UB]: if X ≠ not Y, Y ≠ not X
    - ⇨ [AB]: if X = not Y, Y ≠ not X, where X is negation
    - ⇨ [O]: otherwise (keep this tag for now for the case of X ≠ not Y, Y = not X, where X is negation)
  - Tag-3: tag negation:
    - ⇨ [O]: otherwise, neither antonym-1 or antonym-2 is a negated antonym
    - ⇨ [N]: antonym-2 is the negative antonym
    - ⇨ [N1]: antonym-1 is the negative antonym
    - ⇨ [BN]: antonym-2 is the broad negative antonym
    - ⇨ [BN1]: antonym-1 is the broad negative antonym
- Generate root LexAntonym list
  - Format: [antonym-1| antonym-2|pos|type|negation|domain]
- Expand antonym to Lexeme level
  - Generate antonym pairs from spelling variants of antonyms
  - Generate antonym pairs from LexSynonym of antonyms
    - ⇨ Use nominalization from adj to noun, then from noun to verb
      Nominalizations do not necessarily hold the same antonym tag if both terms in the antonym pair have a noun through nominalization. For instance, “sleepy” and “awake” are antonyms; their nominalizations are sleepiness & awakeness, which are essentially synonyms (both describing the degree of tiredness). For example, medium sleepiness and medium awakeness are semantically identical.

- Release new table for negation detection cue word list (for negation detection):
  - Generate negation word list from Lexicon:
    - Adverbs
    - Modals
    - Auxiliaries
    - Pron
    - Prep
  - Generate negation word list from LexAntonyms:
    - antonym-2 with type of (B) and negation of (N).
  - Generate LexAntonym list
  - Identify negation detection pattern
    - [not + adj]

- [no + noun]
- [without + noun]
- [never + verb]

4. Results

- Negation Detection Cue word list
- LexAntonym list
- Test (TBD)

5. Discussion & Conclusion

- Distribution between antonyms, synonyms, and random word pairs (others)
  - Distribution of antonyms on POS: adj 59%, nouns 19%, Verbs 13% Other 9% (adv), [Murphy 2003?] [Mohammed 2013]
- POS distribution for LexAntonyms
- Our manual verified affixal negations are more precise than automatic computer generated approach. Give examples.
- Negated antonyms could be used for negation scope detection? (Need further study to see the data)
- Antonym pair must have same POS
- Apply existing word2Vec or GloVe and see the similarity score on antonyms (in addition to co-occurrence)
-

6. References:

- Carita Paradis. Good, better and superb antonyms. The Mental Lexicon. 2010
- Carita Paradis, Caroline Willners and Steven Jones. Good and bad opposites. The Mental Lexicon, 2010
- Carita Paradis, Joost van de Weijer, Caroline Willners and Magnus Lindgren. Evaluative Polarity of Antonyms. 2012
- Daniela Katunar. Prepositional antonymy in Croatian: a corpus approach. 2014
- Debela Tesfaye, Carita Paradis. On the use of antonyms and synonyms from a domain perspective. IJCOL, 2015
- Enrico Santus, Alessandro Lenci, Qin Lu, Chu-Ren Huang. When Similarity Becomes Opposition: Synonyms and Antonyms Discrimination in DSMs. IJCOL, 2015
- Enrico Santus, Qin Lu, Alessandro Lenci, Chu-Ren Huang. Unsupervised Antonym-Synonym Discrimination in Vector Space. 2014
- Enrico Santus, Qin Lu, Alessandro Lenci, Chu-Ren Huang. Taking Antonymy Mask off in Vector Space. 2014
- Federico Fancellu and Adam Lopez and Bonnie Webber. Neural Networks for Negation Scope Detection. 2016
- J. Manimaran and T. Velmurugan. Evaluation of lexicon- and syntax-based negation detection algorithms using clinical text data. Bio-Algorithms and Med-Systems, 2017

- Jones Steven, M.L. Murphy, Carita Paradis & Caroline Willners. Antonyms in English: Construals, constructions and canonicity. Cambridge University Press, 2012
- Joost van de Weijer, Carita Paradis, Caroline Willners, Magnus Lindgren. Antonym canonicity: Temporal and contextual manipulations. Brain & Language, 2014
- Joost van de Weijer, Carita Paradis, Caroline Willners and Magnus Lindgren. As lexical as it gets the role of co-occurrence of antonyms in a visual lexical decision experiment. 2012
- Martine Enger, Erik Velldal, Lilja Øvrelid. An open-source tool for negation detection: a maximum-margin approach. 2017
- Michael Roth, Sabine Schulte im Walde. Combining Word Patterns and Discourse Markers for Paradigmatic Relation Classification. Computational Linguistics, 2014
- Natalia Silveira. Learning to identify antonyms. 2013
- Peter D. Turney, Patrick Pantel. From Frequency to Meaning: Vector Space Models of Semantics. 2010
- Roser Morante, Eduardo Blanco. SEM2012 Shared Task: Resolving the Scope and Focus of Negation. 2012
- Saif Mohammad, Bonnie Dorr, Graeme Hirst. Computing Word-Pair Antonymy. Empirical NLP Conference , 2008
- Saif M. Mohammad, Bonnie J. Dorr, Graeme Hirst, Peter D. Turney. Computing lexical contrast. Computational Linguistics, 2013
- Saif Mohammad and Graeme Hirst. Distributional Measures of Concept-Distance: A Task-oriented Evaluation. 2006
- Steven Jones and M. Lynne Murphy. Antonymy in childhood: a corpus-based approach to acquisition. 2002
- Vijay Lingam, Simran Bhuria, Mayukh Nair, Divij Gurpreetsingh, Anjali Goyal, Ashish Sureka. Deep learning for conflicting statements detection in text. PeerJ Preprints. 2018
- Vinodkumar Prabhakaran, Branimir Boguraev. Learning structures of negations from flat annotations. Lexical and Computational Semantics conference, 2015
- Wendy W. Chapman Will Bridewell, Paul Hanbury Gregory F. Cooper, and Bruce G. Buchanan. A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries. Journal of Biomedical Informatics, 2002.