The Ontological Semantics of Antonyms

Max Petrenko
Princess Ekaterina R. Dashkova Moscow Humanities Institute, NLP Lab, Purdue University
maxpetrenko@alumni.purdue.edu

Christian F. Hempelmann
Department of Literature and Languages Texas A&M University–Commerce NLP Lab, Purdue University chempelm@purdue.edu

Abstract – This paper discusses the formal representation and lexical acquisition strategies of antonyms. The methodologies and toolbox of Ontological Semantic Technology (OST) are introduced as a framework. A brief outline of the technology and antonym taxonomies is followed by a discussion of how common and recalcitrant cases of antonyms can be represented and acquired, along with illustrations through pertinent examples.

Keywords: computational lexicon, antonyms, language engineering, ontological semantics, structural semantics

1 Introduction

The relation of antonymy is a commonly used approach to express a key meaning relation between certain word pairs and word sets—or fields—in particular in structural semantics. The purpose here has been to enable a description of meaning without having to actually analyze that meaning, that is, relate senses to each other systemically and label that relationship rather than having to give senses meaningful descriptions in and of themselves.

Providing meaningful descriptions requires an exhaustive and interrelated system, like the ontology underlying the present approach. This ontology is not just a taxonomy and cannot afford to be selective in where it draws lines of distinction, but has to do it across the senses of the lexicon of a language. In other words, the types of antonymy must be systematically related to each other as part of the ontology. The present approach integrates the advantages and insights of the selective contrastive approach afforded by existing research on antonyms and combines it with the exhaustive, descriptive, systematic apparatus of interrelated concepts for meaning representation of ontological semantic technology (OST).

1.1 The semantics of antonyms

While extensive research exists on the semantics and classification of antonyms, the computational aspects of representing and acquiring antonymy received little attention in the literature on computational linguistics. This section will present a brief outline of the semantics of antonyms and their taxonomy.

In general, the concept of antonymy is based on dichotomization. [Murphy 2003, 183] points to binarity as the fundamental conceptualization mechanism underlying antonymy, discussing the tendency to construe of taxonomically similar entities in sets of two, so that triplets like solid/liquid/gas would further bifurcate into solid/liquid, liquid/gas, and gas/solid. Section 2.5 will discuss how the binarity principle of antonymy formation could inform antonymy acquisition within OST.

Not least since Sapir (1944), a major division of antonyms into non-gradable and gradable has been made [cf. Riemer 2010, Murphy 2003, Lyons 1977], which can be exemplified on the antonym pairs dead-alive (non-gradable) and light-dark (gradable; e.g., X is lighter/darker than Y). Further subdivisions, often orthogonal, have been proposed based on linguistic-semantic as well as logical properties. These will be discussed in more detail in relation to their acquisition and detection in section 2. Overall, the group of non-gradable antonyms is an eclectic set exhibiting significant variation at both syntactic and semantic levels, including culture- and context-based variation, which presents significant challenges even for a semi-automatic acquisition system.

Several publications exist that attempt to account for the human acquisition of antonyms by capitalizing on the tendency of antonyms to co-occur in structurally similar sentences with higher than random consistency. [Justeson and Katz 1991] build on the hypothesis introduced in [Charles and Miller 1989] that adjectival antonym pairing is largely informed by statistical co-occurrence within same sentences. Based on the list of 55 antonym pairs frequency-tested on the Brown corpus, [Justeson and Katz 1991] established that an adjectival antonym pair on average co-occurs
once in every 15 sentence in syntactically recurring constructions, which, it is concluded, provides substantial training for antonymous associations. [Fellbaum 1995] and [Jones 2002] have explored syntactic contexts for co-occurring antonym pairs, and [Jones 2006] extended the approach to spoken English. The statistical line of research presented above turns out to be of interest to OST in that it supplies a verifiable list of commonly occurring antonym pairs, which can guide the selection of proper surface word forms during semi-automatic antonym acquisition.

1.2 Ontological Semantic Technology
Ontological Semantic Technology has received extensive coverage [Nirenburg and Raskin 2004, etc.] and substantial research and development has been carried out in a number of academic and commercial projects.

The lexicon entry template below illustrates the structure of lexical entries in the OST lexicon, specific to every language handled and separate from the core resource, the ontology (for a detailed review see Nirenburg and Raskin 2004, Petrenko 2010, for a sketch of the system architecture see Fig. 1):

```
(head-entry
  (sense-1, 2, 3…
   (cat(n/v/adj/pro/prep))
   (synonyms "")
   (anno
    (def "")
    (comments "acquisition time stamp or other notes")
    (ex ")")
  )
)
```

The main information that the sense entry contributes to the OST system for text processing is in the semantic description (sem-struc). In most cases, a concept from the ontology will form the root of the semantic description, usually further refined by specifying additional constraints about certain properties of the root concept. In addition, mappings via the syn-struc to senses of other words found in the sentence can be made. For example, commonly
the syntactic subject of a sentence is mapped to the semantic role of agent.

The ontology on which the sem-strucs are based is a lattice of concepts that are interrelated as mutual property fillers. At the top level it distinguishes objects, events, and properties; objects and events are further distinguished into physical, mental, and social subtypes; properties are distinguished into attributes and relations; etc (see Fig. 2). In previous applications, the ontologies contained 5-7,000 concepts with 15-20,000 unique properties, which, by inheritance, are multiplied into 69-90,000 “facts” across the ontology.

As has hopefully become clear even from this rather cursory overview, the acquisition and maintenance of resources is the crucial effort in OST. Therefore it is guided by a set of tools that facilitate this effort by the use of templates, propagation of similar senses, formal consistency, and semantic consistency, mainly against the constraints of the ontology. For detailed accounts of methodological, practical, and computational aspects of lexicon acquisition see [Taylor et al. 2010, Petrenko 2010, Raskin and Nirenburg 1995, and Raskin et al. 2010].

2 Antonyms in Ontological Semantic Technology

This section outlines representation formalisms and acquisition strategies for various types of antonyms.

2.1 Non-gradable antonyms

Lexical items with non-scalar semantics do not have a unified representation, and there are engineering considerations against having one in the static resources. For a scalable knowledge-based and domain-independent natural language processing system, explicit preference of otherwise unmarked
ontological concepts and properties representing a non-gradable pair like alive-dead (e.g. LIVE vs. LIVE(EPISTEMIC(0)), if implemented consistently, would lead to an unnecessary bias in extracting conceptual relations from input. While it would therefore be undesirable to accommodate the core antonym detection functionality in the static resources, and this functionality being one of the procedural, algorithmic components of an OST application, the ontology with its large descriptive thesaurus provides a comprehensive way to capture non-gradable antonym pairs.

Among the most common formats representing non-gradable antonym pairs are contrastive concept pairs, property fillers, and mutually exclusive facet values. An example below illustrates how a “male-female” pair can be captured via two contrastive literal fillers or values of the GENDER property.

(male-adj1
  (cat(adj))
  (anno(def "of male gender")
    (ex "a male suspect was arrested")
  (syn-struc(root($var0))
    (cat(adj))(np(root($var1))(cat(np))))))
  (syn-struc1
    (subject(root($var1))(cat(np)))
    (root(be))(cat(v))
    (directobject(root($var0))(cat(adj)))))))
  (sem-struc("$var1(gender(sem(male))))))))

(female-adj1
  (cat(adj))
  (anno(def "of female gender")
    (ex "a female suspect was arrested")
  ...
  (sem-struc("$var1(gender(sem(female))))))))

The representation above captures the contrastive gender properties as mutually exclusive values of the non-scalar attribute GENDER. An alternative representation could conceptualize the contrast as that between two full fledged concepts, as per below:

(male-adj1
  (cat(adj))
  (anno(def "of male gender")
    (ex "a male suspect was arrested")
  ...
  (sem-struc(male))

(female-adj1
  (cat(adj))
  (anno(def "of female gender")
    (ex "a female suspect was arrested")
  ...

There are procedural implications of attribute value-based representations and concept-based representations. The former seems more intuitive but makes the semantic structure less accessible to a processor as attribute values are seldom used by semantic processing modules.

2.2 Gradable antonyms
An efficient way of representing lexical items with scalar semantics has been introduced in [Raskin and Nirenburg 1995], where the acquisition of scalar adjectives like good, hot, etc. is acquired by rapid propagation of pertinent values on a single scale (EVALUATION, TEMPERATURE, etc.). Given the scalar nature ofgradable antonym pairs a similar approach could be adopted, which involves anchoring both members of an antonym pair in a single scalar attribute and assigning contrastive values:

(large-adj1
  (cat(adj))
  (anno(def "largely sized")
    (ex "he carried a large bag")
  (syn-struc(root($var0))
    (cat(adj))(np(root($var1))(cat(np))))))
  (syn-struc1
    (subject(root($var1))(cat(np)))
    (root(be))(cat(v))(directobject(root($var0))(cat(adj)))))))
  (sem-struc("$var1(size(greater-than(0.8))))))

small-adj1
  (cat(adj))
  (anno(def "small sized")
    (ex "he carried a large bag")
  (syn-struc(root($var0))
    (cat(adj))(np(root($var1))(cat(np))))))
  (syn-struc1
    (subject(root($var1))(cat(np)))
    (root(be))(cat(v))(directobject(root($var0))(cat(adj)))))))
  (sem-struc("$var1(size(less-than(0.3))))))

[Riemer 2010] and [Murphy 2003] point out the semantic non-comittedness of certain pairs of gradable antonyms. In particular, [Riemer 2010] observes that in the good-bad pair, good merely posits an evaluation scale and is not committed to a specific value, which is evidenced by the acceptability of the example, “How good was that film? – Really bad.” The contrastive member bad, on the other hand, is committed, which is in turn illustrated by the oddness of the example, “*How bad was the film? – The film is worse than the TV series, but they are both really good” [Riemer 2010: 138].
A reasonable adjustment to the representation formalism which would be sensitive to semantic non-comittedness of some gradable antonyms could involve postulating an open-ended variable whose value would be calculated contextually. More specifically, a TMR for “How good was the film” could override the default value of >.5 from the lexical entry and posit an open ended variable instead:

TMR: How good was the film?

exist
  theme $evaluative \text{ value} \ $var99

The value of the high-numbered variable is assigned based on the output of higher-level processing modules capable of aggregating contextual data from ambient input segments across sentences or even across texts in a corpus. In the example above, the ellipsis processing module would establish the co-reference of the EVALUATIVE modality filler in “Really bad.”, and since bad is uncommitted, its value would override the default value in the previous TMR.

2.3 Autoantonyms

Autoantonymous lexical items have received extensive coverage in [Murphy 2003] and [Riemer 2003]. Unlike uncommitted antonym pairs, which posited both representational and processing challenges, autoantonyms are essentially multi-sense entries representation-wise, and the challenge mainly lies in selecting an appropriate sense at the TMR building stage. Consider, for example, ‘sanction-v1’ and sanction-v2:

(sanction-v1
  (cat(v))
  (anno(def "to allow") (comments "")
    (ex "the court sanctioned investigation into bribery allegations in the corporation")
...
  (sem-struc(allow
    (agent(value("$var1$(restricted-to(human))))
    (theme(value("$var2$(restricted-to
      (sem(social-event culture-event)))))))))

(sanction-v2
  (cat(v))
  (anno(def "to penalize, impose sanctions on") (comments "") (ex "the union sanctioned the rogue state for stifling political opposition") (synonyms "")
...
  (sem-struc(allow
    (agent(value("$var1$(restricted-to(human))))
    (patient(value("$var2$(restricted-to(sem(human organization)))))))))

A seeming similarity of the two autoantonymous senses is considerably reduced by the different case roles on their restrictions: if the theme of ALLOW is restricted to EVENTS, ALLOW(EPISTEMIC(0)) only requires an animate PATIENT filler.

2.4 Conversives and reversives

The OST framework offers an explicit formalism for representing both conversive and reversive lexical items. For a conversive pair, the order of the same ontological properties is reversed with fillers remaining in situ. For a reversive pair, sibling properties are invoked to capture the opposition. To illustrate, the canonical conversive pair buy-sell would receive the following representation:

(buy-v1
  (cat(v))
  (anno(def "to purchase") (comments "")
    (ex "he bought a cell phone from a friend")
...
  (sem-struc(buy
    (agent(value("$var1$(restricted-to(human))))
    (theme(value("$var2$(restricted-to
      (sem(artifact food information))))))
    (source(value("$var3$(restricted-to(sem(human))))))))

(sell-v1
  (cat(v))
  (anno(def "to exchange for money") (comments "")
    (ex "he sold his cell phone to a friend")
...
  (sem-struc(sell
    (source(value("$var1$(restricted-to(human))))
    (theme(value("$var2$(restricted-to
      (sem(artifact food information))))))
    (agent(value("$var3$(restricted-to(sem(human))))))))

A reversive pair “enter-leave” would receive the following representation:

(enter-v1
  (cat(v))
  (anno(def "to move into a structure") (comments "") (ex "he entered the room") (synonyms "")
...
  (sem-struc(move
    (agent(value("$var1$(restricted-to(human))))
positives. there would be no scalable way in these cases the machine would over as to which sibling filler would constitute antonymy; siblings. This structure contains head c
of those non consideration exists against automatic representation TMR building and reasoning levels. An additional idiosyncrasies and thus bias the text processor at the
antonyms, their acquisition would lead to unjustified
OST knowledge resources provide for recognition
bitter Japanese, sweet
contextual variation of antonyms of taste for the word
2003: 173] discusses the cross
as well as
pointed in [Murphy 2003], the
problematic and, in some cases, unjustified. As
regarding antonym acquisition.

2.5 Acquisition of antonyms in
Ontological Semantic Technology

This section will discuss a preliminary strategy for automating the acquisition of a limited set of
gradable and all gradable antonyms in OST. The
strategy will allow for semi-automatic acquisition of
antonyms from manually acquired senses.

A number of intrinsic parameters of many non
gradable antonyms make their automatic acquisition
problematic and, in some cases, unjustified. As
pointed in [Murphy 2003], the semantic motivation
for relating many non-gradable antonyms is arbitrary,
as well as culture- and context-sensitive. [Murphy
2003: 173] discusses the cross-cultural and
contextual variation of antonyms of taste for the word
taste with sweet-pungent being preferred in
Japanese, sweet-sour in English in neutral contexts,
and sweet-salty (when describing snack food), sweet-
bitter (chocolate), sweet-dry (wine), etc. While the
OST knowledge resources provide for recognition
and semantic representation of such context-sensitive
antonyms, their acquisition would lead to unjustified
idioms and thus bias the text processor at the
TMR building and reasoning levels. An additional
consideration exists against automatic representation
of those non-gradable antonyms whose semantic
structure contains head concepts with multiple
siblings. This leaves the machine with the uncertainty
as to which sibling filler would constitute antonymy;
in these cases the machine would over-generate, and
there would be no scalable way to eliminate false
positives. To illustrate, multiple sibling fillers of the
attribute HAS-COLOR exist, many of which are
culturally contrastive (e.g. black-white), while others
are context-sensitive (e.g. black-red when describing
a roulette wheel), and there are no a priori intrinsic
linguistic clues for the machine to use to capture
those pairs.

Among non-gradable antonyms, only cases like
male-female, whose semantics involves selection
from a set of two property fillers, lend themselves to
acquisition and can be incorporated into the general
semi-automatic acquisition strategy described below.
The antonym-acquisition strategy takes
advantage of the representational potential of OST’s
knowledge resources, morphological regularities in
antonym formation, the available lists of high-
frequency antonym pairs and, available tagged
corpora.

The general steps of the strategy are outlined
below:

1) **Base antonym type detection**: Based on its
semantic structure, the type of the base antonym is
determined automatically. Senses whose semantic
structure posits a concept in the set of two siblings, or
a property with a set of two sibling fillers are non
gradable senses whose semantic structure contains a
scalar attribute. are gradable, other cases are
eliminated.

2) **Semantic derivation**: Depending on the base
antonym type, a respective rule is applied to derive
an appropriate semantic structure.
The following semantic derivation rules can be
formulated for each base antonym type:

For non-gradable antonyms:

1) (sem-struc(concept)) → (sem-struc(sibling-
concept))

2) (sem-struc(concept(facet(filler)))) → (sem-
struc(concept(facet(sibling-filler))))

For gradable antonyms:

3) (sem-struc(attribute(facet(literal-filler-a)))) →
(sem-struc(attribute(facet(literal-filler-b) {with a
and be being the only literal fillers of that
attribute})))

Rules (1-2) select an alternate sibling for the
head concept or filler, and rule (3) reverses the value
of the scalar attribute of the base antonym. The part
of speech and syntactic structure of both pair
members remain the same.

3) **Form derivation**: Morphological rules are applied
to the base antonym to form a new sense by adding
commonly used affixes. The list of regular antonym-
forming prefixes and suffixes is based on [Justeson
and Katz 1991] and, with minor additions, includes
the following: [a-, ab-, an-, dis-, il-, im-, in-, un-, un-,
non-, anti-, counter-, -less.] At this stage, the set is recursively applied to the base antonym form with each affix attached at a time. All the resulting forms are then passed to the form verification procedure.

4) **Form verification:** The core objective of this procedure is to assign an acceptable word form to an already instantiated sense template. The legitimacy of the newly derived word checked against a tagged corpus (e.g. Brown corpus) and the antonym list developed in [Justeson and Katz 1991] (referred to as JK list) containing 57 pairs of antonyms of higher than random frequency. The JK list comprises a set of 35 antonymous adjectival pairs tested in [Deese 1964] and an additional set of 22 pairs tested in [Justeson and Katz 1991.] The procedure takes all morphological derivatives from procedure (3) as input. If none of the candidate forms returns a match in the tagged corpora, the JK list, as well as online and offline lists of antonyms are taken advantage of. The output of the procedure essentially completes the antonym acquisition cycle: the antonym has been derived semantically, syntactically, and lexically. If no match is returned from the available corpora, the acquisition cycle is aborted.

5) **Rapid propagation:** Once a legitimate antonym has been confirmed, its synonyms are assigned an identical sense template, thus facilitating acquisition. The application of this procedure is largely dictated by the general grain size of lexical coverage. Finer-grained applications may require a more thorough representation capturing purely preferential differences between, say, the more standard big-little and the less acceptable large-little. More generic applications may allow for antonymy to be established between all members of each pair. Procedurally, steps similar to rapid propagation outlined in [Nirenburg and Raskin 1995, 2004] are taken, when the syntactic and semantic structures of a fully derived sense get propagated across a large set of similar entries. A similar procedure was also deployed in [Petrenko 2011] in relation to deriving synonyms of semi-automatically acquired deverbal senses.

As an illustration of the antonym acquisition strategy, a selected pair cold-hot with contrasting TEMPERATURE values is discussed below with each procedure described in detail. Procedure (1) will determine the base antonym cold as gradable:

\[
\text{cold-adj1: [arctic, frigid, gelid, glacial, icy, polar, bleak, cutting, raw-chilly, parky, crisp, frosty, nipping, nippy, snappy, frigorigic, frore, frosty, rimed, rimy, heatless, ice-cold, refrigerant, refrigerating, refrigerated, shivery, stone-cold, unheated, unwarmed]}
\]

Based on the output of procedure (1), semantic derivation rule (3) for gradable antonyms of procedure (2) will be called, which will return the following semantic structure:

\[
\text{(sem-struc(\text{\$var1}(\text{temperature(greater-than(0.8))}))}
\]

None of the word forms (*acold, *abcold, *ancold, *discold, *ilcold, *imcold, *incold, *uncold, *noncold, *anticold, *countercold, *coldless) generated morphologically via procedure (4) will match the existing corpora, so the system will have to resort to existing corpora of antonyms including the JK list. The JK list will return a match “hot” for “cold”, whose sem-struc, if acquired earlier, will match the template derived in procedure (1). After the cold-hot antonym pair has been derived, synonyms will be looked up in available online and offline corpora:

\[
\text{hot-adj1: [baking, baking hot, blistering, blistery, calefacient, warming, calefactory, calefactive, calorificant, calorific, fervent, fervid, fiery, igneous, heatable, heated, heated up, het, het up, hottish, overheated, red-hot, scorching, sizzling, sultry, stifling, sulphurous, sulphurous, sweltering, sweltry, thermal, torrid, tropical, tropic, white, white-hot]}
\]

Under the generic, coarse-grained functionality assumption, the instantiated templates for cold and hot will be propagated among all members of the retrieved sets. The acquirer will be offered an option of manually approving each built sense. Finer-grained acquisition would introduce additional demarcation criteria (stylistic preferential, etc.) and thus substantially increase the amount of manual effort.

4 Conclusion

This paper has presented a first account—descriptive and procedural—of antonyms within Ontological Semantic Technology. We first outlined, based on
relevant literature, the pertinent properties and major types of antonyms which directly resonate with the descriptive potential and computational goals of OST applications. We then discussed how non-gradable and gradable antonyms are represented within OST, how non-comittedness, auto-antonymy, reversion and conversion can be captured with OST formalisms. We concluded the paper by introducing a general strategy for semi-automatic acquisition of a limited set of non-gradable and all gradable antonyms in the OST lexicon, which takes advantage of manually acquired base senses, rich ontological and lexical knowledge resources, and relies on existing corpora.

5 References


