

Modelling of Adjectives in the Ontology-Lexicon Interface

John P. McCrae

Affiliation / Address line 1
Affiliation / Address line 2
Affiliation / Address line 3
email@domain

Francesca Quattri, Chrisitina Unger, Philipp Cimiano

Affiliation / Address line 1
Affiliation / Address line 2
Affiliation / Address line 3
email@domain

Abstract

The ontology-lexicon interface has become an important and successful tool for handling problems in NLP. The foundation of these models based on the separation of the ontological and lexical layers by means of the principle of semantics by reference to an ontology in description logics. However, as noted by other authors, the use of first order logic (hence also description logics), while effective for nouns and verbs breaks down in the case of adjectives. We propose that this is primarily due to a lack of logical expressivity in the ontology. In particular, many adjectives are i) gradable requiring fuzzy or non-monotonic semantics or ii) operator adjectives require second-order logic. We consider how we can handle the ontology-lexicon interface in the face of these more complex logical formalism, and show how these can be backward engineered into OWL based modelling by means of pseudo-classes, with application to question answering.

1 Introduction

Ontology-lexicon models, such as the *lemon* (Lexicon Model for Ontologies) (McCrae et al., 2012), have become an important model for handling a number of tasks in natural language processing. In particular, such ontology-lexica are built around the separation of a lexical layer describing how a word or phrase acts syntactically and morphology and a semantic layer describing how the meaning of a word is expressed in a formal logical model, such as the OWL (Web Ontology Language) (McGuinness et al., 2004). It has been shown that this principle known as *semantics by reference* (Buitelaar, 2010) is an effective model that can be used in tasks such as question answering (Unger and Cimiano, 2011) and natural language generation (Cimiano et al., 2013). In particular, its suitability to the task is driven by the fact that the application of this model to answering facts based on the DBpedia (Auer et al., 2007) knowledge base requires mostly understanding the nouns and the verbs of the sentence. However, as has been shown by the Question Answering over Linked Data (Lopez et al., 2013, QALD) tasks there are many questions that can be asked over this database that require a deeper semantic understanding of the representation of language. For example, questions such as

What is the highest mountain in Australia?

Require understanding of the semantics of ‘high’ in a manner that goes beyond the model of OWL based on classes, properties and individuals. The answer given in the QALD dataset for this question is as follows

```
SELECT DISTINCT ?uri WHERE {  
  ?uri rdf:type dbo:Mountain .  
  ?uri dbo:locatedInArea res:Australia .  
  ?uri dbo:elevation ?elevation .  
} ORDER BY DESC(?elevation) LIMIT 1
```

In particular, the interpretation of this question involves the understanding of how the word ‘high’ relates to the property `dbo:elevation` and how to express this semantics in a formal manner.

It has been claimed that first-order logic and thus by extension description logics, such as OWL, “fail decidedly when it comes to adjectives” (Bankston and others, 2003). To this extent we largely agree that the semantics of many adjectives are difficult or impossible to describe in first-order logic, however from the point of view of the ontology-lexicon interface the logical expressivity of the ontology is not a limiting factor. In fact, due to the separation of the lexical and ontology layers in a model such as *lemon*, it is possible to express the meaning of words without worrying about the formalism used in the ontology. To this extent we will first demonstrate that adjectives are in general a case where the use of description logics (DL) breakdown, and for which more sophisticated logical formalisms must be applied. We then consider to what extent this can be handled in the context of the ontology-lexicon, and introduce pseudo-classes, that is OWL classes with annotations, which we use to express the semantics of adjectives in a manner that would allow reasoning with fuzzy, high-order models. To this extent we base our models on the previously introduced design patterns (McCrae and Unger, 2014) for modelling ontology-lexica. Finally, we show how these semantics can be helpful in practical applications of question answering over the DBpedia knowledge base.

2 Classification of adjectives

There are a number of classifications of adjectives, and first we will start with the most fundamental distinction of *attributive* versus *predicative* usage, that is the use of adjectives in noun phrases (“*X* is a *A N*”) versus as objects of the copula (“*X* is *A*”). It should be noted that there are many adjectives for which only predicative or attributive usage is allowed.

Clinton is a former president.

*Clinton is former.

The baby is awake.

*The awake baby.

One of the principle classifications of the semantics of adjectives (for example (Partee, 2003)) is based on the meaning of adjective noun compounds relative to the meaning of the words by themselves. This classification is as follows (where \Rightarrow means it entails)

Intersective ($X \text{ is a } A N \Rightarrow X \text{ is } A \cap X \text{ is a } N$) Such adjectives work as if they were another noun and indicate that the compound noun phrase is a member of both the class of the noun and the class of the adjective. For example, in the phrase “Belgian violinist” it refers to a person in the class intersection $Belgian \cap Violinist(X)$, and hence we can infer that a “Belgian violinist” is a subclass of a “Belgian person”.

Subjective ($X \text{ is a } A N \Rightarrow X \text{ is a } N, X \text{ is a } A N \not\Rightarrow X \text{ is } A$) Such adjectives do not alter the meaning of the noun phrase itself, but only make sense with knowledge of the noun they refer to. For example, a “skilful violinist” is certainly in the class $Violinist(X)$, but if we knew that that person is a surgeon as well, we cannot conclude they are a skilful surgeon.

Privative ($X \text{ is a } A N \not\Rightarrow X \text{ is a } N$) These adjectives modify the meaning of a noun phrase to create a noun phrase that is incompatible with the original meaning. For example, a “fake gun” is not a member of the class of guns.

This classification is useful, however one further case is important to distinguish and that is of *relational* adjectives which have a meaning that expresses a relationship between two individuals or events, for example:

I would rather go to the sea than the mountains.

He is related to her.

Another important distinction to make with adjectives is whether they are *gradable*, in that whether it makes semantic senses to make a comparative or superlative statement with these adjectives. For example, adjectives such as ‘big’ or ‘tall’ can express relationships such as ‘ X is bigger than Y ’, however it is not possible to say one individual is ‘more former’. Most gradable adjectives are subsective, for example ‘a big mouse’ is not ‘a big animal’. An important group of gradable adjectives are, however, intersective, and we call such adjectives ‘absolute’ (following (Rusiecki, 1985)) as they refer to an ideal point on some scale, for example a ‘dry towel’ is ‘dry’, in that it has little or no water, however we can still talk about a towel being ‘drier’, in the sense of closer to the ideal of having no water than some other object.

Finally, we consider *operator* or *property-modifying* adjectives, which can be considered to be the same as privative adjectives, but in this case are understood as operators that change some property in the qualia structure of the class. For example, we may express the adjective ‘former’ as follows in lambda calculus (Partee, 2003):

$$\lambda C[\lambda x \exists t C(x, t) \cap t < \text{now}]$$

Such adjectives have not only a difference in semantic meaning but can also frequently have syntactic impact, for example in adjective ordering restrictions, as they may be reordered with only semantic impact (Teodorescu, 2006), e.g.,

A big red car.

? A red big car.

A famous former actor.

A former famous actor.

A related syntactic phenomena for Polish is explored by (Partee, 2003).

3 Representation of adjectives in the ontology-lexicon interface

In general it is assumed that adjectives form frames with exactly one argument except for extra arguments given by adjuncts, typically prepositional phrases. Most adjectives are thus associated with a predicative frame, stereotyped in English as:

$$X \text{ is } A$$

And with a second attributive frame, whose stereotype is not directly realizable but we give as:

$$X[\text{attr}]A$$

As such, when we encounter the attributive usage of an adjective such as in

Juan is a Spanish researcher

We understand this as the realization of two frames:

Juan is a researcher.

Juan $[\text{attr}]$ Spanish.

Note we do not provide modelling for adjectives that are part of a noun phrase, such as ‘polar bear’, which we would capture as a normal noun phrase with meaning of *ursus maritimus*.



Figure 1: Modelling of an intersective adjective ‘Belgian’ in *lemon*

3.1 Intersective adjectives

Intersective adjectives are the most straightforward class as in many cases they can mostly be explained as either being noun-like (denominal adjectives such as ‘British’) or verb-like (deverbal adjectives such as ‘broken’). Intersective adjectives have a single argument as with most adjectives and in this case it is natural that they refer to classes, which may be event classes such as described in (McCrae and Unger, 2014). For practical modelling examples we will use the *lemon* model as it is the most prominent implementation of the ontology-lexicon interface.

The primary mechanism of modelling the syntax-semantics interface in the context of *lemon* is by means of assigning a *frame* as a *syntactic behaviour* of an entry and giving it *syntactic arguments*, which can then be linked to the *lexical sense*, which stands in proxy for a true semantic frame in the ontology. For example, the modelling of an adjective such as ‘Belgian’ can be achieved as follows (depicted in figure 1).

```
belgian: a lemon:LexicalEntry ;
  lemon:canonicalForm belgian:Lemma ;
  lemon:synBehavior belgian:AttrFrame ,
                    belgian:PredFrame ;
  lemon:sense belgian:Sense .

belgian:Lemma lemon:writtenRep "Belgian"@eng .

belgian:AttrFrame lexinfo:attributiveArg belgian:AttrSynArg .

belgian:PredFrame lexinfo:copulativeArg belgian:PredSynArg .

belgian:sense lemon:reference [
  a owl:Restriction ;
  owl:onProperty dbpedia:nationality ;
  owl:hasValue dbpedia:Belgium ] ;
  lemon:isA belgian:AttrSynArg ; belgian:PredSynArg .
```

Note, that here we use the external vocabulary LexInfo(Cimiano et al., 2011) to define the meaning of the arguments of the frame as the *attributive argument*, corresponding to the frame stereotype “A [attr] X” and the *copulative argument* for the frame stereotype “X is an A”. Furthermore, the class of Belgians is not named in our reference ontology DBpedia, so we introduce an anonymous class with the axiomatization, $\exists \text{nationality.Belgium}$. It is in fact common that the referent of an adjective is not named

in an ontology and as such we tend to model denominal adjectives as the form $\exists \text{prop.Value}$, where Value is the reference of the noun from which the adjective is derived. This modelling is so common that it has already been encoded as two patterns, called `IntersectiveObjectPropertyAdjective` and `IntersectiveDatatypePropertyAdjective`, see (McCrae and Unger, 2014). Similarly, most deverbal adjectives refer to an event, and as such a common modelling is of the form $\exists \text{theme}^{-1}.\text{EventClass}$, for example ‘vandalized’ may be $\exists \text{theme}^{-1}.\text{VandalismEvent}$.

3.2 Gradable adjectives

Gradable adjectives are naturally defined relative to a particular property¹, that is it is natural to say that big refers to size, however it is clear then that small also refers to size and as antonyms they cannot both refer to the same ontological concept. As such, we introduce the concept of *covariance* and *contravariance*, which refers to whether the comparative form indicates a higher property value for the subject or the object. That is that ‘big’ is covariant with size, as bigger things have a higher size value, and ‘small’ is contravariant with size. We also introduce a third concept of *absolute gradability*, which states that these objects are better described by these adjectives as they approach some ideal value. A common example of this is colours, which where we may say that some object is redder than another if it is closer to some ideal value of red (e.g., RGB 0xff0000).

While these concepts well handle the comparative usage of adjectives, the predicative and superlative usage of adjectives is complicated by three factors that we will outline below. Firstly, we notice that gradable classes are not crisply defined as with the case of many intersective adjectives, that is that while we can clearly define all people in the world as ‘British’ or ‘non-British’, by who holds a passport of the United Kingdom, it is not easy split the world’s population into ‘tall’ and ‘not tall’. In fact, while it may be easy to say that someone who is 6’6” (198cm) tall is ‘tall’ it is not clear whether someone of height 6’ (182cm) is ‘tall’ although they are above average height for a man. As such, the class boundary of a gradable adjective are naturally fuzzy. Secondly, we note that these class boundaries are non-monotonic, that is that with knowledge of more instances of the relative class we must revise our class boundaries. This is especially the case for the superlative as the discovery of a new tallest person in the world would remove the existing tallest person in the world from the class of tallest person in the world. This, non-monotonicity also affects the class boundaries of the gradable class itself, for example in the 18th century the average height of a male was 5’5” (165cm) and as such a male of 6’ would have been considered clearly to be tall. As such, we can conclude that each instance added to our ontology must revise the class boundaries of a gradable class, hence leading to the fact that gradable adjectives are fundamentally non-monotonic. Finally, we must notice that gradability can only be understood relative to the class that we wish to grade, that is that while it is unclear as to whether 6’ is tall for a male, given the current average height of a female of about 5’4” (162cm) it is clear that 6’ is clearly tall for a female.

As such we can clearly conclude that gradable adjectives are *fuzzy* and *non-monotonic* and *context-sensitive*, all of which are incompatible with the description logic used by OWL. Currently there are only limited models for representing fuzzy logic in the context of the web (Zhao and Boley, 2008). As such, we introduce a new model, which we name *lemonOILS* (The *lemon* Ontology for the Interpretation of Lexical Semantics)². This ontology introduces three new classes:

- `CovariantScalar`: Indicating that the adjective is covariant with its bound property
- `ContravariantScalar`: As above, but contravariant
- `AbsoluteScalar`: Indicating that the property represents similarity to an absolute value

In addition, the following properties are introduced to enable the description of gradable adjectives. Note, that all of these properties are typed as *annotation properties* in the OWL ontology, so that they do not interfere with the standard OWL reasoning.

¹Note in many cases the property is quite abstract such as in ‘breakable’ **moar**

²<http://lemon-model.net/oils>

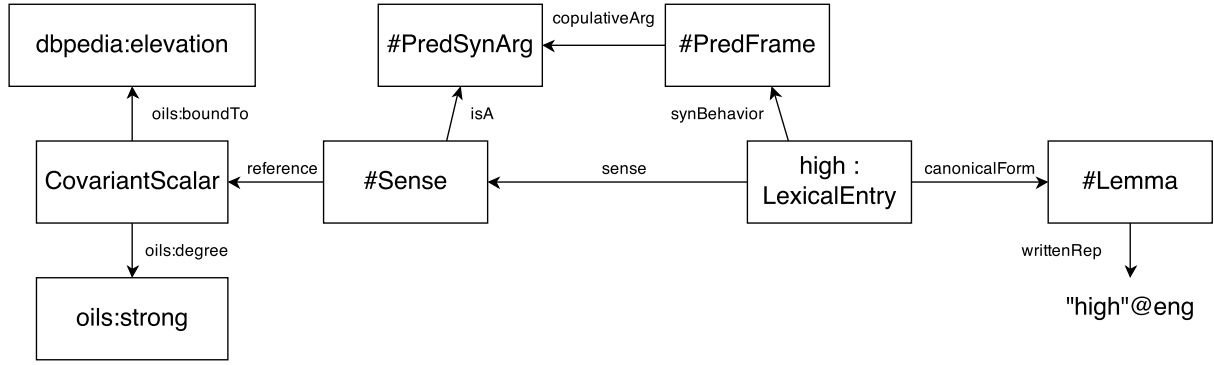


Figure 2: An example of the modelling of ‘high’ in *lemon*

- `boundTo`: Indicates the property that a scalar refers to (e.g., ‘size’ for ‘big’)
- `threshold`: A sensible minimal value by which the adjective can be said to hold.
- `degree`: One of `weak`, `medium`, `strong` or `very strong` corresponding to approximately 50%, 25%, 5% or 1% of all known individuals.
- `comparator`: Indicates an object property that is equivalent to the comparison of the adjective. e.g., an object property `biggerThan` may be considered a comparator for the adjective class `big`.
- `measure` (TODO)
- `default value` (TODO)

Using such classes we can capture the semantics of gradable adjectives syntactically but not formally within an OWL model, as such we call such introduced classes *pseudo-classes*. An example of modelling an adjective such as ‘high’ is given below (figure 2).

```
high: a lemon:LexicalEntry ;
  lemon:canonicalForm high:Lemma ;
  lemon:synBehavior high:PredFrame ;
  lemon:sense high:Sense .

high:Lemma lemon:writtenRep "high"@eng .

high:PredFrame lexinfo:copulativeArg high:PredArg .

high:Sense lemon:reference [
  rdfs:subClassOf oils:CovariantScalar ;
  oils:boundTo dbpedia:elevation ;
  oils:degree oils:strong ] ;
  lemon:isA high:PredArg .
```

As an example of a way in which it would be possible to interpret these annotations, we consider Markov Logic (Richardson and Domingos, 2006), which is an extension of first-order logic in which each clause is given a cost. The process of reasoning is thus transformed into an optimization problem of finding the extension, which minimizes the summed weight of all violated clauses. As such we can formulate a gradable adjective based on the number of known instances as follows:

$$\forall x \in C, y \in C : size(x) > size(y) \rightarrow big_C(x) : \alpha$$

$$\forall x \in C, y \in C : size(x) < size(y) \rightarrow \neg big_C(x) : \beta$$

In this way, the classification of an object into big or small is defined as follows:

Lemma: An individual, $x \in C$, has $big_C(x)$ if and only if

$$|\{y \in C, size(y) > size(x)\}| \alpha > |\{y \in C, size(y) < size(x)\}| \beta$$

The proof of this lemma is trivial, and we see the three properties clearly outlined above, in that we have non-monotonicity (in that more individuals may change whether we consider an individual to be ‘big’ or not, that we have fuzziness, given by the strength of the probability of the proposition $big_C(x)$).

Thresholds, defaults and multiple classes (Francesca to help)

3.3 Operator adjectives

Operator adjectives are those that combine to alter the meaning of the adjective itself. There are two primary issues with the understanding of the adjective in this manner, firstly that the reference of the lexical item does not directly refer to an existing item in the ontology, but rather is novel and productive and secondly that the compositional nature of adjective-noun compounds is no longer simple as in the cases of intersective and gradable adjectives. This means that we must acknowledge operator adjectives in both the lexicon and the adjective. To this extent we define a frame known as the operator adjective frame, whose prototype is:

X is a $A NP$

This leads to the odd case that operator adjectives are then considered the head of such a construction! **hmm...** This understanding is unambiguous assuming there is no attributive frame for the adjective as well. In this case we can understand the reference of the adjective as a property that relates an individual to a class. As such it is clear, that the reference of an operator adjective is a higher-order predicate. Fortunately, in the case of OWL we can cheat on this second-order nature by means of *punning*, which allows a class to also be an individual. If we thus assume that operator adjectives are essentially puns, then it follows that we can assume that the reference of an operator adjective is thus a property. As such, for example, we can model an adjective such as ‘former’ as referring to a property such as `heldRole` whose range is a class of roles punned as individuals. This approach is effective, however it has limits in general **does it???**

We will not be able to easily create a vocabulary that can fully describe the semantics of the adjective within the context of OWL as the second nature order of the logic cannot be captured well within the framework of description logic. However, we can use the punning trick described above to capture the semantics of the adjective. To do this we need to add a frame on the syntax side, that indicates that the argument of the adjective is in fact the noun phrase. We would do this as follows (figure 3):

```
former: a lemon:LexicalEntry ;
lemon:canonicalForm former:Lemma ;
lemon:synBehavior former:OperatorFrame ;
lemon:sense former:Sense .

former:Lemma lemon:writtenRep "former"@eng .

former:OperatorFrame lexinfo:copulativeArg former:Subject ;
lexinfo:attributiveOperator former:Object .

former:Sense lemon:reference onto:heldRole ;
lemon:subjOfProp former:Subject ;
lemon:objOfProp former:Object .
```

The usage of this frame is intended such that a phrase such as

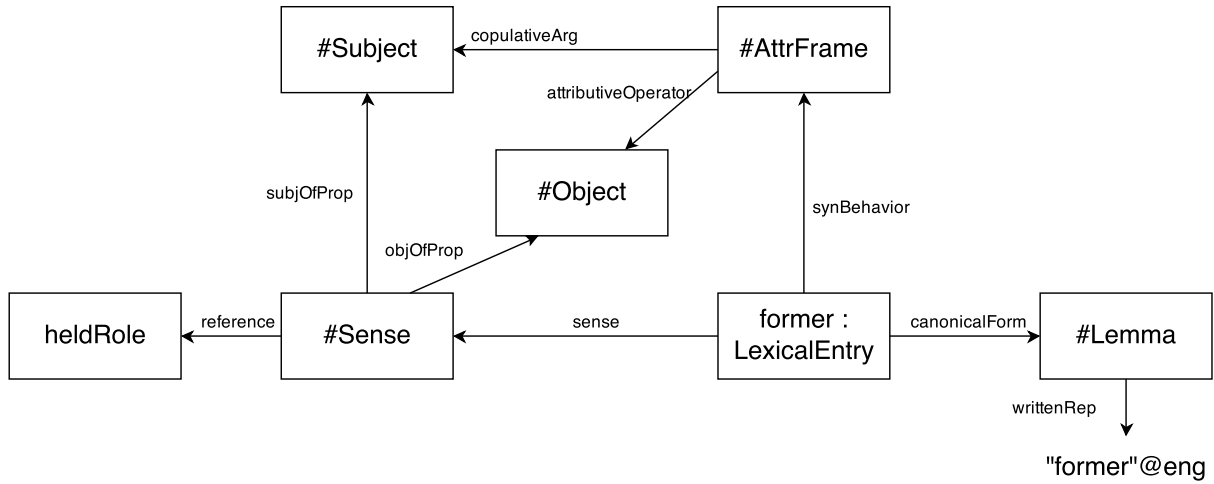


Figure 3: An example of modelling ‘former’ in *lemon*

Clinton is a former president

Is interpreted as

```
ontology:Clinton ontology:heldRole ontology:President .
```

3.4 Relational adjectives

Relational adjectives are among the simplest and modelled with another frame, which extends the attributive frame by allowing for a prepositional phrase adjunct. As such we can model ‘known’ with the frame ‘X is known to Y’ and reference `foaf:knows` as:

```
lexicon:known a lemon:LexicalEntry ;
  lemon:canonicalForm known:Lemma ;
lemon:sense known:Sense ;
lemon:synBehavior known:Frame .
```

```
known:Lemma lemon:writtenRep "known"@eng .
```

```
known:Frame lexinfo:attributeArg known:Subject ;
  lexinfo:prepositionalObject known:Object .
```

```
known:Sense lemon:reference foaf:knows ;
  lemon:subjOfProp known:Subject ;
lemon:objOfProp known:Object .
```

```
known:Object lemon:marker lexicon:to .
```

4 Adjectives in question answering

(Christina, some examples for QALD)

5 Related work

The categorization of adjectives in terms of formal semantics goes back to Montague(1970), however one of the most significant attempts to assign a formal meaning was carried out in the Mikrokosmos project(Raskin and Nirenburg, 1995).

Adjective based inference (Amoia and Gardent) Adjectives A uniform semantic approach (Abdullah Frost) Treatment of adjectives in SIMPLE: Theoretical Observations (Peters and Peters) The Description

of Adjectives for Natural Language Processing: Theoretical and Applied Perspectives (Bouillon and Viegas)

6 Conclusion

multilinguality?

Acknowledgements

References

- Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. Dbpedia: A nucleus for a web of open data. In *The semantic web*, pages 722–735. Springer.
- Paul Bankston et al. 2003. Modeling nonintersective adjectives using operator logics. *The Review of Modern Logic*, 9(1-2):9–28.
- Paul Buitelaar, 2010. *Ontology-based Semantic Lexicons: Mapping between Terms and Object Descriptions*, pages 212–223. Cambridge University Press.
- Philipp Cimiano, Paul Buitelaar, John McCrae, and Michael Sintek. 2011. Lexinfo: A declarative model for the lexicon-ontology interface. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(1):29–51.
- Philipp Cimiano, Janna Lüker, David Nagel, and Christina Unger. 2013. Exploiting ontology lexica for generating natural language texts from rdf data. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 10–19.
- Vanessa Lopez, Christina Unger, Philipp Cimiano, and Enrico Motta. 2013. Evaluating question answering over linked data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 21:3–13.
- John P. McCrae and Christina Unger, 2014. *Design Patterns for the Ontology-Lexicon Interface*, chapter ?, page ??? Springer.
- John McCrae, Guadalupe Aguado-de Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asunción Gómez-Pérez, Jorge Gracia, Laura Hollink, Elena Montiel-Ponsoda, Dennis Spohr, et al. 2012. Interchanging lexical resources on the semantic web. *Language Resources and Evaluation*, 46(4):701–719.
- Deborah L McGuinness, Frank Van Harmelen, et al. 2004. Owl web ontology language overview. *W3C recommendation*, 10(2004-03):10.
- Richard Montague. 1970. English as a formal language. In Bruno Visentini et al, editor, *Linguaggi nella società e nella tecnica*, pages 189–224. Milan: Edizioni di Comunità.
- Barbara H Partee. 2003. Are there privative adjectives. In *Conference on the Philosophy of Terry Parsons, University of Massachusetts, Amherst*.
- Victor Raskin and Sergei Nirenburg. 1995. Lexical semantics of adjectives. *New Mexico State University, Computing Research Laboratory Technical Report, MCCS-95-288*.
- Matthew Richardson and Pedro Domingos. 2006. Markov logic networks. *Machine learning*, 62(1-2):107–136.
- Jan Rusiecki. 1985. *Adjectives and Comparison in English: A Semantic Study*. Longman Linguistic Library.
- Alexandra Teodorescu. 2006. Adjective ordering restrictions revisited. In *Proceedings of the 25th West Coast Conference on Formal Linguistics*, pages 399–407. Citeseer.
- Christina Unger and Philipp Cimiano. 2011. Pythia: Compositional meaning construction for ontology-based question answering on the semantic web. In Rafael Munoz, editor, *Natural Language Processing and Information Systems: 16th International Conference on Applications of Natural Language to Information Systems, NLDB 2011, Alicante, Spain, June 28-30, 2011. Proceedings*, volume 6716, pages 153–160. Springer.
- Jidi Zhao and Harold Boley. 2008. Uncertainty treatment in the rule interchange format: From encoding to extension. In *URSW*.