Modelling the Semantics of Adjectives in the Ontology-Lexicon Interface

John P. McCrae (blind rev.)
Affiliation / Address line 1
Affiliation / Address line 2
Affiliation / Address line 3
email@domain

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

Christina Unger, Philipp Cimiano
Affiliation / Address line 1
Affiliation / Address line 2
Affiliation / Address line 3
email@domain

Abstract

The treatment of the semantics of adjectives is notoriously challenging, particularly in the context of the so called ontology-lexicon interface which attempts to capture the semantics of words by reference to an ontology in description logics or some other, typically first-order, logical formalism. The use of first order logic (hence also description logics), while effective for nouns and verbs, breaks down in the case of adjectives. We argue that this is primarily due to a lack of logical expressivity in the underlying ontology languages. In particular, beyond the straightforward *intersective adjectives*, there exist *gradable adjectives*, requiring fuzzy or non-monotonic semantics, as well as *operator adjectives*, requiring second-order logic for modelling. We consider how we can extend the ontology-lexicon interface as realized by extant models such as *lemon* in the face of the issues mentioned above, in particular those arising in the context of modelling the ontological semantics of adjectives. We show how more complex logical formalisms required to capture the ontological semantics of adjectives can be backward engineered into OWL-based modelling by means of pseudo-classes. We discuss the implications of this modelling in the context of an application to ontology-based question answering.

1 Introduction

Ontology-lexicon models, such as *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 morphologically, and a semantic layer describing how the meaning of a word is expressed in a formal logical model, such as OWL (Web Ontology Language) (Deborah L. McGuinness and others, 2004). It has been shown that this principle known as *semantics by reference* (Buitelaar, 2010) is an effective model that can support the task of developing question answering systems (Unger and Cimiano, 2011) and natural language generation (Cimiano et al., 2013) over backends based on Semantic Web data models. The Pythia system, which builds on the *lemon* formalism to declaratively capture the lexicon-ontology interface, for example, has been instantiated to the case of answering questions from DBpedia (Unger and Cimiano, 2011) However, as has been shown by the Question Answering over Linked Data (Lopez et al., 2013, QALD) benchmarking campaigns, there are many questions that can be asked over this database that require a deeper representation of the semantics of words, adjectives in particular. For example, questions such as 1a require understanding of the semantics of 'high' in a manner that goes beyond the expressivity of OWL. The answer given in the QALD dataset for this question is shown in 1b. In particular, the interpretation of this question involves the understanding of how the word 'high' relates to the property `dbo:elevation`, including ordering and subset selection operations, and how to express this semantics in a formal manner.

1. (a) What is the highest mountain in Australia?
   (b) `SELECT DISTINCT ?uri WHERE {`

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

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, 2003). In fact, 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) breaks down, 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 (M[c]Crae and Unger, 2014) for modeling 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. First we will start with the most fundamental distinction between *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, as shown in 13 and 3.

2. (a) Mary is a technical engineer.
   (b) *Mary is technical.

3. (a) The baby is awake.
   (b) *The awake baby.

One of the principle classifications of the semantics of adjectives (for example (Partee, 2003; Bouillon and Viegas, 1999; Morzycki, 2013b)) is based on the meaning of adjective noun compounds relative to the meaning of the words by themselves. This classification is as follows (where $\Rightarrow$ denotes entailment).

Intersective (X is a $A\ N \Rightarrow X$ is $A \wedge X$ 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 $Belgian \sqcap Violinist(X)$, and hence we can infer that a "Belgian violinist" is a subclass of a "Belgian".

Subsective (X is a $A\ N \Rightarrow X$ is a N, but X is a $A\ N \not\Rightarrow X$ is A) Such adjectives acquire their specific meaning in combination with the noun the modify. For example, a "skilful violinist" is certainly in the class $Violinist(X)$ but the described person is 'skilful as a violinist', but not skilful in general, e.g. as a surgeon.

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

Another important distinction is whether adjectives are *gradable*, i.e. whether a comparative or superlative statement with these adjectives makes sense. 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' (Morzycki, 2013a). An important group of gradable adjectives are intersective; we call such adjectives 'absolute' (following (Rusiecki, 1985)) as they refer to an ideal point on some scale. For example, a 'straight line' is 'straight' in that it has no bends or kinks. However, we can still talk about a line being 'straighter' in the sense of closer to the ideal of straightness 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' in lambda calculus as a function that takes a class C as input and returns the class of entities that were a member of C to some prior time point t (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.,

4. (a) A big red car.
   (b) $^{?}$A red big car.

5. (a) A famous former actor.
   (b) A former famous actor.

One further case that is important to distinguish is that of *relational* adjectives which have a meaning that expresses a relationship between two individuals or events, for example:

6. He is related to her.

7. She is similar to her brother.

8. This is useful for something.

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, which much like the standard noun predicate frame (X is a N) is stereotyped in English as:

$$X \text{ is } A$$

For attributive, usage we associate this with a frame, which is stereotyped as, where the $N?$ argument is not semantically bound, but instead obtained by syntactic unification with a noun predicate frame.

$$X \text{ is } A\ N?$$

As such, when we encounter the attributive usage of an adjective such as in 9, we understand this as the realization of two frames, given in 10.

9. Juan is a Spanish researcher.

10. (a) Juan is a researcher.
    (b) Juan is a Spanish ?.

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

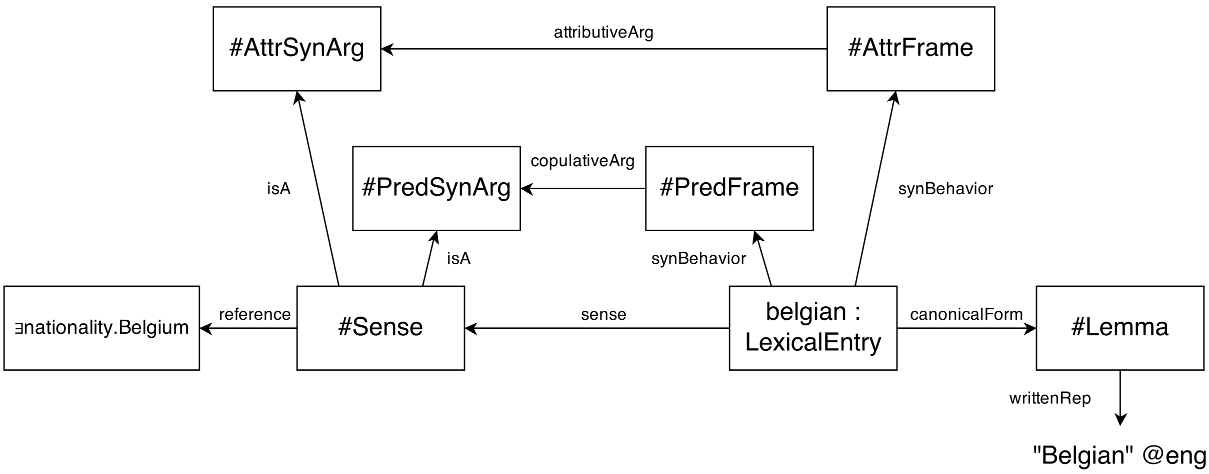


Figure 1: Modelling of an intersective adjective ‘Belgian’ in *lemon*

3.1 Intersective adjectives

Intersective adjectives are the most straightforward class as in many cases they can be modelled essentially as a noun or verb (e.g. deverbal adjectives such as ‘broken’). Intersective adjectives take one argument and can thus be modelled as unary predicates in FOL or classes in OWL as described by McCrae et al. (McCrae and Unger, 2014). For practical modeling 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)[1].

```
lexicon: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 ‘X is $A\ \ N$?’ and the *copulative argument* for the frame stereotype ‘X is A’. Furthermore, the class of Belgians is not named in our reference ontology DBpedia, so we introduce an anonymous class with the axiomatization, $\exists$ *nationality . Belgium*. It is in fact com-

[1] We assume that the namespaces are defined for the lexicon as `lexicon`, e.g., `http://www.example.org/lexicon` and for the entry, e.g., `belgian` is `http://www.example.org/lexicon/belgian#`, other namespaces are assumed to be as usual.

mon that the referent of an adjective is not named in an ontology and as such we tend to model denomial adjectives as classes of the form $\exists prop.Value$, 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 $\exists\, theme^{-1}.EventClass$, for example 'vandalized' may be $\exists\, theme^{-1}.VandalismEvent$.

3.2 Gradable adjectives and relevant observables

Gradable adjectives have a number of properties, which differentiate them from intersective adjectives:

- They occur in comparative constructions, in English with either '-er' or 'more' (Kennedy and McNally, 1999):3), e.g. 'smaller' and 'more frequent', as opposed to intersectives such as '*less geological' and '*more wooden'.

- **Gradable adjectives have a context-dependent truth-conditional variability, meaning that their positive form is the sum of the relation between the degree of the concept possessed by the object (as measured by the predicate) and the context-dependent standard of comparison based on the same concept (Kennedy, 2007). It follows that the properties denoted by adjectives like 'expensive' or 'small' or 'big' vary in intensity according to the context (and time) of use. This part was considered not to be clear by Philipp, he comments: "Examples would help here definitely, e.g. for what we mean with standard of comparison, intensity, sum of relation between the degree and context-depenedent standard etc. These notions are hard to grasp.". We could rephrase it by saying: Gradable adjectives have a context-dependent truth-conditional variability. This means that the value of the adjective depends (and is gradable) according to the object it qualifies (Kennedy, 2007). Adjectives such 'big', 'expensive' or 'small' are given different values according to the objects they are associated with.**

- **The arguments of gradable adjectives are mapped into abstract representations of measurements or degrees (Kennedy, 2007). This part was considered not clear by Philipp. Possible rephrase: It follows that gradable adjectives can be defined as 'scalar' or 'gradable', since their value can ideally be measured on a scale of set degrees, which changes with the object of the adjective ('big' owns a certain degree or value if combined with 'building' and another degree or value if combined with 'body').**

- They are frequently *fuzzy* (or *vague*) (Kennedy, 2007).

- There may be a minimum or maximum of this scale **(Which scale?)**, which can be determined by, for example, whether they can modified by 'completely'.

As such we define gradable adjectives relative to a particular property, these adjectives are also called 'observable' (Bennett, 2006)[2], 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.[3] 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, 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`).

[2]Note in many cases the property is quite abstract such as in 'breakable'.

[3]The use of these terms is borrowed from type systems, and resembles the concept of 'converse observables' as mentioned by ((Bennett, 2006):42). As stated by the author, adjectives often come in pairs of polar opposites (e. g. $conv(tall) = short$, and both refer to the same observable (in this case *size*. Some observables analogously hold converse relationships with other observables (e. g. $conv(flexibility) = rigidity$ or $conv(tallness) = shortness$).

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. 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 ‘Belgian’ or ‘not Belgian’, by who holds a passport of Belgium, it is not easy split the world’s population into ‘tall’ and ‘not tall’ (This is known as *sorites* paradox (Bennett, 2006)). In fact, while it may be easy to say that someone with height 6’6” (198cm) is ‘tall’, it is not clear whether someone of height 6’ (182cm) is ‘tall’, also given that this height should be compared to average (different) height for a man, or a woman. As such, one frequently used way to deal with this class of vague adjectives (and nouns) is via fuzzy logic (Goguen, 1969; Zadeh, 1975; Zadeh, 1965; Dubois and Prade, 1988; Bennett, 2006). **John, you put here the following citations: Goguen 1969, Zadeh 1965, Zadeh 1975, Dubois 1988. Since they had not been added in the bib file, the compiler generated a “?”, I have added them. Check whether these are the papers that you were referring to. For your indicated reference ‘dubois1988’ i found DuboisPrade88en** 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).[4] 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 tall for a female.

So we can conclude that gradable adjectives are *fuzzy*, *non-monotonic* and *context-sensitive*, all of which are incompatible with the description logic used by OWL.

3.2.1 Vagueness

Currently there are only limited models for representing fuzzy logic in the context of the web (Zhao and Boley, 2008). In order to capture the properties of gradable adjectives, we introduce a new model, which we name *lemonOILS* (The *lemon* Ontology for the Interpretation of Lexical Semantics) [5]. This ontology introduces three new classes:

- `CovariantScalar`, indicating that the adjective is covariant with its bound property
- `ContravariantScalar`, indicating that the adjective is contravariant with its bound property
- `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.

- `boundTo` indicates the property that a scalar refers to (e.g., ‘size’ for ‘big’)
- `threshold` specifies a sensible minimal value for which the adjective can be said to hold
- `degree` is 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`)

[4] `https://en.wikipedia.org/wiki/Human_height`
[5] `http://lemon-model.net/oils`

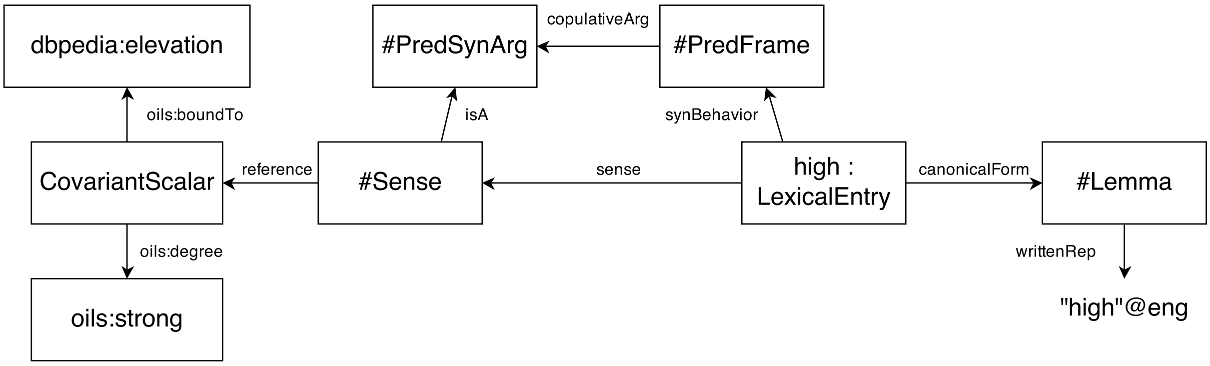


Figure 2: An example of the modelling of 'high' in *lemon*

- `measure` (TODO)
- `defaultValue` (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 (depicted in Figure 2).

```
lexicon: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 .
```

Philipp's comment3.2.1. it becomes now clear that Markov Logic is one way of implementing the oils-based description. Still, it would be nice to provide some alternatives, grounded in the DB-pedia data. Another way would be to construct a distribution of the top 5% (oils:strong) elements according to size for different classes and show the distributions graphically. Fuzzy sets would essentially be similar as they are essentially approximations of distributions without normalization I would say. One problem I see with the modelling in 3.2.1 of high we do not model the actual reference class, i.e. the noun. That's fine, but it is worth mentioning that the semantics (at least the degree and contravariance/variance) is independent of the specific noun (e.g. whether we talk about mountains or mice), but the threshold for instance and defaultValue are surely going to be specific for the noun/class modified. This needs to be pointed out IMHO.

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. For example, we can specify 'big' w.r.t. *size* for some class C as in 11.

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

In this way, the classification of an object into big or small can be defined as follows. For an individual $x \in C$ the property $big_C(x)$ holds if and only if:

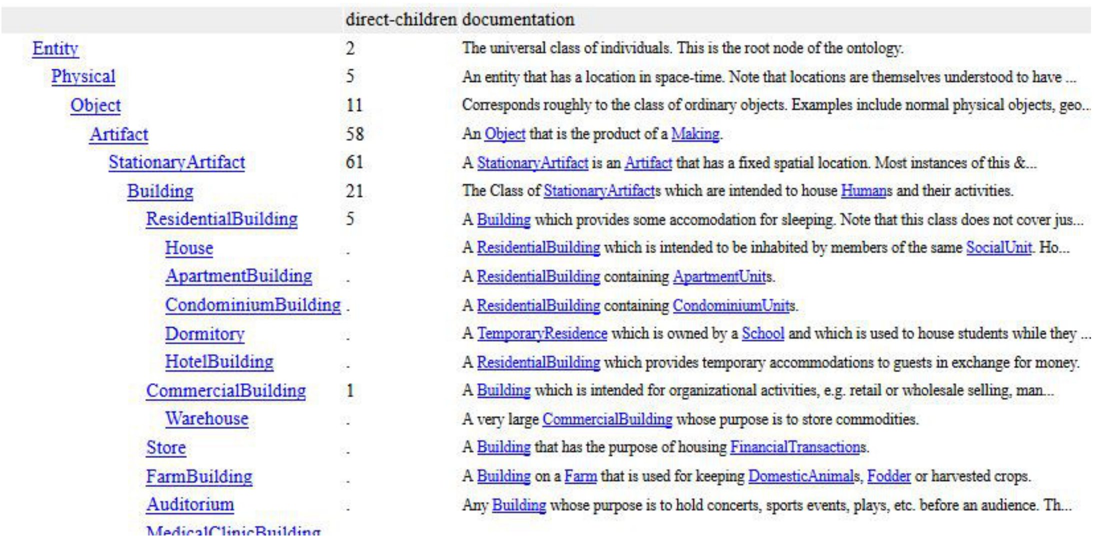
$$|\{y \in C, size(y) > size(x)\}|\alpha < |\{y \in C, size(y) < size(x)\}|\beta$$

Where the values of α and β are related to the degree defined in the ontology.

We see that 'big' defined in this way has the three properties outlined above: it is non-monotonic (in that more individuals may change whether we consider an individual to be 'big' or not), it is fuzzy (given by the strength of the probability of the proposition $big_C(x)$), and it is context-sensitive (as whether an individual counts a big or not depends on the class C).

3.2.2 The supervaluation theory and SUMO

Another way to capture the meaning of these vague terms can be achieved by *supervaluation semantics*. Through supervaluation theory, the modeling or positioning of *sorites* vague concepts is grounded in a judgement or meaning that lies on arbitrary thresholds, but these thresholds are based on a number of *relevant objective measures* (Bennett, 2006). The principle implies that the thresholds used for semantically interdependent concepts (let them be predicates or observables) are used consistently (e. g. the thresholds for 'tall/tallness' closely resemble the thresholds for 'short/shortness'). It also implies that not only the meaning of the gradables, but also the set value of the threshold is context-sensitive.



	direct-children	documentation
Entity	2	The universal class of individuals. This is the root node of the ontology.
Physical	5	An entity that has a location in space-time. Note that locations are themselves understood to have ...
Object	11	Corresponds roughly to the class of ordinary objects. Examples include normal physical objects, geo...
Artifact	58	An Object that is the product of a Making.
StationaryArtifact	61	A StationaryArtifact is an Artifact that has a fixed spatial location. Most instances of this &...
Building	21	The Class of StationaryArtifacts which are intended to house Humans and their activities.
ResidentialBuilding	5	A Building which provides some accomodation for sleeping. Note that this class does not cover jus...
House	.	A ResidentialBuilding which is intended to be inhabited by members of the same SocialUnit. Ho...
ApartmentBuilding	.	A ResidentialBuilding containing ApartmentUnits.
CondominiumBuilding	.	A ResidentialBuilding containing CondominiumUnits.
Dormitory	.	A TemporaryResidence which is owned by a School and which is used to house students while they ...
HotelBuilding	.	A ResidentialBuilding which provides temporary accommodations to guests in exchange for money.
CommercialBuilding	1	A Building which is intended for organizational activities, e.g. retail or wholesale selling, man...
Warehouse	.	A very large CommercialBuilding whose purpose is to store commodities.
Store	.	A Building that has the purpose of housing FinancialTransactions.
FarmBuilding	.	A Building on a Farm that is used for keeping DomesticAnimals, Fodder or harvested crops.
Auditorium	.	Any Building whose purpose is to hold concerts, sports events, plays, etc. before an audience. Th...
MedicalClinicBuilding		

Figure 3: Excerpt of the SUMO graph structure for `Building`, with number of direct children for each entry and related documentation

A recent extension of the SUMO ontology (Niles and Pease, 2001, Suggested Upper Merged Ontology) has add default measurements for the `Artifacts`, `Devices` and `Objects` enlisted in the ontology (currently amounting to 300+). The compilation of `defaultMeasurements` in SUMO has been just conducted on observables. Given for instance an `Artifact` such as `Book`, the compilation of its default measurements would look like this:

```
;;Book
(defaultMinimumHeight Book (MeasureFn 10 Inch))
(defaultMaximumHeight Book (MeasureFn 11 Inch))
(defaultMinimumLength Book (MeasureFn 5.5 Inch))
```

```
(defaultMaximumLength Book (MeasureFn 7 Inch))
(defaultMinimumWidth Book (MeasureFn 1.2 Inch))
(defaultMaximumWidth Book (MeasureFn 5.5 Inch))
```

The example for `Book` shows that the default measurements for the observable reflect a *standard* kind of book, i.e., one of the most commonly known kinds of the same artifact. As for the case of `Book`, SUMO implies it to be a physical object with a certain length, height and width. A weakness here is that the there is no systematic connection between the `defaultMinimumHeight` and `Height`, since **properties like this have solely been defined in SUMO in terms of first-order logic, not in terms of default measurements**. With *lemonOILS* we can add this information as follows:

```
sumo:Book oils:default [
    oils:defaultFor sumo:height ;
    oils:defaultMin "10in" ;
    oils:defaultMax "11in" ] .
```

Then if we understand a lexical entry 'high' as referring to a scalar covariant pseudo-class for `sumo:height`, it is possible to understand that a 'high' object exceeds the default minimum **set established for the same and owns at the same time a value for 'high' which does not go beyond the established default maximum. The decision to set an arbitrary threshold for objective measures to every specific kind of `Artifact` enlisted in SUMO presents advantages and disadvantages or bottlenecks. One obvious advantage is that SUMO poses itself, with this further extension on default physical measurements, as the first ontology of its kind to present defaults in correlation with its `Artifacts`. These approximated values, which ideally should represent the most commonly known form of each of the `Artifact`, can become handy in need of measurements or further information on the `Artifacts`. The approximated values also enable to draw a comparison between the children of the same parent. In this way, it might be possible, given for instance the defaults for `Bed`, to approximate the defaults for `KingBed`, `SingleBed`, `SofaBed`, `Futon` or `DoubleBed`, simply by adding or subtracting height, width, length and weight values with `Bed` as threshold. On the other side though, the SUMO default measurements serve no less and no more the purpose they were originally conceived for, namely to be a reliable (since double-checked), yet arbitrary, approximation of physical measures.**

A further weakness of this approach is captured by the following example:

12. Avery Johnson is a short basketball player.

Here, we see the difficulty in interpreting the sentence as Avery Johnson is in fact of average height (5'10") but for the class of basketball players he is unusually short. **While SUMO has some very specific listings of subsets for the same `Artifact`[6] for which it is then relatively simple to define maximum and minimum defaults for each for them and therefore consequently readjust the defaults for `Car` so to be the most prototypical possible, in some other places of its conceptual taxonomy, the ontology is still behind and needs further extension. For instance, with reference to this example of the basketball player, SUMO does not provide a well-structured subset net for `Person`, which is just enlisted as an 'instance for `Human`'. Further classifications of `Person` (e.g. in terms of gender or age) are still missing.As such we must introduce default values for every possible subclass of person. A way to solve this bottleneck could be to introduce, for every possible subclass of person, default values, as well as to introduce default values for the same Artifact in conjunction with a predicate or adjective. Nevertheless, this can potentially lead** to an explosion in the amount of information we need to encode in the ontology.

3.3 Operator adjectives

To clarify why we introduce operator and relational adjectives at this point, we might want to tell that these kinds of adjectives are not gradable, but they are usually related to their 'classificatory'

[6]For example, the subsets for `Car` are `CrewDormCar`, `GalleryCar`, `MotorRailcar`, `FreightCar`, `BoxCar`, `RefrigeratorCar`, `FiveWellStackCar`, `FlatCar`, `SpineCar`, `HydraCushionFreightCar` and more.

or ‘taxonomic’ meaning (McNally and Boleda, 2004), which should make them particularly suitable for being modeled in ontologies (although we present here some bottlenecks as well). Both operator and relational adjectives are characterized by a not-postnominal position (this is why i don’t think ‘known’ below is an ok example for relational adjective [we can in fact say ‘he is known’ out of ‘he is a known actor’]).

13. **Clinton is a male tall former president**
14. **$\models$ Clinton is president.**
15. **$\models$ Clinton is male (male $\Rightarrow$ intersective gradable adjective)**
16. **$\models$ Clinton is tall (tall $\Rightarrow$ subsective gradable adjective)**
17. **$\emptyset$ Clinton is former. (Since this example was the same as (2a,b) above, I’ve changed 2(a)(b) into ‘Mary is a technical engineer $\Rightarrow$ *Mary is technical)**

Operator adjectives are those that combine with a noun to alter the meaning of the noun itself. **Fran: not clear.** There are two primary issues with the understanding of the adjective in this manner. **Firstly, the reference of the lexical item does not directly refer to an existing item in the ontology, but rather is novel and productive. Secondly, the compositional nature of adjective-noun compounds is no longer simple as in the cases of intersective and gradable adjectives.** To this extent we claim that it is not generally possibly to represent the meaning of an operator adjective, within the context of an OWL ontology. Instead, following Bankston (Bankston, 2003), we claim that the reference of an operator adjective must be a higher order predicate. If we assume that there are operators of the form of a function, then the argument of a operator is the attributed noun phrase. As such we introduce a frame *operator attributive*, that has one argument which is the noun. The interpretation of a phrase like 18a is like 18b.

18. (a) Obama is the president. Clinton is a former president.
    (b) $President(Obama) \sqcap [former(President)](Clinton)$

If we understand *President* as `Artifact` or `Human` that has the role of leading a country (and therefore being true for *Obama* and *Clinton*), then *former* creates a new class in the ontology (i.e. `FormerPresident`) which selects people who were presidents. **Thus it can be said that $[\boldsymbol{former}(\boldsymbol{President})]$ creates a new class in the ontology. Just said.** An alternative interpretation would be that ‘former’ modifies the role of Clinton, however this raises the question of why the role property of presidents would be selected, and as such is also unsatisfactory. **In a way, we have already answered this statement by means of the grammatical definition that we have given for relational and operator adjectives, namely that they cannot take a postnominal position (unlike gradable adjectives). It follows that we can avoid discussing the case that ‘former’ modifies the role of Clinton, given that the statement ‘Clinton is former’ is *de facto* ungrammatical. Not only this, but the adjective ‘former’ (as other operator adjectives such as ‘alleged’ or ‘counterfeit’) is neither intersective, nor subsective, but privative (Partee, 2001), meaning that it *entails the negation of the noun property*. Therefore, to answer the question of why the role property of presidents would be selected, we can state that the operator adjective raised the role of the object, but dismissed it at the same time.** As there is at the moment no agreed representation for such an operator in an ontology, it underlines the fact that more sophisticated ontology representations than first-order logic are required to understood natural language in general.

Philipp’s original comment (prior to my change):On operational adjectives and the specific example ‘Clinton is a former president’. I have problems with this example, I see the class of president as containing a set of individuals representing a role played by someone beginning and ending at some time point. So former would be a class that out of these role individuals selects all those with ending point before now. This seems quite doable in FOL to me ;-)

3.4 Relational adjectives

Relational adjectives are among the simplest and modeled with another frame, which extends the attributive frame by allowing for a prepositional phrase adjunct.Fran: not very clear As such we can model 'known' with the frame 'X is known to Y' and reference `foaf:knows` as: **Another example, apart from 'known', (that you can use in postnominal position, and therefore does not really represent relational adjective) could be: 'Jemma is a single parent' $\Rightarrow$ 'Jemma is a parent', but $\emptyset$ 'Jemma is single'**

```
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 .
```

It is important to specify that relational adjectives can also have another definition in literature (Morzycki, 2013a). They represent a particular class of adjectives that denotes properties of kinds, rather than of simple nouns. In the case for instance of:

19. **Rose is a gastrointestinal surgeon**

'gastrointestinal' refers to the property of a kind, not an individual. Once again, the interpretation of the adjective is context-sensitive, as clearly showed in the following examples:

20. **John is an industrial engineer**

21. **This is an industrial gas turbine**

4 Adjectives in question answering

Of the 150 training and test questions of the QALD-4 benchmark[7] for question answering over linked data, 67 contain adjectives.

Some of the occurring adjectives do not have a semantic contribution, or at least none that is separable from the noun, as exemplified in the noun phrases in 22 and 23 (where $[\![\cdot]\!]$ stands for 'denotes' and the prefixes `dbo` and `res` abbreviate the DBpedia namespaces `http://dbpedia.org/ontology/` and `http://dbpedia.org/resource/`, respectively).

22. (a) $[\![$official website$]\!]$ = `dbo:website`
    (b) $[\![$national anthem$]\!]$ = `dbo:anthem`

23. (a) $[\![$official languages$]\!]$ = `dbo:officialLanguages`
    (b) $[\![$military conflicts$]\!]$ = `dbo:battle`

Otherwise, the most common kinds of adjectives among them are intersective and gradable adjectives.

All intersective adjectives denote restriction classes that are not explicitely named in DBpedia, in correspondence with the modelling proposed in Section 3.1 above, for example:

[7] `http://www.sc.cit-ec.uni-bielefeld.de/qald/`

24. (a) $[\![\text{Danish}]\!] = \exists\,$`dbo:country.res:Denmark`
    (b) $[\![\text{female}]\!] = \exists\,$`dbo:gender.res:Female`
    (c) $[\![\text{Methodist}]\!] = \exists\,$`dbo:religion.res:Methodism`

All gradable adjectives that occur in the QALD-4 question set can be captured in terms of *lemonOILS* as `CovariantScalar` (e.g. 'high') or `ContravariantScalar` (e.g. 'young'), cf. Section 3.2 above, bound to a DBpedia datatype property (e.g. `elevation` or `birthDate`). The positive form of those adjectives only occurs in 'how' questions, denoting the property they are bound to, for example:

25. (a) $[\![\text{deep}]\!] =$ `dbo:depth` in 'How deep is Lake Placid?'
    (b) $[\![\text{tall}]\!] =$ `dbo:height` in 'How tall is Michael Jordan?'

The comparative form denotes the property they are bound to together with an aggregation operation, usually a filter invoking a term of comparison that depends on whether the adjective is covariant or contravariant.

26. (a) $[\![\text{Which mountains are higher than the Nanga Parbat?}]\!] =$

```
SELECT DISTINCT ?uri WHERE {
 res:Nanga_Parbat dbo:elevation ?x .
 ?uri rdf:type dbo:Mountain .
 ?uri dbo:elevation ?y .
 FILTER (?y > ?x)
}
```

    (b) $[\![\text{Was the Cuban Missile Crisis earlier than the Bay of Pigs Invasion?}]\!] =$

```
ASK WHERE {
 res:Cuban_missile_crisis dbo:date ?x .
 res:Bay_of_Pigs_Invasion dbo:date ?y .
 FILTER (?x < ?y)
}
```

Finally, the superlative form denotes the property they are bound to together with an aggregation operation, usually an ordering with a cut-off of all results except the first one, as exemplified in 27. The only exception are two cases where the superlative has a semantic contribution that is not separable from the noun, see 28 (where the prefix `dbp` abbreviates the DBpedia namespace `http://dbpedia.org/property/`).

27. $[\![\text{What is the longest river?}]\!] =$

```
SELECT DISTINCT ?uri WHERE {
 ?uri rdf:type dbo:River .
 ?uri dbo:length ?l .
} ORDER BY DESC(?l) OFFSET 0 LIMIT 1
```

28. (a) $[\![\text{highest place}]\!] =$ `dbo:highestPlace`
    (b) $[\![\text{lowest rank in the FIFA World Ranking}]\!] =$ `dbp:fifaMin`

In addition, there are gradable adjectives that are bound to a property that constitutes the semantic contribution of the noun. Examples are 'frequent' and 'long'; the meaning of their superlative forms is captured in terms of SPARQL aggregation operations, as shown in 29 and **??**.

29. $[\![\text{What is the most frequent death cause?}]\!] =$

```
SELECT DISTINCT ?uri WHERE {
 ?x dbo:deathCause ?uri .
} ORDER BY DESC( COUNT(DISTINCT ?uri) )
  OFFSET 0 LIMIT 1
```

30. $[\![$What is the bridge with the longest span?$]\!]$ =

```
SELECT DISTINCT ?uri
WHERE {
        ?uri rdf:type dbo:Bridge .
        ?uri dbo:mainspan ?s .
}
ORDER BY DESC(?s)
OFFSET 0 LIMIT 1
```

There are some instances of operator adjectives, e.g. 'former', which is used in a context where it does not have a specific semantic contribution:

31. $[\![$the former Dutch queen Juliana$]\!]$ = `res:Juliana`

Another example is 'first', which has a context-dependent and highly ontology-specific meaning, often tightly interwoven with the meaning of the noun, as in the following examples:

32. (a) $[\![$first president of the United States$]\!]$ = $\exists$`dbo:office`.'1st President of the United States'
    (b) $[\![$first season$]\!]$ = $\exists$`dbo:seasonNumber`.1

Finally, there is only one instance of a non-empty subsective adjective, denoting a property:

33. $[\![$professional$]\!]$ = `dbo:occupation` (as opposed to, e.g., hobby),
    for example $[\![$professional surfer$]\!]$ = $\exists$`dbo:occupation.res:Surfing`

We miss some examples for relational adjectives. An example (as taken from DBpedia) could be: "Give me all cars that are produced in Germany" ('all' relates to 'produced in Germany'). This the only example I found in the list of questions proposed in one of the first drafts of this paper.

5 Related work

The categorization of adjectives in terms of formal semantics goes back to Montague(1970) and Vendler(1968), however one of the most significant attempts to assign a formal meaning was carried out in the Mikrokosmos project (Raskin and Nirenburg, 1995). This was one of the first works to treat the case of a micro-theory of adjectives, in which the results were "machine-tractable", in that they could be formally defined by a computer. The applications of this were limited however and no formal logic was attached to the semantic representations. Nevertheless much of the modeling resembles ours. In particular, scalar adjectives **in the Microkosmos project?** are modeled by association with an attribute and a range, e.g., 'big' is described as being `>0.75` (i.e., 75% of all known instances) on the `size-attribute`. These classifications do not however clearly separate meaning and syntax and also require a separate modeling of comparatives and class-specific meanings for many adjectives.

Amoia and Garden (2006) handled the problem of adjectives in the context of textual entailment. They analyzed 15 classes that show the subtle interaction between the semantic class (e.g., 'privative') and the issues of attributive/predicative use and gradability.

Abdullah and Frost (2005) tackles the privative nature of adjectives by arguing that the adjectives modify the set themselves, in a manner that is naturally second-order. Similarly, Partee (2003) proposed a limited second-order model by means of an 'head primary principle' requiring that adjectives are interpreted within their context, **here is not clear then what 'head' in the principle's name stands for; is**

it the head noun they are bound to?. Bankston's analysis (2003) however shows that the fundamental nature of many adjectives is higher-order, and provides a very sophisticated formal representation framework for this syntactic class. A more thorough discussion of non-gradable, non-interesective adjectives is given by Morzycki (2013a).

The Generative Lexicon (Pustejovsky, 1991) provides another approach to the representation of semantics, and the case of adjectives has also been considered in this context **in this context means in the semantics' context presented by Pustejovksy, right?**. **Bouillion (1999) consider the case of the French adjective 'vieux' ('old'), which he interprets as selecting two different elements in the event structure of an attributed noun, that is whether the state, e.g., being a 'mayor' for 'mayor', is considered old or the individual itself. In this way, the introduction of two senses for 'vieux' is avoided, however it remains unclear if such reasoning introduces more complexity than the the extra senses.More than Buillon's study (we could then also cite the German study of (Beck, 2000) or the Spanish study of (McNally and Boleda, 2004)) we could spend some more words on Pustejovsky. Here what I would suggest:**

In his analysis of adjectives, Larson (1998) suggests that many adjectives denote properties of *events*, rather than of simple heads or nouns (which does not fall very far from the statement, made above, that relational adjectives denote properties of kinds). Pustejovsky (1992; 1991) and (Lenci and others, 2000) state that lexical and semantic decomposition can be achieved *generatively*, assigning to each lexical item a specific *qualia structure*. For instance, in an expression like:

34. The round, heavy, wooden, inlaid magnifying glass

- **'round' represents the *Formal* role (giving indications of shape and dimensionality)**
- **'heavy' and 'wooden' is the *Constitutive* role and indicates the relation between the object and its parts (e. g. by specifying weight, material, parts and components)**
- **'inlaid' is the *Agentive* role of the lexical item, denoting the factors that have been involved in the generation of the objects, such as creator, artifact, natural kind, and causal chain**
- **'magnifying' describes the *Telic* role of 'glass', since it shows its purpose and function**

Peters and Peters (2000) provide one of the few other practical reports on modelling adjectives with ontologies, in the context of the SIMPLE lexica. This work is primarily focussed on the categorization of by means of intensional and extensional properties, rather than due to their logical modelling.

6 Conclusion

In this paper we have presented a method for modelling adjectives with the ontology-lexicon model, *lemon*. In particular, we found that adjectives frequently go beyond the first-order logic model used by OWL, but instead require models that are non-monotonic, fuzzy and second-order. As such, we conclude that more sophisticated semantic models are required to represent the semantics of such words. However the separation of syntax and semantics remains a robust model, which can easily be adapted to the task of representing adjectives. As a final note, we consider the claims that not all languages have adjectives [8] and as such we must wonder to what extent this analysis is applicable beyond English. We contend that the underlying semantics of the words we discuss here is representable in all nearly languages. Based on our analysis of realistic questions as applied in QALD, we furthermore believe that this model should be applicable to a range of domains and languages with little issue, even if further validation is naturally necessary.

Acknowledgements

References

Nabil Abdullah and Richard A Frost. 2005. Adjectives: A uniform semantic approach. In *Advances in Artificial Intelligence*, pages 330–341. Springer.

[8]http://linguistlist.org/issues/4/4-442.html

Marilisa Amoia and Claire Gardent. 2006. Adjective based inference. In *Proceedings of the Workshop KRAQ'06 on Knowledge and Reasoning for Language Processing*, pages 20–27. Association for Computational Linguistics.

Paul Bankston. 2003. Modeling nonintersective adjectives using operator logics. *The Review of Modern Logic*, 9(1-2):9–28.

Sigrid Beck. 2000. The semantics of different: Comparison operator and relational adjective. *Linguistics and Philosophy*, 23(2):101–139.

Brandon Bennett. 2006. A theory of vague adjectives grounded in relevant observables. In John Mylopoulos Patrick Doherty and Christopher A. Welty, editors, *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning*, pages 36–45. AAAI Press.

Pierrette Bouillon and Evelyne Viegas. 1999. The description of adjectives for natural language processing: Theoretical and applied perspectives. In *Proceedings of Description des Adjectifs pour les Traitements Informatiques. Traitement Automatique des Langues Naturelles*. Citeseer.

Pierrette Bouillon. 1999. The adjective "vieux": The point of view of "generative lexicon". In *Breadth and depth of semantic lexicons*, pages 147–166. Springer.

Paul Buitelaar, 2010. *Ontology-based Semantic Lexicons: Mapping between Terms and Object Descriptions*, pages 212–223. Cambridge University Press.

Philipp Cimiano, Paul Buitelaar, John M^{c}Crae, 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.

Frank Van Harmelen Deborah L. M^{c}Guinness et al. 2004. Owl web ontology language overview. *W3C recommendation*, 10(2004-03):10.

Didier Dubois and Henri Prade. 1988. *Possibility theory*. Plenum Press, New York.

Joseph H. Goguen. 1969. The logic of inexact concepts. *Synthese*, 19:325–373.

Christopher Kennedy and Louise McNally. 1999. Deriving the scalar structure of deverbal adjectives. *Catalan Working Papers in Linguistics*, 7:125–139.

Christopher Kennedy. 2007. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and philosophy*, 30:1–45.

Richard K. Larson. 1998. Events and modification in nominals. In Devon Strolovitch and Aaron Lawson, editors, *Proceedings from Semantics and Linguistic Theory (SALT) VIII*, pages 145–168. CLC Publications, Itaca, New York.

Alessandro Lenci et al. 2000. Simple work package 2, linguistic specifications, deliverable d2.1.

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.

Louise McNally and Gemma Boleda. 2004. Relational adjectives as properties of kinds. *Empirical issues in formal syntax and semantics*, 5:179–196.

Richard Montague. 1970. English as a formal language. In Bruno Visentini et al, editor, *Linguaggi nella societa e nella tecnica*, pages 189–224. Milan: Edizioni di Comunità.

Marcin Morzycki. 2013a. The lexical semantics of adjectives: More than just scales. Ms., Michigan State University. Draft of a chapter in *Modification*, a book in preparation for the Cambridge University Press series *Key Topics in Semantics and Pragmatics*.

Marcin Morzycki. 2013b. *Modification*. Cambridge University Press.

John P. M^{c}Crae and Christina Unger. 2014. Design patterns for the ontology-lexicon interface. In Paul Buitelaar and Philipp Cimiano, editors, *Towards the Multilingual Semantic Web: Principles, Methods and Applications*. 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.

Ian Niles and Adam Pease. 2001. Towards a standard upper ontology.

Barbara H. Partee. 2001. Privative adjectives: Subsective plus coercion. In *Studies in Presupposition*. Thomas E. Zimmermann.

Barbara H Partee. 2003. Are there privative adjectives. In *Conference on the Philosophy of Terry Parsons, University of Massachusetts, Amherst*.

Ivonne Peters and Wim Peters. 2000. The treatment of adjectives in simple: Theoretical observations. In *LREC*.

James Pustejovsky. 1991. The generative lexicon. *Computational linguistics*, 17(4):409–441.

James Pustejovsky. 1992. The syntax of event structure. In Bett Levin and Steven Pinker, editors, *Lexical & Conceptual Semantics*, pages 47–83. Oxford: Blackwell.

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.

Zeno Vendler. 1968. *Adjectives and nominalizations*. Number 5 in Papers on formal linguistics. Mouton.

Lofti A. Zadeh. 1965. Fuzzy sets. *Information and Control*, 8:338–353.

Lofti A. Zadeh. 1975. The concept of linguistic variable and its application to approximate reasoningi. *Information Sciences*, 8:199–249.

Jidi Zhao and Harold Boley. 2008. Uncertainty treatment in the rule interchange format: From encoding to extension. In *URSW*.