Modelling the Semantics of Adjectives in the Ontology-Lexicon Interface

John P. McCrae
Universität Bielefeld
Bielefeld
Germany
jmccrae@cit-ec.uni-bielefeld.de

Francesca Quattri
Hong Kong Polytechnic University
Hong Kong
francesca.quattri@connect.polyu.hk

Christina Unger
Universität Bielefeld
Bielefeld
Germany
cunger@cit-ec.uni-bielefeld.de

Philipp Cimiano
Universität Bielefeld
Bielefeld
Germany
cimiano@cit-ec.uni-bielefeld.de

Abstract

The modelling of the semantics of adjectives is notoriously challenging. We consider this problem 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 that are 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 application to ontology-based question answering.

1 Introduction

Ontology-lexicon models, such as *lemon* (Lexicon Model for Ontologies) (McCrae et al., 2012) model the semantics of open class words by capturing their semantics with respect to the semantic vocabulary defined in a given ontology. 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). As such, the modelling is based around a lexical entry which describes the morphology and syntax of a word, and is linked by means of a lexical sense to an ontology entity defined in a given ontology described in formal logic. 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 formalization of this question as an executable query formulated with respect to the SPARQL query language is provided in (1b). In particular, the interpretation of this question involves the formal interpretation of the word ‘high’ as relating to the property `dbo:elevation`, including ordering and subset selection operations.

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
```

In the above query, we select an entity denoted by the query variable `?uri` that has the properties that i) the entity’s type is a mountain, ii) it is located in Australia, and iii) it has an elevation bound to the variable `?elevation`. We then sort the query in descending order by the value of the elevation and limit so the query returns only the first result, in effect choosing the largest value in the data set. 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 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. First we will start with the most fundamental distinction between *attributive* and *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 (3a) and (3).

2. (a) Clinton is a former president.
   (b) *Clinton is former.
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 single words that form the compound. 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 class denoted by the noun and the class denoted by 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”. Furthermore, we could conclude that if the same person were a surgeon, he/she would also be a “Belgian Surgeon”.

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 that one individual is 'more former'. Most gradable adjectives are subsective (e. g.'a big mouse' is not 'a big animal' (Morzycki, 2013a)).

Finally, we consider *operator* or *property-modifying* adjectives. They can be understood along the lines of privative adjectives but differ in that they represent operators that modify some property in the qualia structure (Pustejovsky, 1991) of the class. For instance, 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.

Finally, we define *object-relational* adjectives as those adjectives which have a meaning that expresses a relationship between two individuals or events[1], 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 provided 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:

$$X \text{ is } A$$

The attributive usage of an adjective is associate to a stereotypical frame where the $N?$ argument is not semantically bound, but can instead be obtained by syntactic unification to 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 $N?$

Note that we do not provide modelling for adjectives where the meaning is unique for a particular noun phrase, such as 'polar bear', which we would capture as a normal noun phrase with meaning *ursus maritimus*.

[1] Our definition of relational here is borrowed from the idea of relational nouns (De Bruin and Scha, 1988) as a word that requires an argument. Our definition is also different from the one for 'relational adjectives' as proposed by (Morzycki, 2013a).

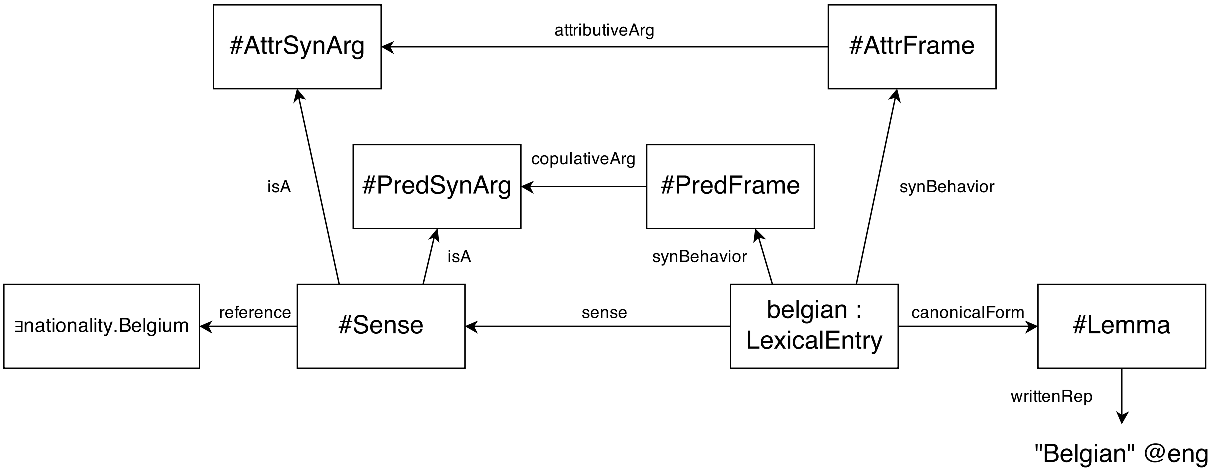

Figure 1: Modelling of an intersective adjective 'Belgian' in *lemon*

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 first-order logic or classes in OWL, as described by McCrae and Unger (2014). For practical modelling examples, we will use the *lemon* model, since 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 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)[2].

```
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 .
```

In this example, the word 'Belgian' is associated with a lemma with representation 'Belgian', two frame objects and a lexical sense. The frame objects describe the attributive and predicative usage, and are associated with an attributive and copulative argument respectively. The sense links the word to the anonymous ontological class for objects that have 'Belgium' as the value of their 'nationality' property and furthermore the arguments of each frame are linked to the sense in order to establish a correspondence between the ontology class and the syntactic frames. Note that here we use the external vocabulary defined in the LexInfo ontology (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

[2]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.

class of Belgians is not named in our reference ontology DBpedia, so we introduce an anonymous class with the axiomatization, i.e. $\exists\, \mathit{nationality}\,.\, \mathit{Belgium}$. It is in fact common that the referent of an adjective is not named in an ontology. An obvious choice is thus to model denominal adjectives as classes of the form $\exists \mathit{prop}.\mathit{Value}$, where *Value* is an individual that represents the semantics of the noun from which the adjective was derived. This modelling is so common that it has already been encoded as two design patterns, called `IntersectiveObjectPropertyAdjective` and `IntersectiveDatatypePropertyAdjective` (see (M^{c}Crae and Unger, 2014)). Similarly, most deverbal adjectives refer to an event, and as such a common modelling is of the form $\exists\, \mathit{theme}^{-1}.\mathit{EventClass}$. For example, 'vandalized' may be $\exists\, \mathit{theme}^{-1}.\mathit{VandalismEvent}$.

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), e.g. 'smaller' and 'more frequent', as opposed to intersectives such as '*less geological' and '*more wooden'.
- Gradable adjectives can be defined as 'scalar', since their value can ideally be measured on a scale of set degrees
- They have a context-dependent truth-conditional variability, meaning that their positive form is understood in relation to the class of the object modified by the adjective. For example, an 'expensive watch' has a different price scale to an 'expensive bottle of water'.
- They are frequently *fuzzy* (or *vague*) (Kennedy, 2007).
- There may be a minimum or maximum of the adjective's scale, which can be determined by, for example, whether they can modified by adverbs such as 'completely' or 'utterly'.

As such, we define gradable adjectives relative to a particular property. These adjectives are also called 'observable' (Bennett, 2006)[3] as they are related to some observable or measurable property, e.g. *size* in the case of *'big'*. However, a specification of the observable property is clearly not sufficient to differentiate between the meaning of antonyms such as *big* and *small.* Thus, we introduce the notions of *covariance* and *contravariance*, which specify whether the comparative form indicates a higher property value for the subject or the object. In this sense 'big' is covariant with size, as bigger things have a higher size value, and 'small' is contravariant with size.[4] We also introduce a third concept, i.e. the one of *absolute gradability*, which expresses the fact that the degree of membership in the denotation of the adjective is stronger the more it approaches a prototypical or 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`).

While these notions can handle the comparative structure of the semantics 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 like in the case of many intersective adjectives. In fact, while we can clearly define all people in the world as 'Belgian' or 'not Belgian', according to whom holds a Belgian passport or not, it is not easy to split the world's population into 'tall' and 'not tall' (This is known as *sorites* paradox (Bennett, 2006)). Furthermore, while it may be easy to say that someone with height 6'6" (198cm) is 'tall', it is not clear whether someone with height 6' (182cm) is 'tall', although compared to an average (different) height for a man, they are 'taller'. 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). Secondly, we notice 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 superlatives, as the discovery of a new tallest person

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

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

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)[5]; as such a male of 6' would have clearly been considered tall.

It follows from this that each instance added to our ontology might lead to a revision of the class boundaries of a gradable class, hence leading to the fact that gradable adjectives are fundamentally non-monotonic. We must also notice that gradability can only be understood relative to the class that we wish to grade. Thus, while it is a priori unclear whether 6' is tall for a male, it is clear that 6' is tall for a female given the current average height of a female being about 5'4" (162cm).

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

Pseudo-classes in lemonOILS

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)[6]. 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 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
- `absoluteValue` is the ideal value of an absolute scalar
- `degree` is specified as `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` indicates a unit that can be used as a measure for this adjective, e.g., 'John is 175 *centimetres* tall'.

Using such classes we can capture the semantics of gradable adjectives syntactically but not formally within an OWL model. As such, we call these introduced classes *pseudo-classes*. An example of modelling an adjective such as 'high' is given below (and 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 .
```

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

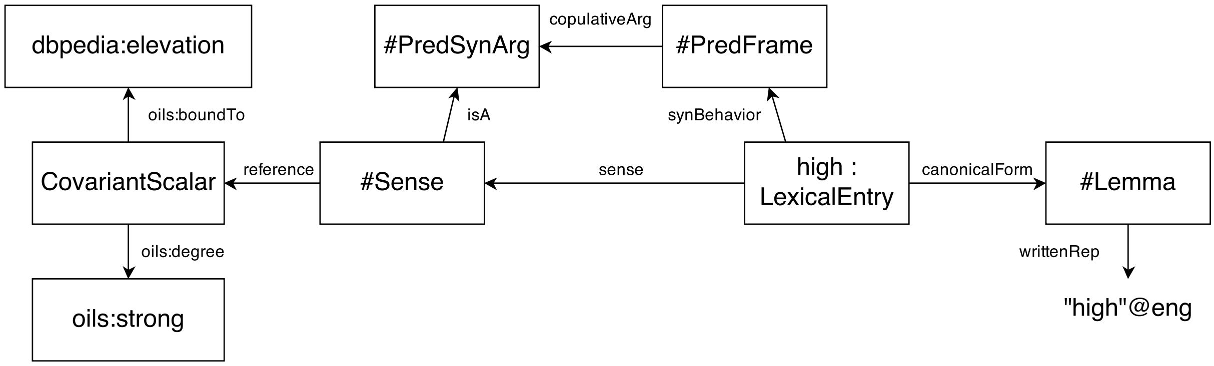

Figure 2: An example of the modelling of ‘high’ in *lemon*

As an example of a logic in which these annotations could be interpreted, 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’ with respect to *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 as big or not depends on the class C). Furthermore, our definition does not rely on defining ‘big’ for a given class, but instead is inferred from some known number of instances of this class. This eliminates the need to define a threshold for each individual class, or even to define the predicate big_C on a per-class basis.

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 modelling 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).

A recent extension of the SUMO ontology (Niles and Pease, 2001, Suggested Upper Merged Ontology)[7] includes default measurements (currently amounting to 300+) added to the `Artifacts`, `Devices` and `Objects` enlisted in the ontology (and marked with capitals). The compilation of `defaultMeasurements` in SUMO has been just conducted on observables, not on predicates. Given for instance an `Artifact` such as `Book`, the compilation of its default measurements would look like:

```
;;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))
```

[7] `www.ontologyportal.org`

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 this case, SUMO implies `Book` to be a physical object with a certain length, height and width (and possibly weight). A weakness here is that the there is no systematic connection between the `defaultMinimumHeight` and `Height` or `Width`, since these physical properties have been defined in SUMO just in terms of first-order logic, and have not been assigned default measurements yet. 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 object and owns at the same time a value for 'high' which does not go beyond the established default maximum. 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`[8], SUMO does not provide a well-structured subset net for e. g. `Person`.As a way to address this bottleneck, we could introduce default values for every subclass of `Person`, as well as to introduce default values for the same Artifact in conjunction with a predicate or adjective (e. g. `BigPerson`, `BulkyPerson`). The creation of such *ad hoc* subclasses is not feasible in general, as we would have to introduce a new class into the ontology for every combination of an adjective and a noun. On the other side though, the SUMO default measurements serve the purpose they were originally conceived for, namely to be an arbitrary, yet computable approximation of physical measures.

3.3 Operator adjectives

Operator adjectives are those that combine with a noun to modify the meaning of the noun itself. There are two primary issues with the understanding of the adjective in this manner. Firstly, the reference of the lexical item does not generally refer to an existing item in the ontology, but rather is novel and productive, in the sense that it generates a new class. Secondly, the compositional nature of adjective-noun compounds is no longer simple, as in the cases of intersective and gradable adjectives. This means that, in order to understand a concept such as a 'fake gun', we must first derive a class of `FakeGuns` from the class of `Guns`. Thus the modified noun phrase must be an argument of the operator adjective.

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 an operator is the attributed noun phrase. As such, we introduce a frame *operator attributive*, that has one argument which is the noun. Thus we understand that the interpretation of 'fake gun' is by means of an operator $fake$, which is a function that takes a class and produces a new class, i.e., $[fake(Gun)](X)$. Capturing such an operator lies beyond the expressivity of first-order logic. To fully capture the semantics of such an operator adjective, formalisms beyond first-order logic are thus clearly needed.

3.4 Object-relational adjectives

Object-relational adjectives are those that require a second argument, such as 'known', which can only be understood as being 'known' to some person, in comparison to 'famous'. Thus, the modelling of the relational adjective *known* is quite similar to the semantics of the corresponding verb *know*. It can be modelled for instance via the frame 'X is known to Y' and reference `foaf:knows` as:

[8]For example, some of the subsets `Car` are: `CrewDormCar`, `GalleryCar`, `MotorRailcar`, `FreightCar`, `BoxCar`, `RefrigeratorCar`, `FiveWellStackCar`, and more.

```
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

In this section we empirically analyze the adequacy of the modelling proposed in this paper with respect to the QALD-4[9] dataset, a shared dataset for Question Answering over Linked Data. The 250 training and test questions of the QALD-4 benchmark contain 76 adjectives in total (not counting adjectives in names such as ‘Mean Hamster Software’).

18 of the occurring adjectives do not have a semantic contribution w.r.t. the underlying DBpedia ontology, or at least none that is separable from the noun, as exemplified in the noun phrases in (13) and (14).[10]

13. (a) $[\![\text{official website}]\!] = \texttt{dbo:website}$
    (b) $[\![\text{national anthem}]\!] = \texttt{dbo:anthem}$
14. (a) $[\![\text{official languages}]\!] = \texttt{dbo:officialLanguages}$
    (b) $[\![\text{military conflicts}]\!] = \texttt{dbo:battle}$

Otherwise, the most common kinds of adjectives among them are gradable (27) and intersective (13) 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:

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

In some cases these intersectives have a context-dependent and highly ontology-specific meaning, often tightly interwoven with the meaning of the noun, as in the following examples:

16. (a) $[\![\text{first president of the United States}]\!] = \exists\,\texttt{dbo:office}.$ ‘1st President of the United States’
    (b) $[\![\text{first season}]\!] = \exists\,\texttt{dbo:seasonNumber}.1$

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 (much)’ questions, denoting the property they are bound to, for example:

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

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

[10] $[\![\cdot]\!]$ stands for ‘denotes’ and the prefixes `dbo` and `res` abbreviate the DBpedia namespaces `http://dbpedia.org/ontology/` and `http://dbpedia.org/resource/`, respectively.

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.

18. (a) ⟦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)
}
```

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 (19). In some cases, the superlative property is already encoded in the ontology, e.g., in the case of the property `dbo:highestPlace`.

19. ⟦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
```

There are three instances of operator adjectives. Examples are ‘former’, as in 20, which does not refer to an element in the DBpedia ontology but is instead a disambiguation clue in the given query, and ‘professional’, which refers to the property `dbo:occupation`, see 21.

20. ⟦the former Dutch queen Juliana⟧ = `res:Juliana`
21. ⟦professional surfer⟧ = ∃`dbo:occupation.res:Surfing`

Finally, there were 8 remaining adjectives totalling 15 occurrences, which do not correspond to meaning in an ontology, but instead are part of the discourse structure, each ‘same’, ‘other’.

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). The approach to adjective modelling in the Mikrokosmos provided one of the first computational implementations of a microtheory of adjective meaning. The modelling of adjectives presented in this paper is clearly inspired by the modelling of adjectives adopted in the Mikrokosmos project. 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`. Still, these classifications do not clearly separate meaning and syntax and also require a separate modelling of comparatives and class-specific meanings for many adjectives.

Amoia and Gardent (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) focused on the modelling of privative adjectives by arguing that these adjectives modify the underlying set itself in a manner that is naturally second-order. Similarly, Partee (2003) proposed a limited second-order model by means of the ‘head primary principle’ requiring that adjectives are interpreted within their context. 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 adjectives. A more thorough discussion of non-gradable, non-intersective adjectives is given by Morzycki (2013a). Bouillion and Viegas (1999) consider the case of the French adjective ‘vieux’ (‘old’), which they interprete 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 extra senses. 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 (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:

22. The round, heavy, wooden, inlaid magnifying glass
    - 'round' represents the *Formal* role (giving indications of shape and dimensionality)
    - 'heavy' and 'wooden' related to the *Constitutive* role and indicate 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

Finally, 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 proposed an approach to model the semantics of adjectives in the context of the lexicon-ontology interface with a focus on the ontology-lexicon model *lemon*. We have argued that the semantics of adjectives, in particular gradable and privative adjectives, is beyond what can be expressed in first-order logics, OWL in particular. Instead, capturing the semantics of such adjectives requires formalisms that are non-monotonic, second-order and can represent fuzzy concepts. We have proposed an extension of *lemon* by the *lemonOILS* vocabulary that adds 'syntactic sugar' that allows us to represent the semantics of adjectives in a way that abstracts from the actual representational formalism used. This work has been used in the construction of lexical resources to support a question answering system, and we found that this framework is sufficient to enable tractable computation of natural language to SPARQL mapping over at least a small but varied set of test questions used in the QALD evaluation task. Future work will show whether this model is scalable and applicable to most adjectives as well as domains and natural languages.

References

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

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.

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 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.

Jos De Bruin and Remko Scha. 1988. The interpretation of relational nouns. In *Proceedings of the 26th annual meeting on Association for Computational Linguistics*, pages 25–32. Association for Computational Linguistics.

Frank Van Harmelen Deborah L. McGuinness 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. McCrae 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. 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.

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*.