Modelling of Adjectives in the Ontology-Lexicon Interface

John P. McCrae — Affiliation / Address line 1, Affiliation / Address line 2, Affiliation / Address line 3, email@domain

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

Abstract

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

1 Introduction

Ontology-lexicon models, such as *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) (McGuinness et al., 2004). It has been shown that this principle known as *semantics by reference* (Buitelaar, 2010) is an effective model that can be used in tasks such as question answering (Unger and Cimiano, 2011) and natural language generation (Cimiano et al., 2013). In particular, its suitability to the task is driven by the fact that the application of this model to answering questions based on the DBpedia (Auer et al., 2007) knowledge base requires mostly understanding the nouns and the verbs of the sentence. However, as has been shown by the Question Answering over Linked Data (Lopez et al., 2013, QALD) tasks, there are many questions that can be asked over this database that require a deeper semantic understanding of the representation of language. For example, questions such as 1a require understanding of the semantics of 'high' in a manner that goes beyond the model of OWL based on classes, properties and individuals. The answer given in the QALD dataset for this question is 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) and similarly we reach the conclusion that this is due to the issues of semantically modelling adjectives. 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 (McCrae 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 of *attributive* versus *predicative* usage, that is the use of adjectives in noun phrases ("X is a A N") versus as objects of the copula ("X is A"). It should be noted that there are many adjectives for which only predicative or attributive usage is allowed, as shown in 2 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 words by themselves. This classification is as follows (where $\Rightarrow$ denotes entailment).

Intersective (X is a $A\ N \Rightarrow X$ is $A \cap X$ 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". Thus if we also knew that the "violinist" was a "surgeon" we could conclude they were 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 do not alter the meaning of the noun phrase itself, but only make sense with knowledge of the noun they refer to. For example, a "skilful violinist" is certainly in the class $Violinist(X)$, but if we knew that that person is a surgeon as well, we cannot conclude that the person is a skilful 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.

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

He is related to her.

Another important distinction to make with adjectives is whether they are *gradable*, in that whether it makes semantic senses to make a comparative or superlative statement with these adjectives. For example, adjectives such as 'big' or 'tall' can express relationships such as 'X is bigger than Y', however it is not possible to say one individual is 'more former'. Most gradable adjectives are subsective, for example 'a big mouse' is not 'a big animal'(Morzycki, 2013a). An important group of gradable adjectives are, however, intersective, and we call such adjectives 'absolute' (following (Rusiecki, 1985)) as they refer to an ideal point on some scale, for example a '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.

Another important distinction to make with adjectives, which is orthogonal to the above classification, is whether they are *gradable*, in that whether it makes semantic senses to make a comparative or superlative statement with these adjectives. For example, adjectives such as 'big' or 'tall' can express relationships such as 'X is bigger than Y', however it is not possible to say one individual is 'more former'. It should thus be noted that most gradable adjectives are mostly intersective, for example a 'big mouse' is not a 'big animal' (Morzycki, 2013b). An important group of gradable adjectives are, however, intersective, and we call such adjectives *absolute* (following (Rusiecki, 1985)) as they refer to an ideal point on some scale, such as 'straight', and frequently refer to an endpoint of a scale, which Kennedy (1999) calls a *trivial standard*.

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[1] is stereotyped in English as:

$$X \text{ is } A$$

[1] X is a N

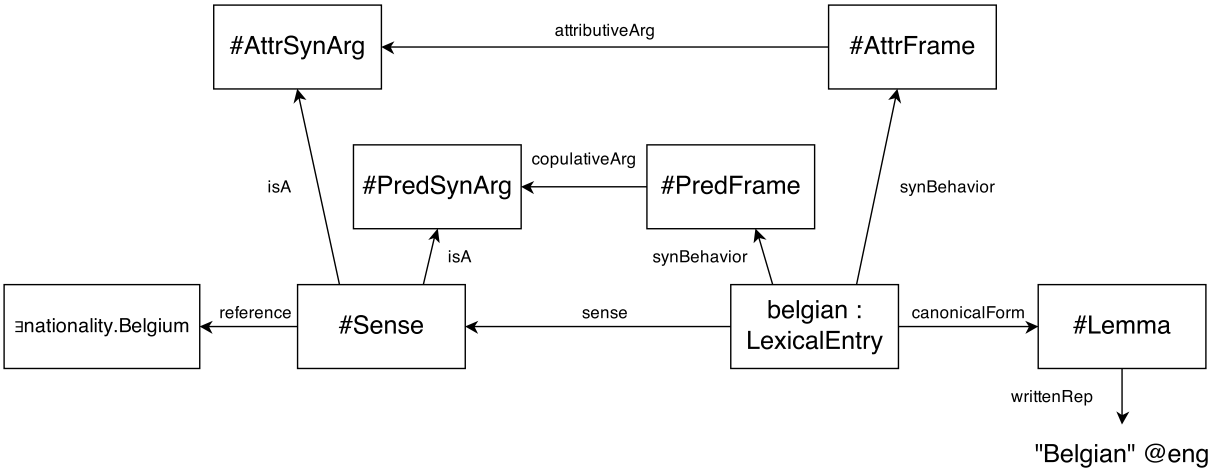

Figure 1: Modelling of an intersective adjective ‘Belgian’ in *lemon*

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*.

3.1 Intersective adjectives

Intersective adjectives are the most straightforward class as in many cases they can be explained as either being noun-like (denominal adjectives such as ‘Belgian’) or verb-like (deverbal adjectives such as ‘broken’). Intersective adjectives have a single argument as with most adjectives and in this case it is natural that they refer to classes, which may be event classes such as described in (McCrae and Unger, 2014). For practical 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)[2].

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

[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.

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

belgian:AttrFrame lexinfo:attributiveArg belgian:AttrSynArg .
belgian:PredFrame lexinfo:copulativeArg  belgian:PredSynArg .

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

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

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

- They have a comparative constructions with either '-er' or 'more' (Kennedy and McNally, 1999):3), c.f. '*less geological', '*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.
- They are frequently *fuzzy* (or *vague*) (Kennedy, 2007) (**more below**).
- The arguments of gradable adjectives are mapped into abstract representations of measurements or degrees (Kennedy, 2007).
- There may be a minimum or maximum of this 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, **hereby also called 'observable' (from (Bennett, 2006)).**[3], 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.[4] We

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

[4]**As clearly demonstrated in this example, the use of the terms 'covariant' and 'contravariant', although borrowed from mathematics and theoretical physics (covariance and contravariance of vectors), is not attributed the same original meaning the terms own in these two fields. The concept of covariance and contravariance we want to introduce here much more 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 $conv(tallness) = shortness$).**

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`).

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 kind of dilemma is know as the *sorites* paradox or *sorites* vagueness ((Bennett, 2006):36), for which many terms and adjectives in natural language are vague given the non-existence of a specific threshold to relate them to other properties or observables.** 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, the class boundary of a gradable adjective is naturally fuzzy. [Here i would say it: "As such, one recurrently used way to deal with this class of vague adjectives (and nouns) is via fuzzy logic", which enables a probabilistic approach to concepts (Goguen 1969; Zadeh 1965, 1975; Dubois & Prade 1988; in: (Bennett, 2006)).] 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).[5] 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 Fuzzy logic and 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 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

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

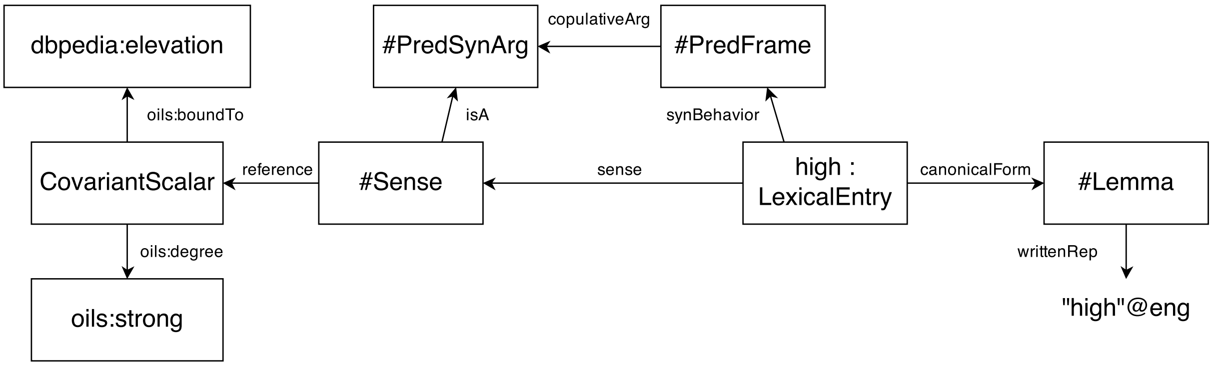

Figure 2: An example of the modelling of 'high' in *lemon*

- `comparator` indicates an object property that is equivalent to the comparison of the adjective (e.g., an object property `biggerThan` may be considered a comparator for the adjective class `big`)
- `measure` (TODO)
- `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 .
```

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

Despite the successful application of fuzzy logics in the case of lemonOILS, we acknowledge at least two bottlenecks in the use of this kind of logic for modeling vague or gradable adjectives.

- **the values which are assigned to the thresholds or degrees to which an object satisfies a vague predicate are assigned *ad hoc*, meaning that there is no objectivity in appointing them.**

- **Fuzzy logic is inferentially weak. Its fixed, domain-independent operators do not extensively work for uncertain information as in the case of vague predicates and observables, and could lead to unsatisfied or incorrect inferences ((Elkan, 1994):5); (Elkan, 1993). Given for instance a vague piece of information such as 'The mountain is far from the sea; and my house is beside the mountain" ((Bennett, 2006):36), one can automatically conclude 'My house is far from the sea', a kind of inference that fuzzy logic cannot provide. [John, it would be useful to try to represent this with fuzzy logic operators]**

Another way to measure 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):36). The principle implies (a) 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 (b) that not only the meaning of the gradables, but also the set value of the threshold is context-sensitive.

In the case of (as inspired by ((Kennedy and McNally, 1999):129)):

12. **Michael Jordan is tall.**

    ◇ **The threshold for 'height' can either be set (1) with reference to the average height for baseball players, or (2) with reference to the average height for Human, namely Person (where the capital starting letter of the word stands for ontological concept or KB term).**

13. **The building is tall.**

    ◇◇ **The threshold for 'height' can either be set with reference to the average height of a specific kind of Building, or with reference to a specific building.**

The possibilities mentioned in ◇(2) and ◇◇ have already been contemplated in one recent extension of the SUMO ontology[7], namely the addition of defaultMeasurements for the Artifacts, Devices and Objects enlisted in the ontology (currently amounting to 300+). The SUMO ontology can be browsed either via English terms (as derived from the Princeton WordNet ®), or KB terms, where the upper-level concepts have been converted in the formal first-order logic language SUO-KIF.

The compilation of defaultMeasurements in SUMO has been just conducted on observables, not predicates. This means that, given an observable such as Book, listed under `Artifact > Document`, only the defaults of a *standard* book have been measured, not those for `<predicate>`+Book.

[7]The Suggested Upper Merged Ontology (SUMO); (Niles and Pease, 2001); `www.ontologyportal.com`

	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**

```
;;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))
```

SUMO's pros

✓ bla bla

✓ bla bla

SUMO's cons

× **almost all children of respectively Artifacts, Objects and Devices in SUMO could be assigned default values ('Train', 'AircraftCarrier', 'Fish'), basically because selective (i. e. rather precise) count nouns. Nevertheless, default measurements for their parents could not be defined, given their extreme vagueness. While it is in fact still feasible to determine, among for instance all different kinds of 'Book', a standard volume (with implied arbitrary properties), the choice is to vast in the case of 'Artifact' or 'Device'.**

× bla bla

Thresholds, defaults and multiple classes (Francesca to help)

3.3 Operator adjectives

Operator adjectives are those that combine to alter the meaning of the adjective itself. There are two primary issues with the understanding of the adjective in this manner. Firstly, 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 then the attributed noun phrase, as such we introduce a frame *operator attributive*, that has one argument which is the noun. As such the interpretation of a phrase such as

Clinton is a former president.

Is understood as:

$$[former(President)](Clinton)$$

Where $former$ applies to $President$ to create a new class in the ontology. 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.

this kind of wusses out, but what else can we do?

3.4 Relational adjectives

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

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

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

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

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

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

4 Adjectives in question answering

Most common adjective kinds in QALD-4:

Intersective adjectives

- denoting a restriction class, e.g.
  - Danish films ($\exists$ `dbo:country.res:Denmark`)
  - female given names ($\exists$ `dbo:gender.res:Female`)
  - Methodist politicians ($\exists$ `dbo:religion.res:Methodism`)
- empty semantic contribution, e.g.
  - official website (`dbo:website`)
  - artistic movement (`dbo:movement`)
  - national anthem (`dbo:anthem`)
- non-separable semantic contribution (whole NP corresponds to class or property), e.g.

- American inventions (`yago:AmericanInventions`)
- official languages (`dbo:officialLanguages`)
- military conflicts (`dbo:battle`)

Subsective adjectives

- denoting a property
  - professional (`dbo:occupation`, as opposed to hobby)
    e.g. professional surfer = $\exists$ `dbo:occupation.res:Surfing`

Privative adjectives

- not treated
  - former Dutch queen Juliana = `res:Juliana`

Gradable adjectives

- positive form only occurs with `how`, denoting a property, e.g.
  - how deep (`dbo:depth`)
  - how heavy (`dbo:mass`)
  - how tall (`dbo:height`)
  - how high (`dbo:elevation`)
- comparative denotes a property plus aggregation, e.g.
  - higher than (`dbo:elevation` + `FILTER` with comparision operator that depends on polarity)
  - earlier than (`dbo:date` or `year` + `FILTER` with comparision operator that depends on polarity)
- superlative denotes a property plus aggregation, e.g.
  - the highest (`dbo:elevation` + `ORDER BY DESC(·) OFFSET 0 LIMIT 1`)
    the second highest (`dbo:elevation` + `ORDER BY DESC(·) OFFSET 1 LIMIT 2`)
    the highest after (`dbo:elevation` + `FILTER` + `ORDER BY DESC(·) OFFSET 0 LIMIT 1`)
  - the longest (`dbo:length` + `ORDER BY DESC(·) OFFSET 0 LIMIT 1`)
  - the youngest (`dbo:birthDate` + `ORDER BY DESC(·) OFFSET 0 LIMIT 1`)
  - the most frequent (ordering `COUNT`)
- superlative denotes an aggregation operation (whereas the property is contributed by the noun)
  - highest population density
  - lowest rank
  - longest span
- superlative has a non-separable contribution
  - the highest place (`dbo:highestPlace`)

Others

- temporal
  - first album (`releaseDate` + `ORDER BY ASC(·) OFFSET 0 LIMIT 1`)
  - first president ($\exists$ `dbo:office`.'1st President of the United States')
  - past two years (`year` + `FILTER`)
- first season ($\exists$ `dbo:seasonNumber`.1)
- alive (`deathDate` + `FILTER !BOUND`)

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 modelling resembles ours. In particular, scalar adjectives are modelled by association with an attribute and a range, e.g., 'big' was 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 as they also required a seperate modelling of comparatives and class-specific meanings for many adjectives.

Amoia and Garden (2006) handled the problem of adjectives in the context of textual entailment and 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 their 'head primary principle' requiring that adjectives are interpreted within their context. The analysis of Bankston (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.

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. 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.

(Morzycki, 2013b)

(McNally and Boleda, 2004)

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 fact that not all languages even have adjectives (**?**) 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 and based on our analysis of realistic questions as applied in QALD, we believe that this model should be applicable to a range of domains and languages with little issue, however 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.

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.

Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. Dbpedia: A nucleus for a web of open data. In *The semantic web*, pages 722–735. Springer.

Paul Bankston. 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 Patrick Doherty, John Mylopoulos, 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.

Charles Elkan. 1993. The paradoxical success of fuzzy logic. In *Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI-93)*, pages 698–703. MIT Press.

Charles Elkan. 1994. The paradoxical success of fuzzy logic. *IEEE Expert: Intelligent Systems and Their Applications*, 9(4):1–16.

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.

Vanessa Lopez, Christina Unger, Philipp Cimiano, and Enrico Motta. 2013. Evaluating question answering over linked data. *Web Semantics: Science, Services and Agents on the World Wide Web*, 21:3–13.

John P. McCrae and Christina Unger, 2014. *Design Patterns for the Ontology-Lexicon Interface*, chapter ?, page ??? Springer.

John McCrae, Guadalupe Aguado-de Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asunción Gómez-Pérez, Jorge Gracia, Laura Hollink, Elena Montiel-Ponsoda, Dennis Spohr, et al. 2012. Interchanging lexical resources on the semantic web. *Language Resources and Evaluation*, 46(4):701–719.

Deborah L McGuinness, Frank Van Harmelen, et al. 2004. Owl web ontology language overview. *W3C recommendation*, 10(2004-03):10.

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.

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.

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. Walter De Gruyter Inc.

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