

On the Role of Senses in the Ontology-Lexicon

Philipp Cimiano, John McCrae, Paul Buitelaar, Elena Montiel-Ponsoda

1 Introduction

Ontology-based natural language processing (NLP) applications interpret language with respect to the vocabulary of a given (domain) ontology. Take the example of an ontology-based question answering system [26] and the following input question: “*Who painted the Mona Lisa?*”. A query in SPARQL that represents the semantics of this question with respect to the DBpedia ontology [2] could look as in Figure 1.

The interpretation of linguistic input is a compositional process and requires knowledge about how the lexical atoms – words or phrases – which occur in a given domain are interpreted in the context of a given ontology. An important assumption that we build on in this paper is that the meaning of a word can not be specified universally, i.e. independently of any application or domain, but that the meaning of a lexical entry is specified with respect to the vocabulary defined by the given ontology. We refer to this principle as “*semantics by reference*”. The ontology-dependent meaning of a lexical entry (word or phrase) is captured in what we call an *ontology-lexicon*.

The principle of ‘semantics by reference’ implies that the expressivity and the granularity at which the meaning of words can be expressed is dependent on

Philipp Cimiano

Semantic Computing Group, CITEC, Universität Bielefeld, e-mail: cimiano@cit-ec.uni-bielefeld.de

John McCrae

Semantic Computing Group, CITEC, Universität Bielefeld, e-mail: jmccrae@cit-ec.uni-bielefeld.de

Paul Buitelaar

Unit for Natural Language Processing, DERI, National University of Ireland, Galway, e-mail: paul.buitelaar@deri.org

Elena Montiel-Ponsoda

Ontology Engineering Group, Universidad Politécnica de Madrid, e-mail: emontiel@delicias.dia.fi.upm.es

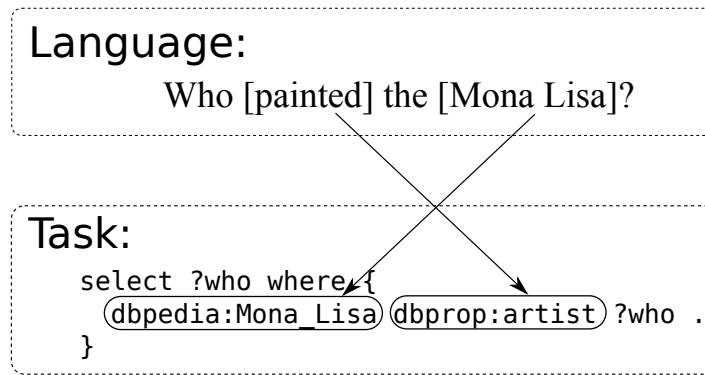


Fig. 1 Example of mapping of natural language into a task vocabulary

the meaning distinctions made in the ontology. Consequently, there might be aspects of the meaning of lexical entries that can simply not be represented with respect to the semantic vocabulary of a given ontology ([6], [7]). We will give two examples for this, a monolingual and a cross-lingual one. Consider the words ‘*mosque*’ and ‘*synagogue*’. In an ontology on urban planning, there might exist no specific concepts to represent the meaning of these words other than the more general concept `ReligiousBuilding`. Thus, with respect to this ontology, both ‘*mosque*’ and ‘*synagogue*’ will be interpreted as referring to the concept `ReligiousBuilding`. Clearly, this does not capture the ‘*full lexical meaning*’¹ of these words, but for the application and domain ontology in question the differences between a mosque and a synagogue might be irrelevant.

Let us now consider a cross-lingual example. Consider a geographic ontology which includes the concept of a `Watercourse`. With respect to this ontology, the full lexical meaning of the French words ‘*rivière*’ and ‘*fleuve*’ cannot be represented, as they encompass differing aspects that are not captured in this ontology. An ontology in our sense can be essentially seen as an artifact that represents a particular conceptualization of such a domain, limiting the representation of word meaning to those distinctions that are actually relevant in the context of the given ontology and/or domain. As such in the context of ontology-based NLP, it follows that the meaning of words is highly specific for a given ontology and that it should be possible to make principled choices concerning the number and granularity of senses that a word has as these senses may be said to correspond directly to explicit ontological distinctions made in the ontology. Overall, we thus need an ontology-lexicon that specifies how words and phrases occurring in the domain in question are interpreted with respect to the semantic vocabulary of the given ontology.

¹ By full lexical meaning we refer to the meaning that an average speaker of a language who shares common knowledge with his/her community associates with a word in their mental lexicon with respect to some language of thought.

A key question is how *senses* in an ontology-lexicon differ from those employed in traditional lexical resources such as WordNet. In fact, we could ask if explicit senses are needed at all in the context of the ontology-lexicon interface where ontology elements (classes, relations, individuals) essentially capture the meaning of words. We argue, however, that in the context of an ontology-lexicon, the role of a sense is to reify the link between a lexical entry and the concept it *evokes*, at the same time providing a hook into the ontology, thus allowing the larger context of the evoked concept to be exploited for interpretation. In addition, this reified link allows to model properties, including register as well as pragmatic constraints and implications related to the usage of the given work as evoking the concept in question.

Thus, we argue that in the context of the ontology-lexicon interface, a *sense* can be understood as a three-faceted entity. On the one hand, a sense can be understood as a reification of a pair of lexical entry and its corresponding reference in the ontology, i.e. the evoked concept so to speak. This is useful to state conditions under which it is possible to interpret the word as referring to that concept. Secondly, a sense can be regarded as a subset of all the uses of the given lexical entry that refer to the concept in question. Thirdly, a sense represents also a “hypothetical” concept that, if added to the ontology, would be a subclass of the evoked concept. This hypothetical concept accounts for the full lexical meaning of the word in question, but neither exists explicitly in the lexicon nor the ontology.²

The inclusion of explicit senses in the lexicon – as reifications of lexical entry/concept pairs – does not imply that the meanings of a word are fixed. In fact, through the interplay between the ontology as background knowledge and the given linguistic context (i.e. a specific sentence in which the word in question appears), further aspects of the meaning of a word can be brought to the foreground by a process that produces a semantic interpretation of the sentence. Lexical meanings in the ontology-lexicon can therefore be generated upon need, given the constraints of lexical context and semantic scope of the ontology. In this sense, an ontology thus supports a generative process in the sense of Wierzbicka [27] and Pustejovsky [25] by which further aspects of the meaning of a word can be derived from the ontology. Yet, these additional meaning aspects need not form part of the semantics of the word in the narrow sense, but are part of the larger ontological context and can be ‘recruited’ to support the interpretation of the word or phrase in a particular sentence. In line with this, our analysis allows us to provide a new account of the phenomenon of polysemy and metonymy in the context of the ontology-lexicon interface.

In this chapter, we will elaborate on these issues, providing a theoretical account of the role of senses in the ontology-lexicon interface. We will further analyse how aspects of the interpretation of words that have been traditionally subsumed under the phenomenon of polysemy. We argue that such aspects can be accounted for by a process that brings into the foreground further aspects of the meaning of words

² As our reviewer has pointed out, the hypothetical concept thus needs to exist in some ontology. In fact, it does, but not in the actual domain ontology, but rather in our ontology of the lexicon-ontology interface.

as a byproduct of the interpretation of a sentence given a specific linguistic context and the ontology as background knowledge. We will also briefly discuss the *lemon* model for representing an ontology-lexicon ([11], [22], [21]) and how the above theoretical considerations have influenced the design of this model.

2 Senses: universal or context-specific?

Traditionally, senses are regarded as specifying the various meanings of a word. Approaches differ essentially in whether they assume a finite and fixed amount of senses or postulate an open and highly context-specific set of senses. The specification of the set of senses (interpretations, meanings) for a given lexical entry is a central task in lexicography, involving decisions on whether to ‘lump’ potential senses together or to ‘split’ them into individual senses [15]. In practice this ends up being a very subjective task, in which lexicographers are guided by factors such as the purpose of the lexicographic resource, its envisioned users, the frequency of use of a certain meaning, or its predictability from other senses [19] and as such it has been questioned whether this is useful for NLP applications [17]. This view seems to be validated by the task of Word Sense Disambiguation (WSD) where it seems that many sense distinctions are not natural even for humans, as inter-annotator agreement for WSD seems to have a limit of about 80% [13]. A traditional approach to defining senses is by cross-lingual comparison [12]. The distinction between ‘*paper*’ (as a material) vs. a (news)-paper (as an information container) is for example inherent in many languages³ and one could thus argue that these two senses are language-independent or ‘natural’ (if not to say universal). Furthermore, it has been shown in the context of machine translation that this approach is helpful [8] and outperforms approaches based on fixed catalogues of senses such as found in WordNet. However, cross-lingual differences are not a solid basis to identify different word senses. Take the example of ‘computer’ with its two senses ‘a machine for performing calculations automatically’ and ‘an expert at calculation (or at operating calculating machines)’⁴. By cross-lingual comparison, we might identify these two different senses if in some language two different words are available for each of these two senses. If we compare German and English, however, we will find out that both languages use the same word for these two senses, i.e. *computer* in English and *Rechner* in German. Furthermore, it is often the case that languages make distinctions that are not considered fundamental to the native speaker of a language. German distinguishes for instance different types of ‘going’, using ‘*gehen*’ in the case of moving under one’s own power and ‘*fahren*’ when using mechanical assistance (e.g. a bicycle). This distinction seems unnatural to a native English speaker⁵. Following the tradition of linguistic relativism, it has been argued that concepts are

³ e.g., German “Papier” and “Zeitung”, French “papier” and “journal” and Japanese “kami” and “shimbun”

⁴ These are the glosses of the two corresponding senses from WordNet 3.1.

⁵ At least from the opinion of the native English-speaking author of this paper

language-specific, counterfeiting the assumption that cross-linguistic comparison can help us to establish a universal set of senses. It thus seems that relying on cross-lingual commonalities and differences as a basis to build a catalogue of senses will lead to extreme fragmentation and to overly specific senses that are not relevant in the context of a given application.

It has been indeed argued by some researchers that a small set of senses per word might suffice for practical applications. Ide and Wilks [18] for instance propose that the meaning of *'paper'* can be captured with only three senses: i) as a material, ii) as a written article or document and iii) as a newspaper. They argue that other senses such as *'publisher of the newspaper'* identified in some dictionaries are unnecessary and can in fact be derived from background knowledge (see Section 4 on this). While distinguishing such core senses of a word might be enough for general purposes, for certain domains we might need a much more fine-granular and domain-specific set of senses. Leon Araúz et al. [20] have for example argued that in their application, they need to distinguish three different senses of *"accretion"*, as i) accretion of snow flakes in the atmosphere, ii) the accretion of ground in a tectonic plate and iii) the accretion of sand in the formation of coastal bars. While these three senses have a similar basic meaning, that of accumulation of materials, it is necessary in this domain to distinguish them. However, such fine-grained distinctions will certainly not be included in a domain-independent lexicon such as WordNet, for reasons that should be clear.

At the other extreme are approaches that claim that any approach postulating a finite set of senses is unsatisfactory from a theoretical point of view. This view is connected to the assumption that there are as many senses as there are different contexts in which a specific word is used and is rooted in and supported by insights from philosophy and linguistic study. Wittgenstein famously claims that *"for a large class of cases - though not for all - ... the meaning of a word is its use in the language"* [29]. Cruse states that the meaning of a word form is different in every distinct context in which it occurs [9]. Later on, Cruse and Croft even maintain that word senses are created *at the moment of use*, in what they consider a dynamic approach [10].

While theoretically appealing, approaches which assume an infinite inventory of senses – one for each usage context – are less useful from an NLP point of view as automatic processing requires an inventory of senses that generalizes beyond specific examples and contexts observed. From an NLP perspective it is crucial to have i) either a finite set of senses, or ii) a specification of the core meaning of a word together with a set of generative processes that allow to derive new meanings from this core. The latter is essentially the underlying idea of the Generative Lexicon of Pustejovsky [25]. According to Pustejovsky, lexical items have a semantic representation of their conventionally assumed meaning, which is accessed in language understanding and production and can produce context-specific interpretations (senses) due to certain constraints that activate one sense or the other. An essential aspect of this theory is that lexical meaning is not decomposed into individual senses, but instead that different context-specific interpretations (roughly corresponding to senses) are activated on demand out of an underspecified lexical semantic representation that Pustejovsky refers to as *Qualia Structure*.

Our standpoint is that there is no universal set of senses for a word that will suit all purposes and applications. In some applications, it might suffice to interpret both ‘*synagogue*’ and ‘*mosque*’ as `ReligiousBuilding`. Other domains might require very specific senses as in the case of ‘*accretion*’ discussed above. In other domains, some of the senses that are generally distinguished might not be relevant. For instance, in the domain of scientific publishing, the material sense of ‘*paper*’ might not be relevant. Overall, it is thus clear that there is no set of universal senses that are valid independently of the domain and application. It is thus legitimate to assume that the senses of a word are specified with respect to a given ontology that models the domain in question. The meanings of a word are thus ontology-specific and need to be specified in an ontology-lexicon for the given ontology.

3 Senses in the Ontology-Lexicon Interface

As stated in the introduction, we regard a sense as a three-faceted entity with three roles. We elaborate on these roles in this section. In the following, we will assume that there is a given ontology $O = (\Lambda_O, V_O)$ expressed in logical language Λ_O and vocabulary V_O consisting of a set of concepts C_O , a set of relations R_O and a set of individuals I_O as well as a lexicon L .

In this paper we focus on tasks where natural language needs to be interpreted with respect to a given task and ontology such as the question answering task as illustrated in Figure 1. Thus, we consider that for a given task we need to find a mapping to an ontology-based representation, which we consider to be a formula in some task language \mathcal{T} which uses the symbols of the vocabulary and the language of the ontology as well as some additional task-specific symbols. The interpretation of natural language with respect to task language \mathcal{T} is given by the following function $[\cdot]_T$ ⁶:

$$[\cdot]_T : \mathcal{L} \rightarrow \mathcal{T}$$

where \mathcal{L} is the natural language in question (for the purposes of this definition we assume all words of relevance to the task are contained in the lexicon, and as such that the sentence can be considered as a sequence of lexical entries).

3.1 Senses as reification

We denote a *sense* as $\sigma^{(l,c)}$, where $l \in L$ is a lexical entry and $c \in V_O$ is the ontological concept, or *reference* and we define the *ontologically interpretable words* in the sentence λ as $W_\lambda \subseteq \{l_i : l \text{ is the } i^{\text{th}} \text{ word of } \lambda\}$. Furthermore, we define the *meaning of a word in sentence* λ as a function

⁶ Without loss of generality we simplify to the case where there is only a single result for the task

$$[\cdot]_T^\lambda : W_\lambda \rightarrow V_O$$

And we assume that this function satisfies the *compositionality principle* given as:

$$\forall l_i \in W_\lambda : [l_i]_T^\lambda = c \Rightarrow c \in [\lambda]_T$$

This means that if the lexical entry l_i is interpreted as c in sentence λ , then c should be part of the interpretation of λ with respect to task T .

Finally, we can define a sense, $\sigma^{(l,c)}$ to be *valid* with respect to a meaning function if the following holds:

$$\exists i \in \mathbb{N}, \lambda \in \mathcal{L} : l_i \in W_\lambda \wedge [l_i]_T^\lambda = c$$

With this definition, we consider the sense to be a reified pair capturing the cases under which it is valid to interpret l as having meaning c , where we understand validity to mean that the sense is used in at least one interpretation for the given task.

Consider the question in our introduction: Who painted the Mona Lisa? In this case: $\lambda = \textit{Who painted the Mona Lisa}$ and:

$$\begin{aligned} [\textit{painted}]_T^\lambda &= \text{dbprop:artist} \\ [\textit{Mona Lisa}]_T^\lambda &= \text{dbpedia:Mona.Lisa} \end{aligned}$$

Hence we have the following valid senses for our example sentence:

$$\begin{aligned} \sigma^{\textit{paint}, \text{dbprop:artist}} \\ \sigma^{\textit{Mona Lisa}, \text{dbpedia:Mona.Lisa}} \end{aligned}$$

In this role, we can understand a sense as the ‘glue’ between a pair of lexical entry and ontology concept, and also as the container for those pragmatic features (usage, register, etc.) whose role is neither purely ontological nor lexical.

3.2 Sense as subset of uses

As we have already argued above, the semantic distinctions made in the ontology provide a principled basis for defining a partition of the uses of a certain lexical entry. We define U – the usage set of a lexical entry l – as

$$U(l) = \{(l_i, \lambda) : l_i \in W_\lambda, \lambda \in \mathcal{L}\}$$

For each sense $\sigma^{(l,c)}$ we assume that the usage set u_i^c is defined as:

$$u_i^c = \{(l_i, \lambda) : l_i \in W_\lambda, \lambda \in \mathcal{L}, [l_i]_T^\lambda = c\}$$

For a set of senses Σ , we denote by $\Sigma|_l$ all senses whose lexical entry is l and we say that a set of senses is *complete* for a lexical entry l if its usage sets for a task T

satisfy:

$$U(l) = \bigcup_{\sigma^{(l,c)} \in \Sigma_l} u_i^c$$

The existence of a sense linking the lexical entry l and the concept c implies that the lexical entry can be used with this meaning, which is supported by at least one interpretation in the context of the given task. It then follows that a complete set of senses for a lexical entry constitutes a (non-disjoint) partition of all the uses ($U(l)$). One facet of a sense is that it can be specified as a subset of uses. By specifying which ontological distinctions are relevant in a given domain, the ontology thus provides a principled criterion to define the senses or meanings of a lexical entry in relation to the given ontology. Thus, the sense represents a subset of the uses of the lexical entry l for which l can be understood as meaning concept c .

3.3 Sense as a subconcept

The sense can be also understood as an implicit concept that captures further aspects of the meaning of the lexical entry that can not be captured by the ontology. In the following, we try to formalize this idea, borrowing several notions from Guarino [14], including the notion of a conceptualization and an ontological commitment.

Definition: A conceptualization is a triple $D = (D, W, R)$ with D a universe of discourse, W a set of possible worlds and R a set of conceptual relations on the domain space $\langle D, W \rangle$, where a conceptual relation ρ on $\langle D, W \rangle$ is a function $\rho : W \rightarrow D^*$ from the set W into D^* , the set of all n -ary (extensional) relations on D .

We define the lexical extension of a lexical entry l as a mapping from worlds to its extension, i.e.

$$lex(l) : W \rightarrow D^*$$

We now consider an ontology $O = (\Lambda_O, V_O)$ where the vocabulary can be further divided into $V_O = \{I_O \cup C_O \cup R_O\}$ consisting of a set of individuals I_O , a set of concepts C_O and a set of relations R_O . In line with Guarino we also consider an ontological commitment K for an ontology O and conceptualization $D = (D, W, R)$ as a pair $K = (O, I)$ where I is a function $I : V \rightarrow D \cup R$, i.e. I is an interpretation function that *interprets* the vocabulary of the ontology with respect to the vocabulary of the conceptualization.

We shall now assume that $c \in C_O$ and for each sense $\sigma^{(l,c)}$ define its *ontological projection* as follows:

$$\pi_c^l = lex(l) \cap I(c)$$

Here, $lex(l)$ and $I(c)$ are functions from possible worlds to D^* ⁷. From this it follows that:

⁷ We denote the intersection of two functions f and g here as $(f \cap g)(w) = f(w) \cap g(w)$

$$\forall w: \pi_c^l(w) \subseteq I(c)(w)$$

If we then add to our ontology a concept $c_{\pi_c^l}$ and extend the ontological commitment such that:

$$I(c_{\pi_c^l}) = \pi_c^l$$

then we get that:

$$O \models_M \forall x (c_{\pi_c^l}(x) \Rightarrow c(x))$$

where M is an *intended model* in the sense of Guarino [14]. For the intended models of the ontology, it follows that $c_{\pi_c^l}$ is thus a subclass of c , which we represent as $c_{\pi_c^l} \sqsubseteq c$ following the description logics notation. This result is similar for the case that $c \in R_O$, in that we derive a similar $c_{\pi_c^l}$ that is a sub-property of c and for $c \in I_O$ we obtain that $c_{\pi_c^l} = c$ (as there is no sub-division of an individual, naturally). These projected or hypothetical concepts $c_{\pi_c^l}$ thus represent the full lexical meaning of entry l when interpreted as concept c .

3.4 The three facets

As such, we have defined the sense in terms of three closely related entities or roles:

1. $\sigma^{(l,c)}$: The pair representing a correspondence between a lexical entry and vocabulary item in the ontology that it can be interpreted as in a given task.
2. u_l^c : The set of uses of a particular lexical entry l when used as referring to c .
3. π_c^l : a hypothetical concept representing the full lexical meaning of l when interpreted as c .

4 Systematic polysemy in the ontology-lexicon interface

In the previous sections we have argued that there is no universal set of senses for a word, but that the meanings of a word are specific for a given task and domain as described by a given ontology. We think that this is compatible with the view of Ide and Wilks [18] who propose using a small set of senses that suit a given task. Ide and Wilks for example propose to use three senses for ‘*paper*’: i) material, ii) daily newspaper and iii) article, thus excluding the sense of ‘*publisher of a newspaper*’. However, this raises the question how we would interpret the following sentence:

The paper was sued by the Workers Revolutionary Party.

A reasonable interpretation for this sentence might look as follows:

$$\exists x \text{ sued}(\text{WorkersRevolutionaryParty}', x) \wedge \text{NewspaperPublisher}(x)$$

Now, if *paper* does not have the sense of `NewspaperPublisher`, where does then the part of the above formula – `NewspaperPublisher(x)` come from? The meaning of *paper* in the above sentence might be derived as a result of a generative process which brings to the foreground the knowledge that a *newspaper* always has a publisher which publishes it, thus yielding the following meaning:

$$\exists x \exists y \text{ sued}(\text{WorkersRevolutionaryParty}', y) \wedge \text{Newspaper}(x) \wedge \text{publisher}(x, y)$$

The meaning of *paper* in the above sentence can thus be approximated by $\lambda y \exists x \text{ Newspaper}(x) \wedge \text{publisher}(x, y)$. Now where does this meaning come from? Following Hobbs et al. [16] we understand that this can be obtained by a principle of abduction in that we can introduce the property `publisher` as we have axioms within the ontology which state that the range of the `sued` property must be the of a type `IndividualOrOrganization`, which is a disjoint class with `Newspaper`. Furthermore, we suppose that the property `publisher` has domain and range of `Publication` and `Organization` which are super/sub classes of `Newspaper` and `IndividualOrOrganization` respectively. Hence, from an ontological point of view, the above interpretation is thus plausible and can thus be deduced abductively by a system performing the interpretation task.

In the linguistics literature, different examples of such a systematic relation between different meanings of a word has been studied under the label of *systematic polysemy*. Examples of classes of words the senses of which are systematically related are the following (see [1] or [24])⁸:

- animal/meat (The lamb is running in the field vs. John had lamb for dinner)
- plant/food (Mary watered the fig in the garden vs. Mary ate a fig)
- producer/product (The newspaper fired its editor vs. John spilled coffee over the newspaper)
- institution/building (The university became established in the early medieval ages. vs. The university is close to the capitol)

As an example of the institution/building class, consider the following uses of ‘*school*’ based on [4]:

- “Daddy drove me to school this morning”. \Rightarrow Daddy drove me to the location of the school building this morning.
- “They painted the school over the holidays” \Rightarrow They painted the walls of the building which hosts the school.
- “The school was built in 1950.” \Rightarrow The building which hosts the school (as institution) was built in 1950.
- “The school decided to fire the teacher.” \Rightarrow The executive board of the school (as institution) decided to fire the teacher.

⁸ Buitelaar [5] gives an overview of many systematic polysemy classes derived from WordNet 1.6

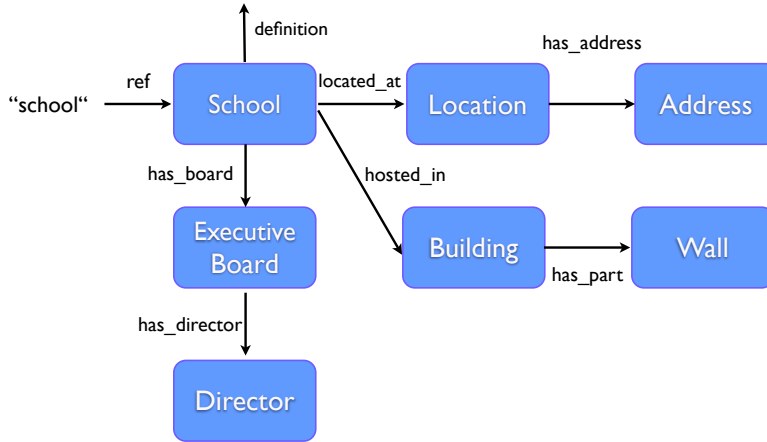


Fig. 2 Concept of ‘school’ in the ontology

Logically, the meanings of the above sentences can be expressed as follows⁹:

- “Daddy drove me to school this morning”: $\exists x, l, s \text{ drive}(daddy, x, l) \wedge speaker(x) \wedge location(l) \wedge located_at(s, l) \wedge educational_institution(s)$
- “They painted the school over the holidays”: $\exists p, b, s \text{ paint}(they, p) \wedge has_part(b, p) \wedge building(b) \wedge hosted_in(s, b) \wedge educational_institution(s)$
- “The school was built in 1950.”: $\exists s, e, b \text{ built}(e, b) \wedge happensAt(e, t) \wedge year(t, 1950) \wedge building(b) \wedge hosted_in(s, b) \wedge educational_institution(s)$
- “The school decided to fire the teacher.”: $\exists b, t, s \text{ fire}(b, t) \wedge teacher(t) \wedge has_board(s, b) \wedge educational_institution(s)$

The above paraphrases suggest that there are (at least) the following different senses of school: i) address where the school building is located, ii) building which hosts the school as institution, iii) walls of the building in which the school (as institution) is located, iv) executive board of the school (as institution). Thus, the meanings of ‘school’ in the above sentences could be formalized as follows:

1. $\lambda l \text{ location}(l) \wedge locatedAt(s, l) \wedge educational_institution(s)$
2. $\lambda p \text{ has_part}(b, p) \wedge building(b) \wedge hosted_in(s, b) \wedge educational_institution(s)$
3. $\lambda b \text{ building}(b) \wedge hosted_in(s, b) \wedge educational_institution(s)$
4. $\lambda b \text{ has_board}(s, b) \wedge educational_institution(s)$

Now, would we include all of the above senses in a lexicon? Definitely not, for good reasons. In fact, it seems that all of the above mentioned meanings are related in the ontology to the concept *educational_institution* and can thus be generated when interpreting the corresponding sentences by some process of abduction that

⁹ For the sake of simplicity, we do not represent the temporal adverbials and we model definites through existential quantifiers.

exploits the ontological neighbourhood of *educational_institution* to bring to the foreground additional – systematically related – meaning aspects as required by the linguistic context.

In fact, we argue that *educational_institution* is the primary meaning of ‘school’ and that the different linguistic contexts above select one particular or related aspect of the primary meaning of school, emphasizing the building in which it is located, the executive board, the activities that are typically offered at school, etc. If all of the aspects that are relevant for a school are modelled within an ontology (as sketched in Figure 2), most of the above systematically related concepts can be derived through appropriate coercion operations that traverse the ontology to find an entity that is related and that fits the linguistic and semantic context of the sentence in question.

Such a viewpoint is still in agreement with our definition of sense given in section 3 as it follows that the senses are still complete as although not every usage of a word is directly interpreted as *educational_institution*, every interpretation uses the symbol *educational_institution* and as such this sense is complete for the examples above. It is in this way that we argue that systematic polysemy is not a phenomenon that needs to be modeled in the context of the ontology-lexicon interface but exclusively at the ontological level.¹⁰

As such, from the perspective of the lexicon-ontology interface, the role of a sense is essentially to provide a hook into the ontology graph, reifying the fact that a word evokes a certain concept in the ontology. In our case, we merely need a link between the word ‘school’ and the concept *educational_institution*.

Thus, from the point of view of ontologies we can thus distinguish two types of lexical ambiguities:

- **Systematic Polysemy:** This case corresponds to the case where a word has different meanings that are systematically related through an ontology. In this case, the different interpretations of the lexical entry can be obtained by a process of abduction based on the axioms present in the ontology.
- **Homonymy:** According to our understanding, homonymy refers to the case where the different meanings of a word are not related through axioms in the ontology, so that two different senses are indeed required.

A consequence, however, is that whether a lexical ambiguity is systematic or irregular depends ultimately on the specific ontology. From a NLP perspective, we think that this represents indeed a principled approach, reducing the number of senses of a word to a minimum while being able to generate systematically related meanings exploiting background knowledge.

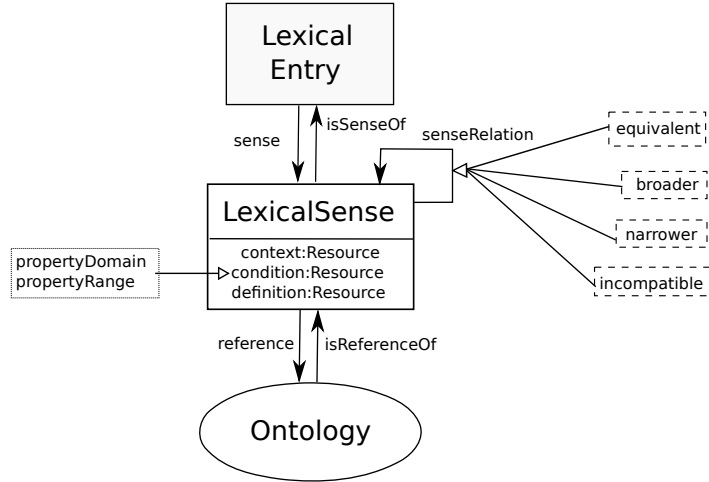


Fig. 3 The modelling of senses within the *lemon* ontology-lexicon model.

5 Application of senses in the ontology-lexicon model *lemon*

In earlier work [22], we have proposed *lemon* as a model for representing and sharing ontology lexica using Semantic Web standards such as OWL and RDF. In *lemon*, senses as considered in this paper are implemented through the class `LexicalSense`, which reifies the link between a `LexicalEntry` and some entity in the ontology (see Figure 3). This link is established via two properties (*reference* and *sense*) which are specified as functional properties, such that it can be inferred that the sense is unique to the pair (l, c) as in the definition.¹¹

As a simple example of the representation of senses in *lemon* we consider the case of translation. In most cases, a certain lexical entry l_1 is not a translation of some other lexical entry l_2 in all contexts. Rather, the “translation property” is dependent on the meaning of lexical entry l_1 . Thus, it might be that l_2 is only a translation of l_1 when l_1 is interpreted as concept c . Translation is thus a multi-valued function $trans : L_O \times V_O \rightarrow \mathcal{P}(L_O \times V_O)$, i.e. defined on pairs of lexical entry and concept. For example, the German word ‘*Krebs*’ is translated into English as ‘*cancer*’ when referring to the illness and as ‘*crab*’ when referring to the animal. In this example, it holds that $(cancer, illness) \in trans((Krebs, illness))$ but $(crab, illness) \notin trans((Krebs, illness))$. Instead, it holds that $(cancer, illness) \in trans((Krebs, illness))$.

¹⁰ Of course, we admit that for a certain application, specifying some of the senses that are derived from the primary one may improve performance.

¹¹ Technically, as we follow an open world assumption, if a sense is stated as reference to ontology entities or lexical entries, these entities or entries are considered to be exactly equivalent (pseudonymous)

The property of being a translation is thus a property between senses and not between lexical entries. This is formally stated in *lemon* as follows:

```
:Krebs lemon:sense [ lemon:reference dbpedia:Cancer ] ;
  lemon:sense [ lemon:reference dbpedia:Crab ] .

:Cancer lemon:sense [ lemon:reference dbpedia:Cancer ] .

:Cangrejo lemon:sense [ lemon:reference dbpedia:Crab ] .
```

As a sense is a reified pair of the lexical entry and ontology concept, we do not require an explicit translation link to capture the *trans* function defined above (although the schema does not preclude the inclusion of such a link if desired).

5.1 Sense properties

lemon allows additional properties to be attached to sense objects that can be used to describe aspects related to the usage of this lexical entry. In this sense, the reification of sense is crucial to express i) certain pragmatic implications of using a certain lexical entry to refer to the concept in question and ii) to state conditions under which it is legitimate to interpret the word as referring to the concept. An example of this might be the subjective or emotional associations that a certain language and culture makes when using a certain lexical entry to refer to a concept, i.e., connotations. Consider the noun *'retardation'*, which was earlier used to refer to people with learning and developmental difficulties. However, this use is considered extremely pejorative in modern usage¹². However, according to the current, 4th Edition of 'Diagnostic and Statistical Manual of Mental Disorders' (DSM-IV)¹³, there is a pathology called "Mental Retardation." The senses that we might distinguish for *'retardation'* are the following ones:

- A meaning of developmentally disabled that is used primarily in texts before 1970s.
- A meaning of unintelligent that is pejorative
- A reference to DSM-IV disorders 317, 318 or 319.

The associated aspects of the usage of the word *'retardation'* can neither be attached to the lexical entry itself nor the corresponding classes in an ontology as they describe the pair of lexical entry and concept and in particular constrain i) in which cases and under which conditions the lexical entry can refer to the concept in question, but also ii) what connotations the use of this lexical entry has when referring to the concept. In the case of *'retardation'* as *'developmentally disabled'* we would

¹² As a corollary many charities have changed their original name, for example the "Association for the Help of Retarded Children" is now just the "AHRC", which officially is not an intialism.

¹³ Standard international reference for mental health disorders

attach to the pair the information that this interpretation was mainly valid before the 1970s; this is neither a property of the lexical entry itself nor of the actual concept `DevelopmentallyDisabled`. With respect to the fact that ‘retardation’ is considered offensive in some contexts, this is neither a property of the lexical entry, as the word can be used in a non-offensive manner, nor of the actual concept, as the concept can be expressed in a non-offensive manner, but of the lexical entry when referring to the concept. A reification of the pair of lexical entry and concept is thus needed to express the pragmatic connotations that the word has when interpreted as a certain concept. We model this in *lemon* as follows:

```

:Retardation_entry
  lemon:sense [
    lemon:reference dbpedia:Developmental_Disorder ;
    lexinfo:dating lexinfo:old
  ] ;
  lemon:sense [
    lemon:reference dbpedia:Stupidity ;
    lexinfo:register lexinfo:perjorative
  ] ;
  lemon:sense [
    lemon:reference dsmiv:317
  ] .

```

As can be seen the *lemon* model requires an explicit sense object in its graph as otherwise there would be no sensible place to attach the properties required.

5.2 Contexts and conditions

In order to specify contextual conditions and constraints under which it is legitimate to interpret a lexical entry as referring to a given concept, *lemon* allows to model such contextual conditions using two properties: `context` and `condition`. The property `context` constrains the domains under which the interpretation of the lexical entry as the concept in question is permissible. For example, for the case of ‘*retardation*’ discussed above, the interpretation as referring to a disorder from the DSM-IV is valid in the medical domain. Two further properties called `dating` and `register` are subproperties of `context` and allow to constrain the time (e.g. before 1970) or register (e.g. informal, colloquial, ...) as conditions under which the lexical entry can be interpreted as referring to the concept in question.

The second property `condition` is used to state an evaluable expression describing the circumstances that need to be fulfilled such that the lexical entry can be interpreted as the ontological concept in question. The property is abstract and specific properties instantiating it need to be defined. The *lemon* model has two built-in subproperties of `condition`: `propertyDomain` and `propertyRange`. They restrict the usage of the lexical entry, requiring that the domain or range of the on-

tological property is of a specify type. For example, we could model that the verb ‘essen’, when interpreted as eat, requires the eater to be human, while ‘fressen’, when interpreted as eat, requires the eater to be an animal.¹⁴

The semantics of ‘fressen’ and ‘essen’ are thus modelled in lemon as follows:

```
:essen a lemon:LexicalEntry;
  lemon:canonicalForm [ lemon:writtenRep "essen"@de];
  lemon:synBehaviour [ a lexinfo:TransitiveFrame;
    lemon:subject :essen_subj;
    lemon:object :essen_obj];
  isocat:partOfSpeech lexinfo:Verb;
  lemon:sense [ lemon:reference myOntology:eat;
    lemon:subjOfProp :essen_subj;
    lemon:objOfProp :essen_obj;
    lemon:propertyDomain myOntology:Human].

:fressen a lemon:LexicalEntry;
  lemon:canonicalForm [ lemon:writtenRep "fressen"@de];
  lemon:synBehaviour [ a lexinfo:TransitiveFrame;
    lemon:subject :fressen_subj;
    lemon:object :fressen_obj];
  isocat:partOfSpeech lexinfo:Verb;
  lemon:sense [ lemon:reference myOntology:eat;
    lemon:subjOfProp :essen_subj;
    lemon:objOfProp :essen_obj;
    lemon:propertyDomain myOntology:Animal].
```

While *lemon* provides these two built-in properties, many other properties that model contextual conditions are possible. However, these need to be introduced by taking into account specific tasks and have thus not been included in the general model.

5.3 Sense relations

lemon also has properties for the representation of relationships between senses that are defined based on the facets as defined above. In particular the properties are defined as follows based on the usages u_i^c and the projection $\pi_c^!$.

- **equivalent**: The usages of the two senses are equal and the projections are equal.
- **broader**: The usages of the first sense is a superset of the second sense’s usage and projections are similarly a superset.
- **narrower**: The usages of the first sense is a subset of the second sense’s usage and projections are similarly a subset.

¹⁴ *lemon* actually allows to model the corresponding (subcategorization) frames of these verbs and their mapping to ontological properties. This aspect of the model is however not discussed in the present paper. The interested reader is referred to the *lemon* cookbook [21].

- `incompatible`: The usages of the two senses are disjoint and the projections are disjoint.

These properties have a very different status compared to the properties in the ontology. The properties in the ontology are defined between concepts, while the properties considered here are defined between senses as three-faceted entities introduced in this paper. These sense relations thus model (lexical) meaning aspects that are not included in the ontology but might nevertheless be important to model for a number of reasons. For example, one might be able to establish relations between different ontologies with different conceptualizations at the sense level if they are difficult to align at the conceptual level (e.g. because they vary substantially in granularity and modelling detail). Consider the property of antonymy for instance. Antonymy is typically a property between words that is not to be confused with the disjointness property between concepts used in many ontology languages. *lemon* introduces the property `antonymy` at the sense level as a subproperty of `incompatible`. The only ontological consequence is that the two projections π_c^l of the senses are regarded as being (ontologically) disjoint.

We can also use the *lemon* properties to capture the relationships between particular interpretations of lexical entries. For example, we consider the example of the French words ‘*rivière*’ and ‘*fleuve*’, which may be mapped to an ontology that only contains a concept corresponding to the English word ‘*river*’, while still ensuring that the terms are considered as not interchangeable. This can be achieved in *lemon* by mapping both words to the same ontology class but indicating that they are incompatible

```
:Riviere lemon:sense :Riviere_sense .
:Riviere_sense lemon:reference dbpedia:River .

:Fleuve lemon:sense :Fleuve_sense .
:Fleuve_sense lemon:reference dbpedia:River .

:Riviere_sense lemon:incomptable :Fleuve_sense .
:Fleuve_sense lemon:incomptable :Riviere_sense .
```

While sense relations described above do strictly speaking affect neither the actual ontology nor its conceptualization, they are crucial for NLP applications. Take the example of natural language generation and assume that we want to describe a given `river` in French. We might choose to generate the lexicalization ‘*rivière*’, but then we should remain consistent and not refer to the same river as ‘*fleuve*’. A NLP system thus needs to know that both senses are incompatible.

From a more general perspective, sense relations allow to represent cultural and linguistic differences in terminology and meaning granularity to be encoded in the lexicon.

6 Conclusions

In this paper, we have revisited the notion of *sense* in the context of the ontology-lexicon interface and argued that the senses that a word has are specific to a task and domain as modelled by a given ontology. Following a principle we call *semantics by reference*, the goal of an ontology lexicon is to define the meaning of a lexical entry relative to the meaning distinctions made in a given ontology. We have argued that in the context of the ontology-lexicon interface, the intrinsic subjectiveness of the answer to the question of which and how many senses a certain word has, can be overcome in a principled way by resorting to the meaning distinctions in the ontology. We have then discussed whether, under this assumption, senses are still meaningful entities in the context of the ontology-lexicon interface. We argue that the notion of sense is necessary in the context of ontology-lexicon interface and that in this context senses can be understood as a three-faceted entity that has the following roles: firstly, a sense can be understood as a reification of a pair of lexical entry and its corresponding reference in the ontology (concept). This is useful to state conditions under which it is permissible to interpret the word as referring to that concept. Second, the senses represent a set of disambiguated uses of an entry when used as referring to a certain concept in a given interpretation task. Third and finally, a sense represents also a “hypothetical” concept that, if added to the ontology, would be a subclass of the reference. We have further discussed what implications this has for the traditional notion of systematic polysemy, arguing that this is a phenomenon that should be resolved by means of abduction on the axioms in the ontology instead of by recording them in the lexicon. From this perspective, the role of a sense is to provide a hook to a concept in the ontology, providing an access route to other (systematically related) aspects of the meaning of a word. This hook can be exploited in the process of interpretation of a sentence in order to bring additional meaning aspects into the foreground as required by the linguistic context and to yield a well-defined interpretation. Finally, we have provided formal definitions of what it means for a sense to exist in the ontology-lexicon as well as details of how this understanding is implemented in the ontology-lexicon model *lemon*.

References

1. Apresjan, J.: Regular polysemy. *Linguistics*, 142:5–32. (1974)
2. Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R. and Ives, Z.: Dbpedia: A nucleus for a web of open data. *Semantic Web*. 4825:722–735, Springer. (2007)
3. Berlin, B., Kay, P. *Basic Color Terms: Their Universality and Evolution* (1969), University of California Press
4. Bierwisch, M.: Semantische und konzeptuelle Repraesentation lexikalischer Einheiten. In R. Ruzicka and W. Motsch (eds) *Untersuchungen zur Semantik*. Berlin. (1982)
5. Buitelaar, P.: *CoreLex: Systematic Polysemy and Underspecification*. PhD Thesis, Brandeis University. (1998)

6. Buitelaar, P.: Semantic lexicons: Between terminology and ontology. In: Simov K., Kiryakov A. (eds.) *Ontologies and Lexical Knowledge Bases. Proceedings of the First International OntoLex Workshop*. Sofia: OntoText Lab (2000)
7. Buitelaar, P.: Ontology-based Semantic Lexicons: Mapping between Terms and Object Descriptions. In: Huang C-R., Calzolari N., Gangemi A., Oltramari A., Lenci A., Prevot L. (eds.) *Ontology and the Lexicon: A Natural Language Processing Perspective. Cambridge Studies in Natural Language Processing*. CUP (2010)
8. Carpuat, M. and Wu, D.: Improving statistical machine translation using word sense disambiguation. In: *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)* pp. 61–72. (2007)
9. Cruse, D. A.: *Lexical Semantics*. Cambridge University Press. (1986)
10. Croft, W. and Cruse, D. A.: *Cognitive Linguistics*. Cambridge University Press. (2004)
11. Cimiano P., Buitelaar P., McCrae J., Sintek M.: LexInfo: A declarative model for the lexicon-ontology interface. *Journal of Web Semantics* 9(1), 29–51 (2011)
12. Dagan, I. and Itai, B.: Word sense disambiguation using a second language monolingual corpus. *Computational Linguistics* 20(4), pp 563–596. (1994)
13. Edmonds, P. and Kilgariff, A.: Introduction to the special issues on evaluating word sense disambiguation systems. *Journal of Natural Language Engineering* 8(4) pp 279–291. (2002)
14. Guarino N., Oberle D., Staab S.: In: Staab S., Studer R. (eds.) *Handbook on Ontologies, International Handbook on Information Systems*, chapter What is an Ontology?, pp. 1-17 (2009)
15. Hanks, P.: Do Word Meanings Exist?. *Computers and the Humanities* 34(2) pp 205-215. (2000)
16. Hobbs, J.R., Stickel, M.E., Appelt, D.E. and Martin, P.: Interpretation as abduction. *Artificial Intelligence* 63(1):69–142, Elsevier. (1993)
17. Ide, N. and Véronis, J.: Mapping Dictionaries: A spreading activation approach. In *Proceedings of the 6th Annual Conference of the Centre for the New Oxford English Dictionary*, pp. 52–64. (1990)
18. Ide, N. and Wilks, Y.: Making sense about sense. In: *Word Sense Disambiguation* pp 47–73. (2006)
19. Kilgariff, Adam: “I don’t believe in word senses”. *Computers and the Humanities*, 31(2), 91–113. (1997)
20. Leon Araúz, P., Faber, P., Magaña Redondo, P. J.: Linking Domain-Specific Knowledge to Encyclopedic Knowledge: an Initial Approach to Linked Data. In: *Proceedings of the 2nd International Workshop on the Multilingual Semantic Web*, pp 68–73. (2011)
21. McCrae, J. et al.: *The Lemon Cookbook*, <http://lexinfo.net/lemon-cookbook.pdf>. (2010)
22. McCrae J., Aguado-de-Cea G., Buitelaar P., Cimiano P., Declerck T., Gomez-Perez A., Gracia J., Hollink L., Montiel-Ponsoda E., Spohr D., Wunner T.: *Interchanging lexical resources on the Semantic Web. Language Resources and Evaluation* (accepted for publication)
23. Miller, G. A.: WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39. (1995)
24. Nunberg, G.: The non-uniqueness of semantic solutions: polysemy. In: *Linguistics and Philosophy*, 3.1. (1979)
25. Pustejovsky, J.: *The Generative Lexicon*. MIT Press, Cambridge, MA (1995)
26. Christina Unger, Philipp Cimiano: Pythia: Compositional Meaning Construction for Ontology-Based Question Answering on the Semantic Web. *NLDB 2011*: 153-160
27. Wierzbicka, A.: Semantic primitives and lexical universals. *Quaderni di Semantica* X(1) pp 96–111. (1989)
28. Wierzbicka, A.: The semantics of color: A new paradigm. In: Pitchford, N.J. & Biggam, C.P. (Eds.), *Progress in Colour Studies: Volume I. Language and Culture* (pp. 1–24). (2006)
29. Wittgenstein L.: *Philosophische Untersuchungen* (1953)