Generative Lexicon Theory and Linguistic Linked Open Data

Fahad Khan **Francesca Frontini** **Riccardo Del Gratta** **Monica Monachini** **Valeria Quochi**
Consiglio Nazionale delle Ricerche, Istituto di Linguistica Computazionale "A. Zampolli"
Via Moruzzi 1, Pisa, Italy
name.surname@ilc.cnr.it

Abstract

In this paper we look at how Generative Lexicon theory can assist in providing a more thorough definition of word senses as links between items in a RDF-based lexicon and concepts in an ontology. We focus on the definition of lexical sense in *lemon* and show its limitations before defining a new model based on *lemon* and which we term *lemonGL*. This new model is an initial attempt at providing a way of structuring lexico-ontological resources as linked data in such a way as to allow a rich representation of word meaning (following the GL theory) while at the same time (attempting to) re-main faithful to the separation between the lexicon and the ontology as recommended by the lemon model.

1 Introduction

The linked data movement aims to make it easier to publish and to use collections of data stored at different online locations by providing a standardized way of structuring, describing, and interlinking this data. One of the main tools in the linked data movement's arsenal is the Resource Description Framework (RDF)[1] a general purpose language which organises data on the basis of *subject-predicate-object* triples. These triples are used to link together data stored at different locations using Unique Resource Idenfiers. The RDF model also serves as the basis for Web Ontology Language (OWL)[2] a family of formal knowledge representation languages of varying degrees of expressivity developed for the purpose of building ontologies.

For the Language Resources and Technology (LRT) community the popularity of linked data makes it far easier to carry out its traditional aims of standardising, linking, and re-using linguistic data. This has led to a trend towards the conversion of existing lexicons using the RDF format and other linked data tools. It also opens up the way for a greater linking up of distributed lexical and ontological resources offering both greater access to the knowledge explicitly stated in a lexicon as well as increased possibilities for inferring new knowledge from associated ontologies (Smrž and Sinopalnikova, 2003).

In this paper we look at *lemon*, an increasingly popular RDF-based model for sharing lexical information online. We focus in particular on how *lemon* handles the link between the lexical entries contained in a lexicon and the associated semantic data in an ontology. We examine some of the shortcomings of *lemon* in this respect and suggest an altered version of *lemon*, *lemonGL*, which attempts to present a more accurate model of the interaction between lexical entries and their meanings. This model is based on Generative Lexicon theory (GL) which treats lexical senses as complex structured semantic objects which can enter into a range of generative semantic operations with other senses in order to generate meanings for linguistic expressions.

In the next section, Section 2.1 we look at the *lemon* view of senses, and argue that it leads to difficulties when it comes to modelling logical polysemy. In Section 2.2 we give a brief overview of the GL approach to word senses. In Section 3, we look at a previous example of the conversion of a GL inspired lexical resource, PAROLE-SIMPLE-CLIPS, using *lemon*, and explore some of the issues related to that conversion. We give a definition of *lemonGL* in Section 4. In the final section we present our conclusions.

[1] http://www.w3.org/RDF/
[2] http://www.w3.org/OWL/

2 *lemon* and GL

2.1 *lemon* senses and their limitations

lemon is a model that provides an RDF-based standard for publishing lexical data online (McCrae et al., 2011). As such it has fast gained both acceptance and widespread popularity within the LRT community[3]. At its heart *lemon* defines a set of core modules that help to describe the basic aspects of the entries in most lexicons such as those aspects relating to morphology, the phrase structure of complex expressions, and the syntactic frames associated with predicative lexical items. It also allows the addition of semantic information to any given lexical entry by mapping the entry to a concept in an ontology via an intermediate lexical sense object. This is all based on the idea of semantics by reference and entails a clear separation between the linguistic and ontological levels of a lexical resource as well as facilitating the "plugging-in" of different ontologies into the same lexicon – this turns out to be particularly useful when it comes to modelling the meanings of terms in different domains.

We now provide a brief overview of the theoretical basis for the *lemon* treatment of lexical senses and references as developed in (Cimiano et al., 2012) and as also presented in (McCrae et al., 2010).

Within the *lemon* framework, each lexical entry l in a *lemon* lexicon L can be mapped to an concept c in an ontology O via a *lemon* lexical sense object, $\sigma^{(l,c)}$. The definition of a *lemon* lexical sense object given in (Cimiano et al., 2012) presents three different, complementary, facets to each *lemon* word sense. These are as follows.

Firstly, a *lemon* sense object $\sigma^{(l,c)}$ can be viewed as representing "a subset of the uses of the lexical entry l in which l can be understood as meaning concept c", or in other words as a *disambiguated lexical entry*. So that for example given the lexical item $bank$, and a concept **bank** within an ontology where it is taken to mean something like an incline at the side of a body of water, we can view the sense object $\sigma^{(bank,\mathbf{bank})}$ as standing for the set of uses of the word $bank$ where it has a meaning corresponding to this geographical sense. Secondly, we can understand the sense object $\sigma^{(l,c)}$ as the "reification" of a pairing between the lexical entry l and the ontological element c where $\sigma^{(l,c)}$ is defined as *valid* if there exists evidence of at least one instance of the lexical item l being taken to mean c. Finally, $\sigma^{(l,c)}$ can also be seen as the hypothetical full meaning of the lexical entry l, such that if this full meaning were to be added to the ontology as a concept then it would be a subtype of c.

[3] See http://lemon-model.net/index.html for more details.

The foregoing tripartite definition of a lexical sense object seems to suggest that each time we are able to match the meaning of a lexical entry l to a concept c in an ontology we are also permitted to create a new *lemon* lexical sense object that will serve to mark this pairing of word and meaning. It is also clear that *lemon* sense objects only play a limited role at the intersection of a lexicon and an ontology and correspondingly carry very little structure (although as we will see there is provision in *lemon* for subsenses as well as for mappings between senses and representations of syntactic frames). All of which leaves us with a reasonably clear division of labour: the lexicon should contain all the morpho-syntactic data that relates to a lexical entry, whereas the ontology should contain most if not all of the "purely" semantic data associated with the entry, with *lemon* sense objects serving to map between these two layers.

This view of the sense relation between a lexical entry and an ontological concept can be seen however to lead to difficulties when it comes to modelling examples of logical polysemy. Following (Pustejovsky, 1995, 28) we define logical polysemy as any (syntactically realised) category-preserving semantic ambiguity where the different senses of a word have meanings that overlap or that otherwise clearly depend on one another. For example take the following two sentences.

- She walks through the door.
- She paints the door.

These sentences demonstrate that the noun $door$ can be taken to mean either an aperture allowing passage (as in the first sentence) or a physical object occupying such an aperture (as in the second).

But then, according to the *lemon* model, given any ontology that distinguishes between these two concepts (i.e., between $door$ as aperture and $door$ as physical object) as c_1, c_2 respectively, and given a lexicon with an entry for $door$ which we wish to map to the aforementioned ontology, there should be at least two distinct *lemon* sense objects

$\sigma^{(door,c_1)}$ and $\sigma^{(door,c_2)}$ mapping between the lexicon and the ontology. Indeed one could argue that even were there just one concept c in our ontology O representing the meaning of the word *door* in a vague enough way to cover both meanings of *door* then we would still be justified in creating at least two different sense objects since both of the instances of *door* given above represent different full hypothetical concepts though with the same reference, c, in O.

In other words the *lemon* model seems to necessitate what Pustejovsky calls a sense enumeration lexicon: that is a lexicon in which the multiple senses of each word are stored separately[4]. Pustejovsky argues in (Pustejovsky, 1995) that the sense enumerative approach to lexicon design is problematic precisely because it fails to capture several important aspects of the phenomena of logical polysemy. This inadequacy is addressed under three different heads in (Pustejovsky, 1995).

Firstly, the sense enumerative approach makes it extremely awkward to deal with the creativity of word use. For example an adjective like *fast* is able to appear in different, potentially novel, contexts and to have different (though related) meanings in each: *fast* in the phrase *a fast car* means something different from the use of *fast* in the phrase *a fast motorway*, which in turn has a different meaning from *fast* in *a fast programmer* or *a fast song*. Indeed the possibilities seem open ended, and a simplistic sense enumerative approach in which lexical sense entities are multiplied at every turn seems at the very least impracticable. Secondly, a basic sense enumerative approach fails to capture the relatedness, or in Pustejovsky's terms, the permeability of word senses. For example if we create a distinct sense apiece for the word $lamb$ when taken to mean a young sheep and when it is taken to mean the meat of a young sheep, respectively, then its hard to see how to relate these two senses under a basic sense enumerative approach. Certainly *lemon* sense relations like *equivalent, incompatible, narrower*, or *broader* fail to capture the close relationship between these two different meanings of $lamb$. Thirdly, verbs like *forget* can have different syntactic realisations each of which seem to require a separate sense.

All of this would tend to suggest that another, more nuanced approach to sense relations is in order – or at the very least it means that if a sense is to serve as an intermediary between a lexical item l and its meaning as a concept c then it might not be enough to regard the sense as merely a simple atomic pairing of a word and meaning.

One could argue however that rather than adding extra structure to the sense object there should be sufficient information within the ontology to derive c as a meaning for l, assuming that an already existing lexical sense pairing (l, c') exists, and that c can be somehow derived from c'. So that, for example, given the meaning of the lexical entry of *lamb* as a young sheep in the ontology and given other (commonsense) knowledge contained in the ontology representing the fact that the flesh of young sheep is edible and is commonly eaten by humans, it should be possible to derive the correct sense of lamb in the following sentence.

- I had some delicious lamb last night.

However, since most types of systematic polysemy are only semi productive[5] and since according to many (although by no means all) semantic theories, only limited aspects of commonsense or world knowledge are necessary for disambiguating most cases of logical polysemy, this would necessitate organising the concepts in an ontology in a particular (theory specific, linguistically motivated) way or, in some cases anyway, enriching the kinds of relations that can hold between senses. Thus we might include a relation between senses to represent the fact that they can take part in a systematic polysemic alternation (although, simply adding a polysemy relation between systematically related senses only partially solves the issue, as we would still lack the explanation of how the senses are related; i.e. what the specific dimensions of meaning involved are).

This strategy is problematic however for a number of reasons many of which relate to the fact that it seems to necessitate a certain sort of linguistically based organisation of our ontology in order to make efficient use of the information held therein; it thus very obviously blurs the distinction between what is contained in a lexicon and what is contained in corresponding ontologies. This strategy would also call for a major redefinition of the

[4]Although this might also entail that, for instance, so called complementary related senses are stored under a single entry.

[5]For instance even though the word for a young sheep and the word for the meat of a young sheep are the same in English and even though this is the case for many other animals it is not true of cows and beef.

role played by a *lemon* lexical sense object between a lexical entry l and a concept c so that it represents something along the lines of, say, the concept c's being a prototypical/common reading of l. In Section 3 we discuss an attempt to convert a GL inspired lexical resource PAROLE SIMPLE CLIPS with *lemon* using just this kind of strategy.

Another sort of strategy and one which we will discuss in some detail below is to give *lemon* senses additional structure so that as well as providing a link to a reference in an ontology they enable a more efficient access to particular kinds of "explanatory" information such as are necessary for disambiguating the meanings of polysemous words. This would effectively create an intermediate layer between the lexicon and the ontology and would have the benefit of retaining most of the rest of the *lemon* syntax as well as mimimizing our assumptions as to the structure of the ontology and thereby helping to maintain – at least to a substantial extent – the *lemon*-inspired lexicon-ontology distinction described above[6]. We will detail one potential theoretical foundation for this kind of approach in the next section. We present a model based on this strategy in Section 4.

2.2 The Generative Lexicon Approach

Generative Lexicon theory (GL henceforth) is a theory of lexical organisation that treats senses not as atomic units but instead as formal entities with a complex internal (conceptual) structure which can be described using four different levels of representation. These levels are as follows:

- The **argument structure** - An elaboration of the type and number of logical arguments associated with the entry, along with associated syntactic information;
- The **event structure** - A specification of the event structure associated with the entry;
- The **qualia structure** - A specification of the four qualia roles associated with the entry, see below;
- The **lexical inheritance structure** - The place of the lexical entry within a wider type system.

For each lexical entry GL foregrounds four different, representative, aspects of word meaning, the so called qualia roles, which along with a number of generative mechanisms are, or so the proponents of this approach would claim, sufficient to handle most cases of logical polysemy and creative sense modulation in context. These four qualia roles are regarded as being the modes of explanation for a lexical item and also as generalising the idea of verbal argument structure to apply to nominals, etc. They are defined as follows:

- The **formal**: that which specifies the hierarchical relations of an entity with other entities;
- The **constitutive**: that which specifies what an entity is made of, its relations with its various components;
- The **telic**: that which specifies the function or purpose of an entity;
- The **agentive**, that which specifies the origin of an entity, how it came about.

Pustejovsky (Pustejovsky, 1995) uses these four qualia roles (and the notion of a complex type) along with the information contained in the other representative levels in a lexical entry, as well as a number of generative semantic mechanisms such as type coercion and co-composition, to show how it is possible to disambiguate a variety of different kinds of logical polysemy without having to resort to the division of a word sense into separate senses for each shift in meaning such as is, as we have seen, characteristic of the sense enumerative approach. It is further argued in (Pustejovsky, 1995) that the kind of linguistic knowledge necessary for disambiguating instances of logical polysemy is distinct from the general common sense or pragmatic knowledge that is useful in, say, disambiguating instances of homonymous words such as *bank*, since in the former case rather than choosing between two or more different "contrastive" meanings we instead focus on diverse aspects of a single, complex meaning. This idea will play an important role in our proposed model.

As we noted above the knowledge contained in the qualia structure represents a set of basic building blocks for structuring and generating the concepts expressed by a word sense. The qualia structure can therefore be seen as the main interface to the knowledge of the world such as might be represented by an ontology. We will expand on this in what follows below.

[6]Although as we will see it does necessitate some level of reduplication of what is contained in the ontology.

3 Converting PAROLE SIMPLE CLIPS with *lemon*

In this section we briefly detail an attempt made in previous work to convert a GL based lexical resource, PAROLE SIMPLE CLIPS (PSC), into the RDF format using *lemon*; full details of the conversion are available at (del Gratta et al., 2013)[7]. PSC is a multi-layered Italian language lexicon built up within the framework of three successive projects: the EU-funded PAROLE (Ruimy et al., 1998) and SIMPLE (Lenci et al., 2000) and the Italian national project CLIPS. In particular, the conversion focused on the Italian SIMPLE lexicon, i.e., the lexical semantic layer of the PSC lexical database. As the model used by SIMPLE was strongly informed by GL its conversion is of particular interest for our discussion here.

Each lexical sense object or semantic unit (*USem* for short) in SIMPLE is described using the four different qualia roles, although in this instance the qualia roles are represented as binary relations between the USem in question and other USems in the SIMPLE semantic layer. These relations together comprise a so called *extended qualia structure*. This is a hierarchy of relations arranged at the top level under the four original qualia roles and structured in such a way as to build upon the notion of qualia structure found in the GL literature. For example, the USem corresponding to the semantic type *Vehicle* is associated with the agentive relation $created\_by$, the constitutive relations $made\_of$, has_as_part, and the telic relation $used\_for$; the formal role is given by the is_a relation.

The first part of the conversion of PSC was to define and build the top level ontology which had been described in the specifications of the project and used as a framework for the SIMPLE semantic layer encoding but which hadn't originally been implemented as a separate ontological resource (see (Toral and Monachini, 2007)). Next the SIMPLE lexical database, i.e., the set of USems and the relations which held between them, were converted using the *lemon* model. The main problem here was the fact that the USems in SIMPLE took part both in conceptual relations concerning the meaning and reference of their associated lexical entries, such as for example that a certain kind of tree produces a certain kind of fruit, as well as purely sense based relations such as synonymy. This all needed to be disentangled in order to maintain a separation between the lexicon and the ontology and to thus properly adhere to the philosophy of *lemon*.

Briefly, the solution adopted was to use *lemon* to model the purely lexical part of the resource and then to convert the semantic layer of SIMPLE into an OWL ontology (subsequently linked to the top level ontology). This meant that a new *lemon* lexical resource was created in which each lexical entry was related to its corresponding USem via the *lemon* sense relation. Each one of these USems was duplicated twice, once as a *lemon* sense object and once as an object which was then added to the ontology previously constructed.

This posed some interesting questions about the status of the qualia relations which were now fully transformed into ontological relations. While it is true that the qualia structure encodes knowledge about the world, the relations represented by the qualia structure present only a limited range of the kinds of commonsense knowledge that we might put in an ontology. How then is it possible to distinguish between those relations that are relevant for lexical semantics and those that are just part of encyclopedic knowledge (and which are therefore relevant to language understanding at a more pragmatic level)? The top level ontology in PSC was designed with a focus on structuring the semantic layer of a lexical resource and in fact its design closely follows the notion of qualia structure as found in the GL literature. Therefore there was no neccessity to try and separate out that part of the PSC ontology which dealt with "linguistic" knowledge from the rest. But things in general won't always be so straightforward: we may not be able in other situations to depend on such a closely "linguistic" structuring of ontological knowledge.

4 *lemonGL*

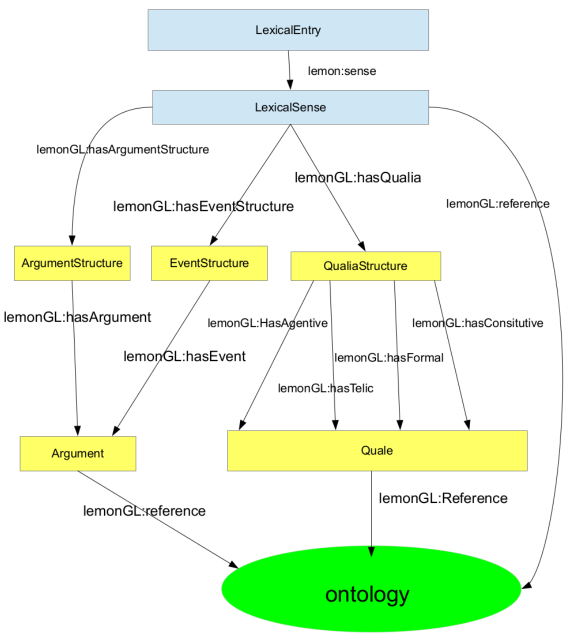
In this section we present *lemonGL* a RDF-based model that builds upon the *lemon* model by presenting a more nuanced version of lexical senses, one that falls in broadly line with the view presented in GL theory. *lemonGL* is an initial attempt at providing a way of structuring lexico-ontological resources as linked data in such a way as to make it easier to access those aspects of a lexical entry's meaning that best serve as modes

[7]The conversion is incomplete: so far only the nouns have been converted.

of explanation for that entry (according to GL theory) while at the same time (attempting to) remain faithful to the separation between the lexicon and the ontology as recommended by the *lemon* model. In what follows we will describe the *lemonGL* model and explain where it diverges from *lemon* before presenting an example to illustrate its use.

lemonGL differs from *lemon* essentially only in its definition of lexical senses and in the kinds of relations into which a lexical sense enters. In *lemonGL* a lexical sense object is still connected to a lexical entry via a *lemon* sense relation which can in turn link the lexical entry to a concept to an ontology that provides a meaning for the lexical entry. On the other hand, a *lemonGL* lexical sense object has a complex structure of its own, and each sense object can be related to an *ArgumentStructure* object via an *hasArgumentStructure* relation; to an *EventStructure* object via an *hasEventStructure* relation both of which are in turn related to *GLArgument* objects; and to a *QualiaStructure* object via a *hasQualia* relation. This *QualiaStructure* is related in its turn to a *Quale* object via the *hasAgentive, hasFormal, hasTelic* and *hasConstitutive* relations. These extra objects then provide a sort of middle layer, or an interface, between the lexical entry and the ontology that serves to isolate certain aspects of the entry's meaning as contained within the ontology.

Figure 1: A diagram of the *lemonGL* model.



This might seem a little like overkill: if the meaning of a lexical entry l is determined by a concept c in an ontology O along with the network of relations that c enters into with other items in the ontology, then what is the purpose of transferring or duplicating a subset of this information into the lexicon? The answer is that it helps to preserve the division between the language specific information in the lexicon and the (relatively speaking) language independent conceptual information contained in the ontology.

As discussed above, in GL theory a word's qualia structure serves to specify the central modes of explanation associated with that word – as distinguished from other more general, common sense, one could say, more purely ontological, knowledge – and that the knowledge encoded in qualia structures is used in dealing with instances of logical polysemy. In fact one could argue that the qualia information is part of a lexical entry in the same way that a verbs argument structure is part of a verb's lexical information and that it therefore doesnt really belong in the ontology.

Another option would be for a separate, linguistically motivated ontology to hold this kind of knowledge which could then be somehow linked up with other ontologies, but this solution is uneconomical both from the theoretical and the practical point of view. Theoretically it would imply that human beings have a subset of their encyclopedic knowledge duplicated as part of their linguistic knowledge; practically speaking it would involve a lot of duplication of labour. With respect to the interaction between the linked data movement and the LRT community it seems necessary to be able both to easily access and to build upon the large amount of formalised knowledge that is currently becoming available online, while at the same time retaining the ability to set the boundaries as to what is relevant to lexical semantics. The *lemonGL* model then attempts to structure word senses in a way that maintains a Pustejovskian linguistic versus commonsense knowledge distinction.

We now present an example modelled using *lemonGL* to illustrate its potential use. We will be using RDF turtle syntax to present our example[8]. The example concerns the noun *wine* which we will represent with the following feature struc-

[8] http://www.w3.org/TeamSubmission/turtle/

ture[9].

$$\left[\begin{array}{lll}\textbf{WINE} & & \\ \text{ARGS} & = & \left[\begin{array}{l}\text{ARG1}= x \\ \text{D-ARG1} = y\end{array}\right] \\ \text{EVENTS} & = & \left[\text{D-E1}= e\right] \\ \text{QUALIA} & = & \left[\begin{array}{l}\text{FORMAL}= liquid(x) \\ \text{AGENTIVE}= make(e,y,x)\end{array}\right]\end{array}\right]$$

Here the argument structure has two logical arguments, x and y and the event structure has one event argument e, all of which are found in the qualia expressions in the qualia structure: these arguments can be understood to play the same role as the bound variables x, y and e in the following lambda expression:

$$\lambda x \lambda y \lambda e[liquid(x) \wedge make(e, y, z)].$$

Only two of the qualia roles are instantiated. The first, the formal quale here expresses the fact that wine is a type of liquid using the *liquid* predicate; the second quale, by referring to the *make* predicate and the variables mentioned previously, to expresses the fact that there is a process of creation associated with each instance of a wine "entity".

In order to represent this feature structure using *lemonGL* we will first assume that our (OWL) ontology contains the following definitions.

```
:hasMadeObject  rdf:type owl:ObjectProperty ;
rdfs:range :Made_Object ;
rdfs:domain :Make_Event .
:hasMaker  rdf:type owl:ObjectProperty ;
rdfs:domain :Make_Event ;
rdfs:range :Maker .
:makes  rdf:type owl:ObjectProperty ;
rdfs:range :Made_Object ;
rdfs:domain :Maker .
:Made_Object rdf:type owl:Class .
:Make_Event rdf:type owl:Class .
:Maker rdf:type owl:Class .
:liquid rdf:type owl:Class .
```

The following lines of RDF structure the sense of the lexical entry for *wine* in the lexicon into an argument structure, an event structure and a qualia structure:

```
:wine rdf:type lemon:LexicalEntry ,
owl:NamedIndividual ;
lemon:sense [rdf:type lemon:LexicalSense ;
lemonGL:hasArgumentStructure :wine_arg_str;
lemonGL:hasEventStructure :wine_ev_str ;
lemonGL:hasQualia :wine_qua_str].
```

[9]The feature structure for *wine* is taken from (Pustejovsky, 1998). Also note that the OWL/RDF code used in the examples has been made up for the purpose of demonstration, and is not drawn from an existing resource. We have not so far converted any non trivial lexical resource into RDF using *lemonGL*.

We refer to the argument structure associated with the lexical sense of the entry for *wine* using the identifier $wine\_arg\_str$. We specify that the first argument associated with $wine\_arg\_str$ has the ontological type of $Made\_Object$, whereas the second argument has the ontological type of $Maker$.

```
:wine_arg_str rdf:type  Glemon:ArgumentStructure ,
owl:NamedIndividual ;
lemonGL:hasArgument
[lemonGL:reference ontology:Made_Object ] ,
[lemonGL:reference ontology:Maker] .
```

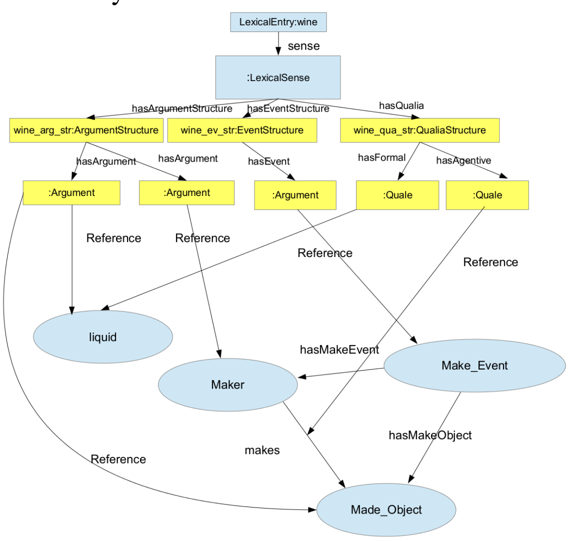
Next we specify that the ontological type of the event associated with the event structure of the sense object is a Make_Event.

```
:wine_ev_str rdf:type lemonGL:EventStrucuture ,
owl:NamedIndividual ;
lemonGL:hasEvent
[lemonGL:reference ontology:Make_Event ].
```

Finally we specify that the agentive role for the qualia structure associated with the sense object refers to a make relation in our ontology, and that the formal role has the ontological type of $liquid$.

```
:wine_qua_str rdf:type lemonGL:QualiaStructure ,
owl:NamedIndividual ;
lemonGL:hasAgentive
[lemonGL:reference ontology:makes] ;
lemonGL:hasFormal
[lemonGL:reference ontology:liquid ].
```

The following figure presents the example schematically:



Even though we have somewhat altered the *lemon* framework, the changes we propose are far from drastic. Indeed, as is helpfully illustrated by the representation of the verb *to give* in the *lemon* cookbook (McCrae et al., 2010), *lemon* senses can have subsenses which can in turn be mapped onto *lemon Argument* objects which are themselves linked to a *Frame* object. This parallel between our GL-inspired representation of the noun *wine* in *lemonGL* and the representation of a verb like *give* in *lemon* is no surprise since, as we've described above, one of the motivations behind GL theory was to provide what is in essence an argument structure for nominals (Pustejovsky and Boguraev, 1993). On the other hand, as we have striven to show throughout this paper, the notion of sense in *lemon* stands in need of substantial revision and part of this means providing extra functionality for other part of speech categories.

5 Conclusion

In this paper we've tried to argue for a revised notion of a sense object in *lemon*, one that both makes it easier to model a range of important linguistic phenomena and that enables the practical implementation (to some extent) of an important and influential theory of lexical semantics. We plan to continue this work by using *lemonGL* to model existing lexical resources, developing the language further as the need arises, and also to investigate the extent to which *lemonGL* makes it easier to reason about such resources.

We have discussed the desirability of maintaining a separation between lexical and ontological knowledge. We believe by adding what is effectively an intermediary layer between the lexicon and the ontology we have created a model for lexical-ontological resources which preserves this separation as far as possible (by limiting what we can assume about the structure of the ontology) while still enabling us to handle the phenomena of logical polysemy.

To take a more general view, we feel that it is of the utmost importance, given the current popularity of LLOD as well as the great potential that it holds out, that the GL community become more active in the definition of the models that are defining the structure of LLOD lexicons and their connections to existing or new conceptual resources. In particular it is important that models that are too geared towards sense enumeration do not become predominant to the detriment of more realistic models of lexical semantics, and that the available lexicon representation schemes allow for the real complexity of lexical semantic relations to be fully represented.

References

Philipp Cimiano, John McCrae, Paul Buitelaar, and Elena Montiel-Ponsoda, 2012. *On the Role of Senses in the Ontology-Lexicon*.

Riccardo del Gratta, Francesca Frontini, Fahad Khan, and Monica Monachini. 2013. Converting the parole simple clips lexicon into rdf with lemon. *Semantic Web Journal (Under Review)*.

Alessandro Lenci, Nuria Bel, Federica Busa, Nicoletta Calzolari, Elisabetta Gola, Monica Monachini, Antoine Ogonowski, Ivonne Peters, Wim Peters, Nilda Ruimy, Marta Villegas, and Antonio Zampolli. 2000. Simple: A general framework for the development of multilingual lexicons. *International Journal of Lexicography*, 13(4):249–263.

John McCrae, Guadalupe Aguado de Cea, Paul Buitelaar, Philipp Cimiano, Thierry Declerck, Asuncin Gmez Prez, Jorge Gracia, Laura Hollink, Elena Montiel-Ponsoda, Dennis Spohr, and Tobias Wunner, 2010. *The Lemon Cookbook*. http://lemon-model.net/lemon-cookbook.pdf.

John McCrae, Dennis Spohr, and Philipp Cimiano. 2011. Linking lexical resources and ontologies on the semantic web with lemon. In *Proceedings of the 8th extended semantic web conference on The semantic web: research and applications - Volume Part I*, ESWC'11, pages 245–259, Berlin, Heidelberg. Springer-Verlag.

James Pustejovsky and Branimir Boguraev. 1993. Lexical knowledge representation and natural language processing. *Artif. Intell.*, 63(1-2):193–223.

James Pustejovsky. 1995. *The Generative Lexicon*. MIT Press, Cambridge, MA.

James Pustejovsky. 1998. The semantics of lexical underspecification.

N. Ruimy, O. Corazzari, E. Gola, A. Spanu, N. Calzolari, and A. Zampolli. 1998. The european le-parole project: The italian syntactic lexicon. In *Proceedings of the First International Conference on Language resources and Evaluation*, pages 241–248.

Pavel Smrž and Anna Sinopalnikova. 2003. Present-day lexical knowledge bases - what they are and what they need. In *Proceedings of the Eight International Symposium on Social Communication*, pages 100–105, Santiago de Cuba. Center of Applied Linguistics of the Santiago de Cuba's branch of the Ministry of Science, Technology and the Environment.

Antonio Toral and Monica Monachini. 2007. Simple-owl: a generative lexicon ontology for nlp and the semantic web. In *Workshop of Cooperative Construction of Linguistic Knowledge Bases, 10th Congress of Italian Association for Artificial Intelligence.*