GLML: Annotating Argument Selection and Coercion

James Pustejovsky, Anna Rumshisky,
Jessica L. Moszkowicz, Olga Batiukova

Abstract

In this paper we introduce a methodology for annotating compositional operations in natural language text, and describe a mark-up language, GLML, based on Generative Lexicon, for identifying such relations. While most annotation systems capture surface relationships, GLML captures the "compositional history" of the argument selection relative to the predicate. We provide a brief overview of GL before moving on to our proposed methodology for annotating with GLML. There are three main tasks described in the paper: (i) Compositional mechanisms of argument selection; (ii) Qualia in modification constructions; (iii) Type selection in modification of dot objects. We explain what each task includes and provide a description of the annotation interface. We also include the XML format for GLML including examples of annotated sentences.

1 Introduction

1.1 Motivation

In this paper, we introduce a methodology for annotating compositional operations in natural language text. Most annotation schemes encoding "propositional" or predicative content have focused on the identification of the predicate type, the argument extent, and the semantic role (or label) assigned to that argument by the predicate (see Palmer et al., 2005, Ruppenhofer et al., 2006, Kipper, 2005, Burchardt et al., 2006, Ohara, 2008, Subirats, 2004).

The emphasis here will be on identifying the nature of the compositional operation rather than merely annotating the surface types of the entities involved in argument selection.

Proceedings of the 8th International Conference on Computational Semantics, pages 169–180,
Tilburg, January 2009.

Consider the well-known example below. The distinction in semantic types appearing as subject in (1) is captured by entity typing, but not by any sense tagging from, e.g., FrameNet (Ruppenhofer et al., 2006) or PropBank (Palmer et al., 2005).

(1) a. Mary called yesterday.
b. The Boston office called yesterday.

While this has been treated as *type coercion* or *metonymy* in the literature (cf. Hobbs et al., 1993 , Pustejovsky, 1991, Nunberg, 1979, Egg, 2005), the point here is that an annotation using frames associated with verb senses should treat the sentences on par with one another. Yet this is not possible if the entity typing given to the subject in (1a) is HUMAN and that given for (1b) is ORGANIZATION.

The SemEval Metonymy task (Markert and Nissim, 2007) was a good attempt to annotate such metonymic relations over a larger data set. This task involved two types with their metonymic variants:

(2) i. **Categories for Locations**: literal, place-for-people, place-for-event, place-for-product;
ii. **Categories for Organizations**: literal, organization-for-members, organization-for-event, organization-for-product, organization-for-facility.

One of the limitations with this approach, however, is that, while appropriate for these specialized metonymy relations, the annotation specification and resulting corpus are not an informative guide for extending the annotation of argument selection more broadly.

In fact, the metonymy example in (1) is an instance of a much more pervasive phenomenon of type shifting and coercion in argument selection. For example, in (3) below, the sense annotation for the verb *enjoy* should arguably assign similar values to both (3a) and (3b).

(3) a. Mary enjoyed drinking her beer .
b. Mary enjoyed her beer.

The consequence of this, however, is that, under current sense and role annotation strategies, the mapping to a syntactic realization for a given sense is made more complex, and is, in fact, perplexing for a clustering or learning algorithm operating over subcategorization types for the verb.

1.2 Theoretical Preliminaries

The theoretical foundations for compositional operations within the sentence have long been developed in considerable detail. Furthermore, type shifting and type coercion operations have been recognized as playing an important role in many formal descriptions of language, in order to maintain compositionality (cf. Partee and Rooth, 1983; Chierchia, 1998; Groenendijk and Stokhof, 1989; Egg, 2005; Pinkal, 1999; Pustejovsky, 1995, and many others). The goal of the present work is to: (a) create a broadly applicable specification of the compositional operations involved in argument selection; (b) apply this specification over a corpus of natural language texts, in order to encode the selection mechanisms implicated in the compositional structure of the language.

The creation of a corpus that explicitly identifies the "compositional history" associated with argument selection will be useful to computational semantics in several respects: (a) the actual contexts within which type coercions are allowed can be more correctly identified and perhaps generalized; (b) machine learning algorithms can take advantage of the mapping as an additional feature in the training phase; and (c) some consensus might emerge on the general list of type-changing operations involved in argument selection, as the tasks are revised and enriched.

For the purpose of this annotation task, we will adopt the general approach to argument selection within Generative Lexicon, as recently outlined in Pustejovsky (2006) and Asher and Pustejovsky (2006). We can distinguish the following modes of composition in natural language:

(4) a. PURE SELECTION (Type Matching): the type a function requires is directly satisfied by the argument;
b. ACCOMMODATION: the type a function requires is inherited by the argument;
c. TYPE COERCION: the type a function requires is imposed on the argument type. This is accomplished by either:
   i. *Exploitation*: taking a part of the argument's type;
   ii. *Introduction*: wrapping the argument with the required type.

Each of these will be identified as a unique relation between the predicate and a given argument. In this annotation effort, we restrict the possible relations between the predicate and a given argument to *selection* and *coercion*. A more fine-grained typology of relations may be applied at a later point. Furthermore, qualia structure values[1] are identified in both argument selection and modification contexts.

The rest of this document proceeds as follows. In Section 2, we describe our general methodology and architecture for GL annotation. Section 3 gives an overview of each of the annotation tasks as well as some details on the resulting GLML markup. A more thorough treatment of the material we present, including the complete GLML specification and updates on the annotation effort can be found at www.glml.org.

2 General Methodology and Architecture

In this section, we describe the set of tasks for annotating compositional mechanisms within the GL framework. The current GL markup will include the following tasks, each of which is described below in Section 3.

(5) a. Mechanisms of Argument Selection: Verb-based Annotation
    b. Qualia in Modification Constructions
    c. Type Selection in Modification of Dot Objects

2.1 System Architecture

Each GLML annotation task involves two phases: *the data set construction phase* and *the annotation phase*. The first phase consists of (1) selecting the target words to be annotated and compiling a sense inventory for each target, and (2) data extraction and preprocessing. The prepared data is then loaded into the annotation interface. During the annotation phase, the annotation judgments are entered into the database, and the adjudicator resolves disagreements. The resulting database representation is used by the exporting module to generate the corresponding XML markup, stand-off annotation, or GL logical form.

These steps will differ slightly for each of the major GLML annotation tasks. For example, Task 1 focuses on annotating compositional processes between the verbs and their arguments. The first step for this task involves (1) selecting the set of target verbs, (2) compiling a sense inventory for each

[1]The qualia structure, inspired by Moravcsik (1975)'s interpretation of the *aitia* of Aristotle, is defined as the modes of explanation of a word or phrase, and defined below (Pustejovsky, 1991): (a) FORMAL: the category distinguishing the meaning of a word within a larger domain; (b) CONSTITUTIVE: the relation between an object and its constituent parts; (c) TELIC: the purpose or function of the object, if there is one; (d) AGENTIVE: the factors involved in the object's origins or "coming into being".

target, and (3) associating a type template or a set of templates with each sense. Since the objective of the task is to annotate coercion, our choices must include the verbs that exhibit the coercive behavior at least in some of their senses.

At the next step, the data containing the selected target words is extracted from a corpus and preprocessed. Since the GLML annotation is intra-sentential, each extracted instance is a sentence. Sentences are parsed to identify the relevant arguments, adjuncts or modifiers for each target. The data is presented to the annotatator with the target word and the headword of the relevant phrase highlighted.

Due to the complexity of the GLML annotation, we chose to use the task-based annotation architecture. The annotation environment is designed so that the annotator can focus on one facet of the annotation at a time. Thus, in Task 1, the verbs are disambiguated by the annotator in one subtask, and the annotation of the actual compositional relationship is done in another subtask. Figure 1 shows an example of the interface for the verb-based annotation task .

Figure 1: Example of Annotation Interface for GLML Annotation

2.2 The Type System for Annotation

The type system we have chosen for annotation is purposefully shallow, but we also aimed to include types that would ease the complexity of the annotation task. The type system is not structured in a hierarchy, but rather it is presented as a set of types. For example, we include both HUMAN and ANIMATE in the type system along with PHYSICAL OBJECT. While HUMAN is a subtype of both ANIMATE and PHYSICAL OBJECT, the annotator does not need to be concerned with this. This allows the annotator to simply choose the HUMAN type when necessary rather than having to deal with type inheritance.

While the set of types for GLML annotation can easily be modified, the following list is currently being used:

(6) HUMAN, ANIMATE, PHYSICAL OBJECT, ARTIFACT, ORGANIZATION, EVENT, PROPOSITION, INFORMATION, SENSATION, LOCATION, TIME PERIOD, ABSTRACT ENTITY, ATTITUDE, EMOTION, PROPERTY, OBLIGATION, AND RULE

3 Annotation Tasks

In this section, we describe the annotation process: the steps involved in each task and the way they are presented to the annotators. In this paper, we focus on the task descriptions rather than an in depth review of the annotation interface and the resulting GLML markup.

The general methodology for each task is as follows: 1) Select a target set of words and compile a sense inventory for each one, 2) Select a set of sentences for each target, 3) Disambiguate the sense of the target in a given sentence, and 4) Answer questions specific to the annotation task in order to create the appropriate GLML link.

3.1 Mechanisms of Argument Selection: Verb-based Annotation

This annotation task involves choosing which selectional mechanism is used by the predicate over a particular argument. The possible relations between the predicate and a given argument will, for now, be restricted to *selection* and *coercion*. In *selection*, the argument NP satisfies the typing requirements of the predicate, as in *The child threw the stone (*PHYSICAL OBJECT*)*. *Coercion* encompasses all cases when a type-shifting operation (exploitation or introduction) must be performed on the complement NP in order to satisfy selectional requirements of the predicate, as in *The White House (*LOCATION → HUMAN*) denied this statement.*

An initial set of verbs and sentences containing them has been selected for annotation. For each sentence, the compositional relationship of the verb with every argument and adjunct will be annotated. The target types for each argument are provided in a *type template* that is associated with the sense of the verb in the given sentence. For example, one of the senses of the verb *deny* (glossed as "State or maintain that something is untrue") would have the following type template: HUMAN *deny* PROPOSITION.

In the first subtask, the annotator is presented with a set of sentences containing the target verb and the chosen grammatical relation. The annotator is asked to select the most fitting sense of the target verb, or to throw out the example (pick the "N/A" option) if no sense can be chosen either due to insufficient context, because the appropriate sense does not appear in the inventory, or simply no disambiguation can be made in good faith.

Next, the annotator is presented with a list of sentences in which the target verb is used in the same sense and is asked to determine whether the argument in the specified grammatical relation belongs to the type specified in the corresponding template. If the argument belongs to the appropriate type, the "yes" box is clicked, generating a CompLink with compType="SELECTION". If "no" is selected, a type selection menu pops up below the first question, and the annotator is asked to pick a type from a list of shallow types which is usually associated with the argument. Consequently, a CompLink with compType="COERCION" is created with the corresponding source and target type.

The following example of GLML markup is generated from the database[2]:

Sir Nicholas Lyell, Attorney General, denies a cover-up.

```
<SELECTOR sid="s1">denies</SELECTOR>
a <NOUN nid="n1">cover-up</NOUN> .
<CompLink cid="cid1" sID="s1" relatedToNoun="n1" gramRel="dobj"
compType="COERCION" sourceType="EVENT" targetType="PROPOSITION"/>
```

3.2 Qualia Selection in Modification Constructions

For this task, the relevant semantic relations are defined in terms of the qualia structure. We examine two kinds of constructions in this task: adjectival modification of nouns and nominal compounds[3].

[2]While we present these examples as an inline annotation, a LAF (Ide and Romary, 2003) compliant offset annotation is fully compatible with GLML.

[3]Since target nouns have already been selected for these two tasks, it is also possible to annotate qualia selection in verb-noun contexts such as *Can you shine the lamp over here?* (TELIC). However, here we focus solely on the modification contexts mentioned here.

3.2.1 Adjectival Modification of Nouns

This task involves annotating how particular noun qualia values are bound by the adjectives. Following Pustejovsky (2000), we assume that the properties grammatically realized as adjectives ”bind into the qualia structure of nouns, to select a narrow facet of the noun’s meaning.” For example, in the NP “a sharp metal hunting knife”, *sharp* refers to the knife as a physical object, its FORMAL type, *metal* is associated with a material part of the knife (CONSTITUTIVE), and *hunting* is associated with how the knife is used (TELIC). Similarly, *forged* in “a forged knife” is associated with the creation of the knife (AGENTIVE).

The task begins with sense disambiguation of the target nouns. Questions are then used to help the annotator identify which qualia relations are selected. For example, the TELIC question for the noun *table* would be ”Is this adjective associated with the inherent purpose of table?” These questions will change according to the type associated with the noun. Thus, for natural types such as *woman*, the TELIC question would be ”Is this adjective associated with a specific role of woman?” Similarly, for the AGENTIVE role, the question corresponding to the PHYSICAL OBJECT-denoting nouns refers to the ”making or destroying” the object, while for the EVENT-denoting nouns, the same question involves ”beginning or ending” of the event. QLinks are then created based on the annotator’s answers, as in the following example:

The walls and the wooden table had all been lustily scrubbed.

```
<SELECTOR sid="s1">wooden</SELECTOR>
<NOUN nid="n1">table</NOUN>
<QLink qid="qid1" sID="s1" relatedToNoun="n1" qType="CONST"/>
```

3.2.2 Nominal Compounds

This task explores the semantic relationship between elements in nominal compounds. The general relations presented in Levi (1978) are a useful guide for beginning a classification of compound types, but the relations between compound elements quickly prove to be too coarse-grained. Warren’s comprehensive work (Warren, 1978) is a valuable resource for differentiating relation types between compound elements.

The class distinction in compound types in language can be broken down into three forms (Spencer, 1991): endocentric compounds, exocentric compounds, and dvandva compounds. Following Bisetto and Scalise

(2005), however, it is possible to distinguish three slightly differently constructed classes of compounds, each exhibiting endocentric and exocentric behavior: subordinating, attributtive, and coordinate.

We will focus on the two classes of *subordinating* and *attributive* compounds. Within each of these, we will distinguish between *synthetic* and *non-synthetic* compounds. The former are deverbal nouns, and when acting functionally (subordinating), take the sister noun as an argument, as in *bus driver* and *window cleaner*. The non-synthetic counterparts of these include *pastry chef* and *bread knife*, where the head is not deverbal in any obvious way. While Bisetto and Scalise's distinction is a useful one, it does little to explain how non-relational sortal nouns such as *chef* and *knife* act functionally over the accompanying noun in the compound, as above.

This construction has been examined within GL by Johnston and Busa (1999). We will assume much of that analysis in our definition of the task described here. Our basic assumption regarding the nature of the semantic link between both parts of compounds is that it is generally similar to the one present in adjectival modification. The only difference is that in nominal compounds, for instance, the qualia of a head noun are activated or exploited by a different kind of modifier, a noun. Given this similarity, the annotation for this task is performed just as it is for the adjectival modification task. A QLink is created as in the following example:

Our guest house stands some 100 yards away.

```
<SELECTOR sid="s1">guest</SELECTOR>
<NOUN nid="n1">house</NOUN>
<QLink qid="qid1" sID="s1" relatedToNoun="n1" qType="TELIC"/>
```

3.3 Type Selection in Modification of Dot Objects

This task involves annotating how particular types within dot objects are exploited in adjectival and nominal modification constructions. *Dot objects* or *complex types* (Pustejovsky, 1995) are defined as the product of a type constructor $\bullet$ (”dot”), which creates dot objects from any two types a and b , creating $a \bullet b$. Complex types are unique because they are made up of seemingly incompatible types such as FOOD and EVENT.

Given a complex type $c = a \bullet b$, there are three possible options: 1) the modifier applies to both a and b, 2) the modifier applies to a only, or 3) the modifier applies to b only. Option 1 would be illustrated by examples such as *good book [+info, +physobj]* and *long test [+info, +event]*. Examples such as *delicious lunch [+food, -event]* and *long lunch [-food, +event]* illustrate options 2 and 3. A listing of dot objects can be found in Pustejovsky (2005).

The sense inventory for the collection of dot objects chosen for this task will include only homonyms. That is, only contrastive senses such as the *river bank* versus *financial institution* for *bank* will need to be disambiguated. Complementary senses such as the financial institution itself versus the building where it is located are not included.

In order to create the appropriate CompLink, the annotator will select which type from a list of component types for a given dot object is exploited in the sentence. The resulting GLML is:

After a while more champagne and a delicious lunch was served.

```
<SELECTOR sid="s1">delicious</SELECTOR>
<NOUN nid="n1">lunch</NOUN>
<CompLink cid="cid1" sID="s1" relatedToNoun="n1" gramRel="mod"
compType="SELECTION" sourceType="[PHYS_OBJ,EVENT]"
targetType="PHYS_OBJ" />
```

4 Conclusion

In this paper, we approach the problem of annotating the relation between a predicate and its argument as one that encodes the compositional history of the selection process. This allows us to distinguish surface forms that directly satisfy the selectional (type) requirements of a predicate from those that are accommodated or coerced in context. We described a specification language for selection, GLML, based largely on the type selective operations in GL, and three annotation tasks using this specification to identify argument selection behavior.

There are clearly many compositional operations in language that have not been addressed in this paper. The framework is general enough, however, to describe a broad range of type selective behavior. As the tasks become more refined, the extensions will also become clearer. Furthermore, as other languages are examined for annotation, new tasks will emerge reflecting perhaps language-specific constructions.

Acknowledgements

The idea for annotating a corpus according to principles of argument selection within GL arose during a discussion at GL2007 in Paris, between one of the authors (J. Pustejovsky) and Nicoletta Calzolari and Pierrette Bouillon. Recently, the authors met with other members of the GLML Working Group in Pisa at the ILC (September 23-25, 2008). We would like to thank the members of that meeting for their fruitful feedback and discussion on an earlier version of this document. In particular, we would like to thank Nicoletta Calzolari, Elisabetta Jezek, Alessandro Lenci, Valeria Quochi, Jan Odijk, Tommaso Caselli, Claudia Soria, Chu-Ren Huang, Marc Verhagen, and Kiyong Lee.

References

N. Asher and J. Pustejovsky. 2006. A type composition logic for generative lexicon. *Journal of Cognitive Science*, 6:1–38.

A. Bisetto and S. Scalise. 2005. The classification of compounds. *Lingue e Linguaggio*, 2:319–332.

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Pado, and Manfred Pinkal. 2006. The salsa corpus: a german corpus resource for lexical semantics. In *Proceedings of LREC*, Genoa, Italy.

Gennaro Chierchia. 1998. Reference to kinds across language. *Natural Language Semantics*, *6(4)*.

Marcus Egg. 2005. *Flexible semantics for reinterpretation phenomena*. CSLI, Stanford.

Jeroen Groenendijk and Martin Stokhof, 1989. *Type-shifting rules and the semantics of interrogatives*, volume 2, pages 21–68. Kluwer, Dordrecht.

Jerry R. Hobbs, Mark Stickel, and Paul Martin. 1993. Interpretation as abduction. *Artificial Intelligence*, 63:69–142.

Nancy Ide and L. Romary. 2003. Outline of the international standard linguistic annotation framework. In *Proceedings of ACL'03 Workshop on Linguistic Annotation: Getting the Model Right*.

M. Johnston and F. Busa. 1999. The compositional interpretation of compounds. In E. Viegas, editor, *Breadth and Depth of Semantics Lexicons*, pages 167–167. Dordrecht: Kluwer Academic.

Karin Kipper. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Phd dissertation, University of Pennsylvania, PA.

J. N. Levi. 1978. *The Syntax and Semantics of Complex Nominals*. Academic Press, New York.

K. Markert and M. Nissim. 2007. Metonymy resolution at semeval i: Guidelines for participants. In *Proceedings of the ACL 2007 Conference*.

J. M. Moravcsik. 1975. Aitia as generative factor in aristotle's philosophy. *Dialogue*, 14:622–636.

Geoffrey Nunberg. 1979. The non-uniqueness of semantic solutions: Polysemy. *Linguistics and Philosophy*, 3:143–184.

Kyoko Hirose Ohara. 2008. Lexicon, grammar, and multilinguality in the japanese framenet. In *Proceedings of LREC*, Marrakech, Marocco.

M. Palmer, D. Gildea, and P. Kingsbury. 2005. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106.

Barbara Partee and Mats Rooth, 1983. *Generalized conjunction and type ambiguity*, pages 361–383. de Gruyter, Berlin.

Manfred Pinkal. 1999. On semantic underspecification. In Harry Bunt and Reinhard Muskens, editors, *Proceedings of the 2nd International Workshop on Computational Semantics (IWCS 2), January 13-15*, Tilburg University, The Netherlands.

J. Pustejovsky. 1991. The generative lexicon. *Computational Linguistics*, 17(4).

J. Pustejovsky. 1995. *Generative Lexicon*. Cambridge (Mass.): MIT Press.

J. Pustejovsky. 2000. Events and the semantics of opposition. In C. Tenny and J. Pustejovsky, editors, *Events as Grammatical Objects*, pages 445–482. Center for the Study of Language and Information (CSLI), Stanford, CA.

J. Pustejovsky. 2005. A survey of dot objects. Technical report, Brandeis University.

J. Pustejovsky. 2006. Type theory and lexical decomposition. *Journal of Cognitive Science*, 6:39–76.

J. Ruppenhofer, M. Ellsworth, M. Petruck, C. Johnson, and J. Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*.

A. Spencer. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Blackwell Textbooks in Linguistics, Oxford, UK and Cambridge, USA.

Carlos Subirats. 2004. FrameNet Español. Una red semántica de marcos conceptuales. In *VI International Congress of Hispanic Linguistics*, Leipzig.

B. Warren. 1978. *Semantic Patterns of Noun-Noun Compounds*. Acta Universitatis Gothoburgensis, Göteborg.