Enriching Wordnets with New Relations and with Event and Argument Structures*

Raquel Amaro[1]**, Rui P. Chaves[1], Palmira Marrafa[1,2], and Sara Mendes[1]***

[1] CLG – Group for the Computation of Lexical and Grammatical Knowledge, Center of Linguistics, University of Lisbon, Portugal
[2] Department of Linguistics of the Faculty of Arts, University of Lisbon, Portugal
{ramaro,rui.chaves,palmira.marrafa,sara.mendes}@clul.ul.pt

Abstract. This paper argues that wordnets, being concept-based computational lexica, should include information on event and argument structures. This general approach is relevant both for allowing computational grammars to cope with a number of different lexical semantics phenomena, as well as for enabling inference applications to obtain finer-grained results. We also propose new relations in order to adequately model non explicit information and cross-part-of-speech relations.

1 Introduction

Wordnets are electronic databases developed along with the same general lines of the so-called Princeton WordNet, an electronic database of English [1,2] containing nouns, verbs, adjectives, and adverbs. This database is structured as a network of relations between *synsets* (a set of roughly synonymous word forms). Several other wordnets have since been developed for many other languages and the number of relations adopted by the system has been enlarged (see for instance EuroWordNet [3]). In this paper we will show how wordnets can be integrated with a finer-grained lexical description framework in order to deal with various complex lexical semantics phenomena in a general and systematic way. Such an extension can be used both for deep lexical semantics analysis in computational grammars, and for a finer-grained linguistic knowledge-base in inference and question answering systems.

In Section 2 we will discuss the hyponymy/hypernymy relation. Following [4] we propose augmenting wordnet synset nodes with rich lexical-semantics descriptions which allow to explicitly capture the semantic inheritance patterns between hyponyms and hypernyms. We discuss some technical issues concerning this approach and provide a more general alternative view of semantic compatibility. Section 3 is dedicated to the verbal lexicon, focusing on argument

* The research reported in this paper has been partially supported by Instituto Camões, Fundação Calouste Gulbenkian, and Fundação para a Ciência e Tecnologia through the WordNet.PT, UniNet.PT, and COMPGRAM projects.
** Supported by FCT grant SFRH/BD/13875/2003.
*** Supported by FCT grant SFRH/BD/8524/2002.

© Springer-Verlag – Lecture Notes in Computer Science
(http://www.springeronline.com/lncs)

structure. We will show that a decompositional approach to troponymy enables us to establish more precisely what is inherited through the hierarchy, accounting for different argument structures. We will show that co-troponyms' incompatibility is accounted for at the argument structure level, and that it is possible to state the exact prepositional complements selected by a verb at the lexical level, through the use of the available lexical structures in the lexicon. Section 4 is also dedicated to the verbal lexicon, but it focuses on event structure. We argue that telicity is not only a compositional property of syntactic structures but also a lexical property to be encoded in the lexicon. The analysis focuses on the behavior of complex telic predicates, in particular those which are deficitary with regard to their lexical-conceptual structure. In order to represent appropriately such predicates in wordnets we propose a new relation, which has strong empirical motivation. In Section 5 we show that, despite the importance of the information that can be extracted from the hierarchical organization of lexical items, extending wordnets to all the main POS involves a revision of certain commonly used relations and the specification of several cross-part-of-speech relations. We focus on the specific case of adjective encoding and we present some strategies in order to mirror definitional features in the network, so that adjective classes emerge from the relations expressed in the database. In Section 6 we present some concluding remarks.

2 The Semantics of Hyponymy

Hyponymy is a relation which concerns not only world-knowledge, but also linguistic knowledge. Evidence for this comes from anaphoric constructions where the hypernym can be used to refer back to a more specific referent previously introduced (see [1]):

(1) He owned a rifle, but the gun had not been fired.

The fact that the relation is hierarchical in nature does not allow hypernyms and hyponyms to be contrasted, as noted in [5]:

(2) a. #A rifle is safer than a gun.
b. #He owned a rifle, but not a gun.
c. #He owned both a rifle and a gun.

Nominal hyponyms thus inherit all the information associated with the hypernym, and in turn further introduce specific semantic properties. Take for instance *school* and *bank*, two of the hyponyms of *institution*. The former is an institution which is dedicated to the teaching of students, and the latter is an institution dedicated to managing monetary funds. Although hyponymy is the main structuring relation in wordnets, there are other relations available such as meronymy and antonymy, but neither these relations nor the extended set of relations adopted in EuroWordNet are sufficiently expressive to adequately capture complex lexical semantic information.

Our goal is thus to enrich wordnets with a lexical semantics framework which allows to better describe the nature of lexical meaning as well as the specific semantic contribution made by a hyponym in relation to its hypernym. With this goal in mind, we have adopted the Generative Lexicon framework (henceforth GL, see [6]). In the approach we adopt, a synset is associated to a complex lexical description that encodes several kinds of semantic information, in particular, the specific semantic contribution of the synset as well as the meaning which is inherited from the hypernym. Such perspective has been put forth by [4], inspired by the distintion between *formal* and *telic* hypernymy in [7]. By enriching synsets with QUALIA descriptions, one can 'define' (to a reasonable extent) in what sense does one synset function as the hyponym of another. Take for instance the words *sword* and *rifle*. While man-made physical objects, both are hyponyms of *artifact*. In GL terms, *artifact* specifies properties about composition (via the FORMAL quale) which are inherited and further elaborated in *sword* and *rifle*. However, both synsets are also hyponyms of *weapon*, in the sense that these are entities devised for violent attack. In GL terms, the meaning of *weapon* is largely *agentive* and *telic*. Although the prototype of *weapon* is an object one can wield, some references are quite different: organic material can be a weapon ("*anthrax and other bio-chemical weapons*"), software ("*a computer virus is a weapon used to attack other computers* (...)"), violent coercion ("*terrorism is the political weapon of choice for some factions*"), and so on. It is not the case that ideas or molecules are weapons, but it is the case that there are in principle no incompatible properties between these concepts. [4] proposes to compute this referential possibility as a consequence of the QUALIA information associated to each synset, via an operation that integrates QUALIA features monotonically (*feature unification of* QUALIA *roles*). More specifically, two QUALIA structures are said to be *compatible* if the values of the pair-wise QUALIA roles are not inconsistent. Consider the multi-inheritance hypernymy structure of quales illustrated below:

$$\begin{bmatrix} \textit{artifact} \\ \textsc{Formal} = \boxed{1} object(x) \end{bmatrix} \qquad \begin{bmatrix} \textit{weapon} \\ \textsc{Telic} = \boxed{2} agress(e,k,w) \wedge with(e,x) \end{bmatrix}$$

$$\begin{bmatrix} \textit{sword} & \\ \textsc{Formal} & = \boxed{1} \\ \textsc{Const} & = has\_meronym(x,y) \wedge blade(y) \wedge has\_meronym(x,z) \wedge hilt(z) \; ... \\ \textsc{Telic} & = \boxed{2} \end{bmatrix}$$

Fig. 1. QUALIA Inheritance and Hypernymy

Since these particular hypernyms are relatively underspecified, the relevant inherited information only concerns the FORMAL and the TELIC roles. Note that all the information present in the hypernym must be inherited, but that the hyponym needs not be confined to it and should be able to add further information to any given quale. For instance, the TELIC role of *sword* coincides with the

Telic of *weapon* in the example above, but it may be the case that a hyponym introduces specific information in addition to the inherited properties.

This approach also allows to distinguish *compatible* from *incompatible* co-hyponyms: synsets X and Y that share the same hypernym (inheriting the same information) but where X introduces specific properties which may or not be consistent with the ones introduced by Y. For instance, *feline* and *canine* are incompatible co-hyponyms because the constitutive quale of *mammal* is extended with mutually inconsistent information about the animal's morphology (cf. [8]). Compatible co-hyponymy obtains whenever Qualia properties are orthogonally extended. E.g. some of the hyponyms of *dog* are compatible: *police dog* (extending the Telic role), and any co-hyponym extending the constitutive role, such as *german shepperd*. Another example is *lap dog* (extending the Telic role) and any co-hyponym that does not extend dog along the same dimensions, such as *poodle* (extending the constitutive role). Virtually every hyponym of *person* or *profession* is compatible with the remaining hyponyms (*man*, *sibling*, *teacher*, *witness*, *biologist*, *musician*, *lawyer*, etc). The alternative to the general approach in [12] seems to be to exhaustively mark all the pairs of compatible co-hyponyms.

However, the method proposed for determining if two given Qualia are consistent – subsumption – is too restrictive. It correctly obtains that *canine* constitutive properties are incompatible with *feline* constitutive properties, but a noun may receive more than one kind of Telic value, for instance. Nouns like *professor* and *biologist* have different Telic properties, and yet are not incompatible. In our view, Qualia should simply be conjoined rather than unified. It is world knowledge that imposes the relevant constraints: nothing needs to be said about entities having more than one function (i.e. Telic role), but different *simultaneous* physical properties along the same dimension ("short fur" and "thick fur" are orthogonal and thus not inconsistent properties, while "thick fur" and "thin fur" are inconsistent) should be prohibited. This can be achieved by background world knowledge rules. We note also that such rules may even be suppressed in hypothetical contexts ("*If square circles existed* (...)"), children stories, or in metaphorical uses.

3 Argument Structure and the Semantics of Movement Verbs

Verbal concepts are related through a hyponymy relation that refers a special subtyping relation: troponymy. Troponymy establishes a relation between verbal concepts concerning types of manner (see [9, p. 79]):

(3) *to V1 is to V2 in some particular manner*

The types of manner denoted in the verbal concepts that determine hypernym/troponym relations can be of different kinds, accounting – as for nouns – for the occurring sets of compatible co-troponym verbs, as explored in [10,11] in a decompositional approach to troponymy within the set of verbs of movement.

(4) a. He *came walking.*
b. He *exited* the house *limping.*
c. *He *walked flying.*
d. *He *exited* the house *entering.*

In the GL framework, the argument structure is the representation level for logical arguments. It is thus natural that the semantic content of a verb meaning is reflected on the number and/or the type of arguments selected. Consider for instance the verbs *to move* (change the location of), hypernym of *to put* (move into a given location), and *to box* (put into a box), troponym of *to put*:

(5) a. {*box*} @–> {*put*} @–> {*move*}

b.
$$\begin{bmatrix} \textit{move} & \\ \text{ARG-STR} = & \begin{bmatrix} \text{ARG}_1 = \textit{x: entity} \\ \text{ARG}_2 = \textit{y: entity} \end{bmatrix} \end{bmatrix}$$

c.
$$\begin{bmatrix} \textit{put} & \\ \text{ARG-STR} = & \begin{bmatrix} \text{ARG}_1 = \textit{x: entity} \\ \text{ARG}_2 = \textit{y: entity} \\ \text{ARG}_3 = \textit{z: goal} \end{bmatrix} \end{bmatrix}$$

d.
$$\begin{bmatrix} \textit{box} & \\ \text{ARG-STR} = & \begin{bmatrix} \text{ARG}_1 = \textit{x: entity} \\ \text{ARG}_2 = \textit{y: entity} \\ \text{S-ARG}_1 = \textit{z: box} \end{bmatrix} \end{bmatrix}$$

The meaning specificity of to *put* – denoting a specified GOAL – is reflected in an increase of the list of true arguments that is inherited from the hypernym to *move*. The expression of the final location (GOAL), ARG_3 – introduced by a preposition –, becomes obligatory in the case of the verb *to put.* Again, this argument structure is inherited by the immediate troponym, to *box*, that expresses a specific GOAL location, a box, through lexical shadowing, changing the predicate type of argument from true argument (ARG_3) to shadow argument (S-ARG_1). This level of representation allows us to establish more precisely what is inherited through the hierarchy and how, making it possible to account for different argument structures within a troponymy tree. Moreover, the account for compatibility issues among co-troponyms makes use of the argument structure of verbs. As stated in Section 2, two compatible nominal co-hyponyms are synsets that share the same hypernym and whose specific properties – information added to any given quale – are consistent with the properties of each other. However, this operation cannot be directly applied to the QUALIA structure of verbs. Verbal QUALIA structure is fulfilled with semantic predicates that establish the relations between the arguments of a verb. The QUALIA are also used to reflect the internal structure of the events. For instance, the agentive and formal QUALIA are typically used, respectively, to represent the causal chain and the final state for accomplishment and achievement type events. This way, the compatibility among co-troponyms is checked at argument structure level.

Considering that it is in the argument structure that the logical arguments of a predicate are listed, reflecting the added information responsible for the meaning specificities between hypernyms and troponyms, we account for co-troponym compatibility indirectly, checking for arguments' incompatibilities, recurring also to the inheritance structure in the wordnet. Following the proposal in Section 2:

(6) Two co-troponym verbs are incompatible *iff* the non-inherited arguments in their argument structure refer to incompatible co-hyponyms, i.e. if the QUALIA values of these arguments refer to opposite *simultaneous* properties along the same dimension. (see Section 2).

This indirect checking enables us to predict that the co-troponym verbs to *exit* (move out of) and to *enter* (move into) are incompatible, see (4d), since the QUALIA structure of out and in – which are co-hyponyms – are not consistent.

(7) a. $\begin{bmatrix} \textit{exit} \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = \textit{x: entity} \\ \text{ARG}_2 = \textit{y: location} \\ \text{S-ARG}_1 = \textit{z: out} \end{bmatrix} \end{bmatrix}$ b. $\begin{bmatrix} \textit{enter} \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = \textit{x: entity} \\ \text{ARG}_2 = \textit{y: location} \\ \text{S-ARG}_1 = \textit{z: in} \end{bmatrix} \end{bmatrix}$

Conversely, co-troponym verbs *exit* (move out) and *limp* (move using a leg deficiently) (in (4b)) are compatible since there are no co-hyponym arguments in their structure. The integration of an argument structure can also provide means for some syntactic mapping. In the GL, it is assumed that the type of arguments and their listing order (from less oblique to more oblique) account for the syntactic mapping of the arguments [6, pp. 62–67]. However, the representation of the arguments in these terms does not state which type of oblique argument – typically expressed by prepositional phrases – is selected by a given verb. For instance, the argument structure of the verb *to exit* in (7a) does not reveal that the ARG_2 is a prepositional phrase, nor which preposition heads this particular phrase.

It is our proposal that the use of the lexical structures available in the lexicon should make possible to state at the lexical level the exact prepositional complement selected by the verb. This proposal assumes the integration of prepositions in the lexicon, following [12] that states that the semantic contribution of prepositional phrases is consistent across uses, regardless of their status as complements or adjuncts. Prepositional lexical entries allow to account for the semantic contribution of the prepositional phrase in sentences such as "*He pulled the box* from here", as well as when the semantic content of the preposition is part of the semantic content of the verb itself, as in the case of the verb *to exit*.

4 Event Structure and the Semantics of Telic Verbs

The semantics of telic verbs involves a change of state of their theme argument. In other words, the sub-event that closes the whole event is an atomic event (i.e.

a state) that affects the theme and is different from its initial state, as briefly represented below.

(8)
$$\begin{bmatrix} \text{EVENT-STR} = [^{\sigma}_{e}\ e_1\text{: } \textit{process} <_{\propto} e_2\text{: } \textit{state}] \\ \text{ARG-STR} = \begin{bmatrix} \text{ARG}_1 = x\text{: } \textit{entity} \\ \text{ARG}_2 = y\text{: } \textit{entity} \end{bmatrix} \\ \text{QUALIA} = \begin{bmatrix} \text{AGENTIVE} = \textit{act}(e_1,\ x,\ y) \\ \text{FORMAL} = \textit{result}(e_2,\ y) \end{bmatrix} \end{bmatrix}$$

It becomes apparent from (8) that the event denoted by the verbs at stake has a typical transition type geometry, with an initial head sub-event (e_1**: process*) and a definite endpoint sub-event (e_2*: state*) which corresponds to the final state of the argument that undergoes the result of the event. In most cases e_2 is shadowed or externalized by means of a subtyping operation.

(9) a. John washed his shirt.
b. John washed his shirt white/*washed.

(10) a. John painted his house.
b. John painted his house yellow/*painted.

Sentence (9b) entails that John's shirt is white as a result of washing. Similarly, sentence (10b) entails that John's house became yellow as a result of painting. Following [13] and previous work, we assume that the constituent that expresses the result of the event denoted by the verb integrates the predicate. In other terms, the verb plus the resultative constitute a complex predicate, as extensively argued in [14]. However this is not an uncontroversial issue. As a matter of fact, despite the general assumption that resultative constructions are telic constructions (i.e. they describe events with a definite endpoint), there is a major controversy on whether or not the telic aspect of such constructions is an inherent feature of the meaning of the corresponding verbs. The compositional hypothesis, defended by [15], has been argued for in more recent works (see, for instance [16]) on the basis of contrasts like the following:

(11) a. John painted his house in one year / *for one year.
b. John painted houses *in one year / for one year.

At a first glance, these examples suggest that (11a) is telic and (11b) is atelic and, consequently, that telicity depends on the nature of the internal argument. Hence, telicity would be a compositional feature of VP and not a lexical feature of V. However, the relevant opposition seems to be transition *vs* process (in the sense of [17]) and not telic *vs* atelic aspect.

As defended in [14], though the global event in (11a) is a process, its main sub-events are not atomic events, but transitions. Let us compare the structure of the global event of (11a) and (11b), represented by (12a) and (12b), respectively (T: Transition; P: Process; e: atomic event):

(12) a. $[_{\mathrm{T}}\ [_{\mathrm{T}}\ e_1\ ...\ e_n]\ e_m]\ \ :\ \ e_m > e_n$
b. $[_{\mathrm{P}}\ [_{\mathrm{T1}}\ [_{\mathrm{P}}\ e_1^1\ ...e_n]\ e_{m1}]...\ [_{\mathrm{Tt}}\ [_{\mathrm{P}}\ e_1^t\ ...\ e_k]\ e_{m2}]\ ...]\ \ :\ \ e_{m1} > e_n, e_{m2} > e_k$

Similarly to e_m, in (12a), e_{m1} and e_{m2}, in (12b), are telic states. This suggests that, although telicity is a compositional feature regarding the whole sentence, it is also an intrinsic feature of the verb. By default, verbs like *paint* are associated to the following Lexical-Conceptual Structure (LCS' in [17]):

(13) $[_{\mathrm{T}}\ [_{\mathrm{P}}$ act(x, y) and $\sim$ Q(y)], $[_{\mathrm{e}}$ Q(y)]] : Q: atomic event

Instantiating the variables with the data of the first example above, we obtain:

(14) [[act(john, his_house) and $\sim$ painted_yellow(his_house)], [painted_yellow(his_house)]]

The absence of the resultative (*yellow*) does not have any impact on the LCS:

(15) [[act(john, his_house) and $\sim$ painted(his_house)], [painted(his_house)]]

However, in the case of verbs like to *make*, discussed below, it seems impossible to assign a value to Q independently of the resultative. Consider the sentence given below in (16). The LCS associated with it seems to be (17a) and not (17b):

(16) He made Mary happy.

(17) a. [[act(he, Mary) and $\sim$ happy(Mary)], [happy(Mary)]]
b. [[act(he, Mary) and $\sim$ made_happy(Mary)], [made_happy(Mary)]]

Therefore, Q is instantiated just with the resultative. The absence of the resultative induces ungrammaticallity, as expected:

(18) *He made Mary.

Along the same lines of [14] and [13], verbs like to *make* are defended here to be LCS deficitary, in the following sense (informal definition):

(19) $\forall$v((v a verb, $\exists\varepsilon$, ε the LCS of v, $\exists\pi$, π the set of content properties of ε, $\pi = \emptyset$) $\Rightarrow$ LCS_deficitary(v))

Since $\pi = \emptyset$, the LCS cannot bear an appropriate interpretation. A syntactic structure that projects an anomalous LCS is, then, expected to be ruled out, since it does not satisfy the commonly accepted requirement of full interpretation. In this case, the resultative fills the gap of the LCS of the verb. Therefore, these facts render evident that the representation of the predicates at issue has to include information regarding the telic expression. Obviously, it would not be adequate to overtly include in the synset all the expressions that can integrate the predicate, among other reasons, because they seem to constitute an open set. Rather, we claim that we can capture the telicity of these verbs by the inclusion of a new relation in the set of the internal relations of wordnets: the telic sub-event relation, which has two inverse counterparts, as exemplified below.

(20) {*make*} has_telic_sub-event {*state*}
{*state*} is_telic_sub-event_of {*make*}

Relating *make* to *state* by means of this relation, we capture the telic properties of the verb and let underspecified the specific nature of the final state. This way, we also account for the weakness of the verb selection restrictions. As expected, we can also use this relation to encode telicity in the case of the troponyms of the class of verbs discussed so far. Let us examine an example:

(21) a. He saddened Mary.
b. He made Mary sad.
c. *He saddened Mary sad.

Verbs like *sadden* incorporate the telic state. This fact justifies that *sadden* can be paraphrased by *make sad* ((21a) is semantically equivalent to (21b)) and cannot co-occur with *sad* (cf. (21c)). In these cases, we use the telic sub-event relation to relate the verb to the expression corresponding to the incorporated telic information:

(22) {*sadden*} has_telic_sub-event {*sad*}
{*sad*} is_telic_sub-event of {*sadden*}

It should be noticed that the existing sub-event relation in the EuroWordNet framework is different from the relation proposed here. It only stands for lexical entailment involving temporal proper inclusion. Therefore, it does not account for the geometry of the event. On the contrary, the telic sub-event relation regards the atomic sub-event that is the ending point of the global event.

As shown, the telic sub-event relation allows straightforwardly the encoding of lexical telicity in wordnets, in accordance with the empirical evidence.

5 Encoding Cross-Part-of-Speech Relations

In section 2, we focused on *hyponymy* since it is the main structuring relation in wordnets. Even if we claim here that more detailed semantic information should be introduced in computational lexica (cf. Sections 2 and 3), it is undeniable that important structural information can be extracted from the hierarchical organization of lexical items, namely of nouns and verbs. However, extending wordnets to all the main POS involves a revision of certain commonly used relations and the specification of several cross-part-of-speech relations. In this section we will focus on adjectives.

As pointed out by [1,18], the semantic organization of adjectives is unlike that of nouns and verbs, as this POS does not generally show a hierarchical organization. Thus, encoding adjectives in wordnets calls for the specification of a number of cross-part-of-speech semantic relations. In the following subsections we will present some strategies in order to mirror adjectives main features in wordnets, namely definitional ones. This way, it is possible to make adjective classes emerge from the relations expressed in the network.

5.1 Adjectives in Wordnets

In Princeton WordNet, descriptive and relational adjectives are distinguished by both being encoded in separate files and by the relations holding between synsets. Descriptive adjectives are organized in clusters of synsets, each cluster being associated by semantic similarity to a focal adjective linked with a contrasting cluster via an *antonymy* relation. Relational adjectives, on the other hand, do not have antonyms and cannot be organized in opposite clusters. Thus, relational adjectives are linked to the nouns they relate to.

[19] discusses this organization of adjectives in GermaNet. It abandons the cluster structuring of adjectives in favor of an uniform treatment of all POS in taxonomic chains. The distinct treatment of relational and descriptive adjectives is also abandoned in GermaNet, as the distinction between these two classes is considered to be 'not at all clear'. Here, along with [18] and [1], we will claim that, even if the distinction between relational and descriptive adjectives is not always clear-cut, it is however a relevant one, as these adjectives differ in terms of their intrinsic meaning, as well as with regard to their syntactic and semantic behavior. To put it somewhat simplistically, descriptive adjectives ascribe a value of an attribute to a noun. We introduce a new relation, the *characterizes with regard to/can be characterized by*,[1] linking each descriptive adjective to the attribute it modifies. Thus, instead of linking adjectives amongst themselves by a similarity relation, all adjectives modifying the same attribute are linked to the noun that lexicalizes this attribute. This way we obtain the cluster effect, argued in [18,1] to be the basis of the organization of adjectives, without having to encode it directly in the network (see [20]).

As shown by word association tests, *antonymy* is also a basic relation in the organization of descriptive adjectives. Nonetheless, this relation does not correspond to conceptual opposition: *antonymy* holds between word forms and not word meanings. We argue that conceptual opposition does not have to be explicitly encoded either, since it is possible to make it emerge from the combination of *synonymy* and *antonymy* relations as in [20]. This way, we are able to define adjective clusters without the *indirect antonymy* relation used in Princeton WordNet, as we manage to obtain the cluster effect via the *antonymy* and the *characterizes with regard to / can be characterized by* relations. In fact, our strategy is more intuitive and descriptively adequate, since many attributes are not bipolar, but can take many values along a continuum.

Concerning relational adjectives, and unlike what is done in other wordnets, we claim that these should be encoded in the same file as descriptive adjectives, avoiding having to decide beforehand whether an adjective is relational or descriptive, for instance. Rather, membership to these classes emerges from the relations expressed in the database. Being, like descriptive adjectives, property ascribing adjectives, relational adjectives usually entail more complex and diversified relations between the set of properties they introduce and the modified

[1] This semantic relation is very close to the *is a value of/attributes* relation used in Princeton WordNet. In WordNet.PT we changed its label in order to make it more straightforward to the common user.

noun, often pointing to a domain exterior to it, the denotation of another noun. We introduce the *is related to* relation to encode this.

Thus, the *characterizes with regard to / can be characterized by* and the *antonymy* relation for descriptive adjectives, and the *is related to* relation for relational adjectives, allow us to encode the basic characteristics of these adjectives in the database, on the one hand, while making it possible to derive membership to these classes from the relations expressed in the database, on the other hand.

5.2 Additional Relations

Ideally, the distinctive syntactic and semantic properties of lexical items would be encoded in lexical models such as wordnets. The SIMPLE project, for instance, addresses the semantics of adjectives (see [21]), identifying a set of features claimed to be relevant for classifying and describing their behavior. Adjectives are organized in terms of semantic fields, but these authors note that, even though similarities exist, the classes proposed in SIMPLE are not homogeneous, as adjectives belonging to the same semantic class often differ from each other in various ways.

We introduce a new relation to encode salient characteristics of nouns expressed by adjectival expressions: *is a characteristic of / has as a characteristic*. Despite the fact that we can object the status of this relation is not clear, concerning the lexical knowledge, it regards crucial information for many wordnet-based applications, namely those using inference systems, allowing for richer and clearer synsets.

Also, it may allow deducing semantic domains from the database: if synsets are encoded in this fine-grained way, it may be possible to identify the typical semantic domains of application of adjectives. The research on the classes and semantic domains emerging from the relations expressed in the database is still ongoing. Future work should include a comparative study between the classes extracted from the database and classes defined by several authors.

6 Conclusion

We have motivated the introduction of information on event and argument structures in wordnets, showing how this general approach is relevant both for allowing computational grammars to cope with a number of different lexical semantics phenomena, as well as for enabling inference applications to obtain finer-grained results. We have also proposed some new relations in order to adequately model non explicit information. Focusing on the specific case of adjective encoding in wordnets, new cross-part-of-speech relations are also introduced in order to mirror definitional features of this POS in the network and to allow for the deduction of adjective classes from the information encoded in the database.

Future work will focus on methods for the specification of the information on QUALIA, event and argument structures, ideally through new relations, in the WordNet model.

References

1. Miller, K.J.: Modifiers in wordnet. In Fellbaum, C., ed.: WordNet: an electronic lexical database. Cambridge, MA: The MIT Press (1998) 47–68
2. Fellbaum, C.: A semantic network of english: the mother of all WordNets. In Vossen, P., ed.: EuroWordNet: a Multiligual Database with Lexical Semantic Networks. Dordrecht, Kluwer Academic Publishers (1998) 137–148
3. Vossen, P.: Introduction to EuroWordNet. Computers and the Humanities (32) (1998) 73–89 Reprinted In: EuroWordNet – A Multilingual Database with Lexical Semantic Networks, Kluwer Academic Publishers (1998), 73–89.
4. Mendes, S., Chaves, R.P.: Enriching wordnet with qualia information. In: Proceedings of the Workshop on WordNets and Other Lexical Resources, NAACL 2001 Conference, Pittsburgh (2001) 108–112
5. Bever, T.G., Rosenbaum, P.S.: Some lexical structures and their empirical validity. In Jacobs, R.A., Rosenbaum, P.S., eds.: Readings in English Transformational Grammar. Waltham, Mass., Ginn (1970)
6. Pustejovsky, J.: The Generative Lexicon. MA: The MIT Press (1995)
7. Pustejovsky, J.: The Generative Lexicon. Computational Linguistics **17**(4) (1991) 409–441
8. Cruse, A.D.: Lexical Semantics. Cambridge University Press, Cambridge (1986)
9. Fellbaum, C.: A semantic network of english verbs. In Fellbaum, C., ed.: WordNet: an electronic lexical database. MA: The MIT Press (1998) 69–104
10. Amaro, R.: Semantic incorporation in a portuguese WordNet of verbs of movement: on aktionsart shifting. In: Proceedings of the Third International Workshop on Generative Approaches to the Lexicon, École de Traduction et d'Interpretation, University of Genéve (2005) 1–9
11. Amaro, R.: Wordnet as a base lexicon model for the computation of verbal predicates. In: Proceedings of the GWA 2006, Global WordNet Association Conference, Jeju Island, Korea (2006) to appear.
12. Verspoor, C.M.: Contextually-Dependent Lexical Semantics. Phd dissertation, University of Edinburgh (1997)
13. Marrafa, P.: The representation of complex telic predicates in WordNets: the case of lexical-conceptual structure deficitary verbs. In nosa, J.C., Gelbukh, A., Tovar, E., eds.: Research on Computing Science. Volume 12. (2005) 109–116
14. Marrafa, P.: Predicação Secundária e Predicados Complexos: Modelização e Análise. Phd. dissertation, University of Lisbon (1993)
15. Verkuyl, H.: On the Compositional Nature of Aspects. D. Reidel (972)
16. Schmitt, C.: Aspect of the Syntax of Noun Phrases. Phd. dissertation, University of Maryland (1996)
17. Pustejovsky, J.: The syntax of event structure. Cognition **41** (1991) 47–81
18. Fellbaum, C., Gross, D., Miller, K.: Adjectives in WordNet. In Miller et al., ed.: Five papers on WordNet, Technical Report, Cognitive Science Laboratory, Princeton University (1993) 26–39
19. Hamp, B., Feldweg, H.: GermaNet – a lexical semantic net for german. In: Proceedings of ACL workshop on Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications, Madrid (1997)
20. Mendes, S.: Adjectives in WordNet.PT. In: Proceedings of the GWA 2006, Global WordNet Association Conference, Jeju Island, Korea (2006) to appear.
21. Peters, I., Peters, W.: The treatment of adjectives in simple: theoretical observations. In: Proceedings of LREC 2000. (2000)