Wordnet-LMF: Fleshing out a Standardized Format for Wordnet Interoperability

Claudia Soria
Istituto di Linguistica Computazionale-CNR
Via Moruzzi 1 – 56126 Pisa Italy
claudia.soria@ilc.cnr.it

Monica Monachini
Istituto di Linguistica Computazionale-CNR
Via Moruzzi 1 – 56126 Pisa Italy
monica.monachini@ilc.cnr.it

Piek Vossen
Faculteit der Letteren - Vrije Universiteit Amsterdam
De Boelelaan 1105 - 1081 HV Amsterdam The Netherlands
p.vossen@let.vu.nl

ABSTRACT
In this paper we present Wordnet-LMF, a dialect of ISO Lexical Markup Framework that instantiates LMF for representing wordnets. Wordnet-LMF was developed in the framework of the EU KYOTO project for the specific purpose of endowing a set of wordnets with a standardized interoperability format allowing the interchange of lexico-semantic information encoded in each of them. The aim of this format is twofold a) to give a preliminary assessment of LMF, by large-scale application to real lexical resources; b) to endow WordNet with a format representation that will allow easier integration among resources sharing the same structure (i.e other wordnets) and, more importantly, across resources with different theoretical and implementation approaches.

Author Keywords
Standards, Lexical Markup Framework, lexical resources, wordnets, intercultural collaboration.

ACM Classification Keywords
E2 Data storage representation H.3.2 Information storage.

INTRODUCTION
Standards are a pre-requisite for interoperability of whatever kind. They are bound to be the communicative channel by means of which diverse data, resources, formats, and models can interact on a common ground, in a controlled way.

Starting in late '90s, standardized formats for lexical resource representation have now reached a high level of sophistication and theoretical consensus, with some of them attaining official international standard status, as the recent proposal for an ISO standard for encoding of lexical resources, the Lexical Markup Framework (LMF, [3]).

At the same time, the fate of a standard crucially depends on how well it is received in a community, the extent to which it gets accepted. There has been a long trail of debate concerning use of standardized representations for representing lexical resources, and the community is now starting to face the issue of usability of standards[1]. The main concern is that that despite their maturity, standards are not always or not properly used. A reason behind this is certainly publicity. A deeper, more difficult one is related to acceptability of a standard, which in turn is related to a number of factors. To be acceptable, a standard should be widely known; it should be adaptable and easy to adopt; it should be efficient (as for ease and accuracy of representation). More importantly, since converting to a standard is not an easy task and it involves deep understanding of the original source and of the model to be adopted, a standard should be useful: people will use a standard if they see a good reason to do that.

The format being presented in this paper is based on LMF, probably one of the most widely recognized standard for the representation of NLP lexicons, yet relatively little used. Wordnet-LMF will thus represent one of the first attempts at practically trying out this format, and an occasion to test LMF on a vast, real scale.

THE KYOTO PROJECT
The representation format that is the topic of this paper is being developed in the framework of the EU KYOTO project (FP7-ICT-2007-1, project nr. 211423[2]).

[1] As shown by the recent LREC 2008 Workshop "Uses and usage of language resource-related standards" (see http://www.sfb441.uni-tuebingen.de/c2/langstanduse/), and the workshop "Toward the Interoperability of Language Resources" at Stanford (http://linguistlist.org/tilr/working-group-reports/Working%20Group%202.pdf).
[2] See http://www.kyoto-project.org and http://www.kyoto-project.eu

The goal of KYOTO is to develop a system that allows people in communities to define the meaning of their words and terms in a shared Wiki platform so that it becomes anchored across languages and cultures but also so that a computer can use this knowledge to detect knowledge and facts in text. Whereas the current Wikipedia uses free text to share knowledge, KYOTO will represent this knowledge so that a computer can understand it. For example, the notion of environmental *footprint* will become defined in the same way in all these languages but also in such a way that the computer knows what information is necessary to calculate a *footprint*. With these definitions it will be possible to find information on footprints in documents, websites and reports so that users can directly ask the computer for actual information in their environment.

The focus of the project is thus on the construction of a system for facilitating the exchange of information across cultures, domains and languages. This endeavour presupposes the sharing of lexical and knowledge bases, both general and domain-related, under the form of lexical repositories and ontologies that need to be accessed both intra- and inter-linguistically.

The lexical resources that will be integrated in KYOTO are seven wordnets, for the English, Dutch, Italian, Basque, Spanish, Chinese and Japanese languages. As these resources need to be shared, linked and accessed in an integrated way, use of interoperability formats is essential.

The Wiki interface to the domain wordnet and ontology supports collaboration, editing and sharing. The domain wordnet and ontology is a plugin extension of the generic wordnet and ontology. The extensions contribute to the development of the Global Wordnet Grid (http://www.globalwordnet.org), which is an initiative to anchor many wordnets for different languages and cultures to a shared ontology backbone.

The aim of this paper is twofold a) to give a preliminary assessment of LMF, by large-scale application to real lexical resources; b) to endow wordnet with a format representation that will allow easier integration among resources sharing the same structure (i.e other wordnets) and, more importantly, across resources with different theoretical and implementation approaches.

LMF AND WORDNET REPRESENTATION
Lexical Markup Framework (LMF) is a model providing a common standardized framework for the description and representation of NLP lexicons. The goals of LMF are to provide a common model for the creation and use of such lexical resources, to manage the exchange of data between and among them, and to enable the merging of a large number of individual resources to form extensive global electronic resources.

We have chosen LMF as a representation because a wordnet is first of all a lexical repository that should be related to a database of lexical units. The focus is on words and their different meanings rather than on concepts per se. Other formats such as RDF and OWL are conceptual repositories representation formats that are not designed to represent polysemy and store linguistic properties of words and word meanings.

We leave the interested reader the opportunity to get a complete description of LMF by looking at [3], [4], [5]. Let us briefly summarize the main features of LMF, which will also help to better understand the sections that follow.

LMF was specifically designed to accommodate as many models of lexical representations as possible. Purposefully, it is designed as a *meta-model*, i.e a high-level specification for lexical resources defining the structural constraints of a lexicon.

It is organised around two main components:

1. The *core package*, i.e. a structural skeleton to represent the basic hierarchy of information in a lexicon, under the form of core classes of objects and relations.
2. A set of modular *extensions* to the core package, i.e. additional classes and relations required for the description of specific types of lexical resources. Available extensions include morphology, syntax, semantics, multilingual notations, paradigm classes, multi-word expression patterns and constraint expressions.

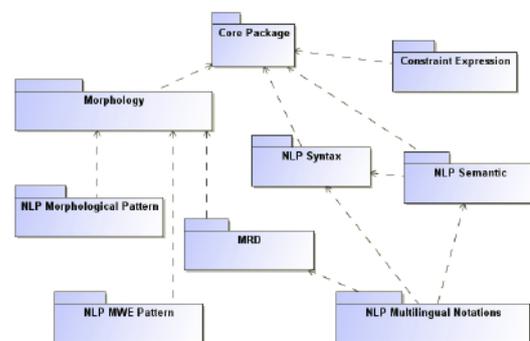The mutual dependencies among the various extensions are illustrated in Figure 1.

Figure 1. Dependencies between the LMF core and extension packages.

LMF Core package
The core lexical objects provide the basis for building LMF-compliant lexicons. *LexicalResource* is intended for representing an entire resource and, in our project, it is the container of the KYOTO wordnet grid. The KYOTO wordnet grid is a domain implementation of the Global Wordnet Grid project. Eventually, the collection of KYOTO grids will make up the modules for the overall Wordnet Grid, when domain wordnets are cumulated and integrated to the central generic repository. Each individual monolingual wordnet lexicon is an instance of the standard

Lexicon class and, on its turn, is the container for words in a given language. The *LexicalEntry* class represents an abstract unit of vocabulary: as a first approximation, it is a word. *LexicalEntry* functions as a bridge among the *Form* (or, in case, *Lemma*)[3], an abstract class representing the way a word is written (or spoken), and its related *Sense(s)*, representing one (or more) meaning(s) of a lexical entry[4].
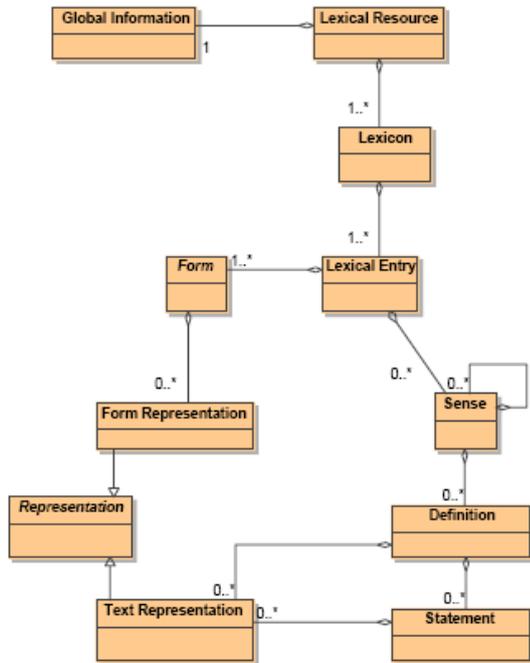

Figure 2. LMF core package.

In wordnet-like lexicons, the *LexicalEntry-Lemma-Sense* triad allows to separately account for each synset variant. *Definition* and *Statement*, tailored to refine sense instances and to facilitate human users in understanding meanings, represent in the KYOTO grid, respectively, instances of glossa and examples of use.

The implementation of Wordnet in LMF also allows for expressing (semantic) relations between specific word meanings rather than just synsets. This is less obvious in the current wordnet representations that focus on relations between synsets.

Data Categories
Data Categories (DCs hereafter) are the linguistic *constants* used to describe instances of the lexical classes, e.g. *PoS*, *noun*. DCs take the form feature structures, or attribute-value pairs used to *adorn* lexical objects, and are not part of

the LMF standard. They constitute the so-called low level specifications, that provide standardized attribute names and constants gathered in the Data Category Registry (DCR), as specified in ISO-12620, cf. [6]. In conformity to ISO philosophy, a lexicon manager, while compiling his LMF-compliant lexical resource, builds his/her own Data Category Selection, partially drawn from the DCR and partially integrated by defining a set of specific DCs needed for the representation of his/her specific resource.

LMF Semantic extension package
The representation of the semantic aspects of words is entrusted to objects related and aggregated to *Sense* class which represents lexical items as lexical semantic units. Each *Sense* instance describes one meaning of a *LexicalEntry*. *Synset* clusters synonymous *Sense* instances[5]. Both may contain information on a specific domain and a link to a semantic type in an ontology which the sense (or synset) instantiate, via the *MonolingualExternalRef* class. Here is the priviledged *locus* where the link between the wordnet grid and the conceptual resource of the project will be established. Semantic relatedness among words is an important property in natural language lexicons and, here, it is expressed through the *SenseRelation* and *SynsetRelation* classes, which encodes (lexical) semantic relationships among instances of the *Sense* or *Synset* class.

In the LMF representation, these relations are stored outside the specific lexical unit repository. Likewise, they represent the semantic relations that apply to all synonymous lexical units, grouped in a synset. It avoids duplication of information and strengthens the notion of synonymy. Implications should hold for all the synonyms.

Figure 3. Semantic extension package.

---

[3] A *Form* subclass is *Lemma*, the conventional word chosen to designate the lexical entry. *Lemma* overrides *Form*.

[4] *SyntacticBehavior*, on its turn linked to *Lexical Entry*, is the basic "brick" to account for syntactic information.
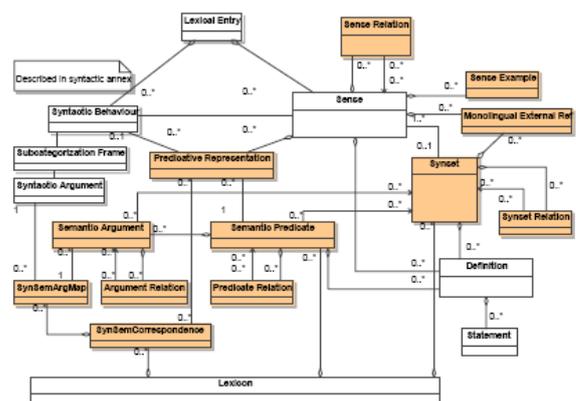
[5] It is not possible to describe *Synset* instances without any *Sense* instance.

LMF Multilingual notation package
A separate package is devoted, in LMF, to multilingual notation, which can be used to represent bilingual and multilingual resources. The framework, based on the notion of Axis, accommodates transfer, *TransferAxis*, and interlingual pivot approaches, *SenseAxis* (cf. Figure 4 below). The interlingual pivot approach, underlying the KYOTO multilingual vocation, induces to use the machinery of *SenseAxes* which are indeed the perfect connectors among nodes belonging to the different monolingual semantic packages and interlingual nodes. In conformity to LMF philosophy, the KYOTO lexical resource is to be seen as a global multilingual grid comprising *SenseAxis* instances which link monolingual *Synset* instances to interlingual nodes.

The multilingual package comes equipped with the possibility to define connections between a node in a lexicon (e.g. a *SenseAxis* instance) and knowledge representation systems, such as ontologies or fact databases, as well. This is allowed by the use of the *InterlingualExternalRef* class.
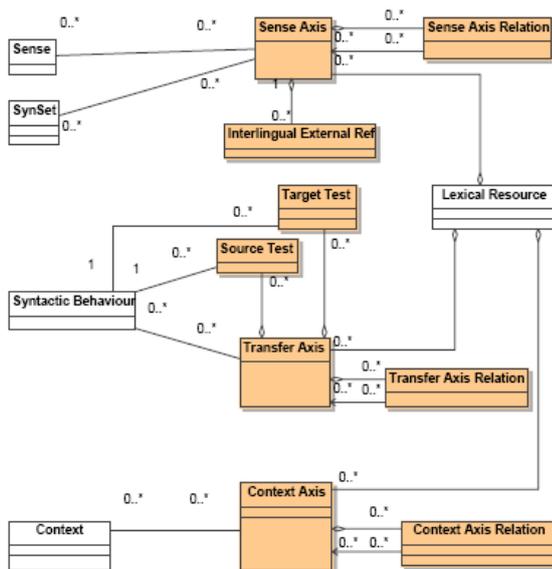

Figure 4. Multilingual notations extension.

WORDNET-LMF
Before being issued as an official ISO standard[6], LMF has passed a range of officially needed stages and has been extensively discussed and commented in a wide community comprising both academia and industry. LMF is thus mature enough to be taken as "the" choice when coming to selecting a standardized format for the representation and

encoding of computational lexicons. Time is ripe now to start assessing LMF, providing the community with real examples of use instead than preliminary examples.

Wordnet-LMF is an LMF dialect tailored to encoding of lexical resources adhering to the WordNet[7] model of lexical knowledge representation.

The WordNet lexical model represents an interesting and challenging case: although WordNet is a de-facto standard in itself, the various wordnets (i.e. the different monolingual versions adhering to the WordNet model) show a good degree of variability among them, and this would prevent immediate conversion or sharing of information.

The format presented here is going to be adopted for the encoding of seven wordnets for English, Dutch, Italian, Basque, Spanish, Chinese and Japanese. This wide spectrum of languages will allow us to take into consideration a broad range of requirements and representational constraints posed by the slightly different yet comparable contents.

As already stated above, LMF specifications are fully compatible with the structural organization of lexical knowledge encoded in wordnet-like lexical resources; actually, WordNet has been one among the pivot models that have informed the design of LMF since its very beginning. However, no real attempt has been made so far in order to fully apply LMF to wordnet-like lexicons.

Moreover, an exploration of the feasibility of LMF adoption to represent full-scale resources is an exercise still to be made. The KYOTO project will represent an ideal test case for this format: going beyond the level of toy examples it will allow to make a crash test, as the various resources will need to be fully integrated. This will put us in the position to both have a preview on any problems we might encounter and assess what acceptance would be given to LMF from a relatively closed community.

Designing Wordnet-LMF
The Wordnet-LMF format builds on the representational devices made available by LMF and tailors them to the specific content requirements of the WordNet model by adopting a user-driven approach. The design procedure of the format has undergone several distinct steps:

1. translation of some exemplifying synsets from various languages into standard LMF format;

2. qualitative assessment of the representations produced by step 1, in terms of both representational adequacy and parsing efficiency;

---

[6] LMF has been published as an ISO International Standard in November 2008. The ISO code number for LMF is ISO-24613:2008.

[7] We use *wordnet* as a generic term and leave *WordNet* (a registered name) for referring to Princeton WordNet.

3. production of a revised format on the basis of the assessment in step 2;

4. translation of synsets in all languages into the revised format.;

5. iteration of steps 2-4 until a consensus is reached.

The format presented here represents phase 3 above.

As a general comment, the purpose of the representation scheme proposed is to represent the information already present in a wordnet. Accordingly, the purpose of the exercise is to assess whether the scheme allows to do it or not, i.e. whether the structure, elements, and attributes are good enough as they are to replicate the information that is already stored in a lexical resource, without altering it, neither adding nor subtracting[8].

LMF Components
Starting form the meta-model provided by LMF, the additional packages used in Wordnet-LMF are the semantics and the multilingual extension packages.

On the basis of a review of the wordnets available in the KYOTO consortium, it turned out that the main conceptual components of WordNet-like lexicons that need to be represented in LMF are the following:

- Synsets, variants and synset relations, including information about synset identifiers and sense-keys;

- Domain attribution, linking to ontologies, administrative information;

- Interlingual information, i.e. mapping of synsets in a given language to Interlingual Index (ILI).

The semantic package naturally lends itself to the representation of wordnet-like resources, since it already contains lexical objects devised for the representation of synsets, their associated gloss and examples, variants, and synset relations.

Most wordnets also contain one or more of the following information: mapping among different versions of the same resource; reference to external or administrative information, such as mapping onto entries of another lexical database and or referencing additional sources. All these kinds of information can be dealt with by the *MonolingualExternalRef* object, which, according to LMF specifications, is an object representing a relationship between a synset instance and an external system, be it a knowledge organisation system or a terminological repository.

Interlingual information in wordnets can be represented via the LMF Multilingual Notation Extension (see [11], p. 49). This package provides a means to encode multilingual information and it is designed as an independent package,

in order not to overload the representation of monolingual lexicons. The model is based on the notion of "Axes" that link synsets pertaining to different languages. For the purposes of creating a grid of WordNets linked via Interlingual Index, the most appropriate device is the *SenseAxis* object, since it is specifically designed to implement approaches based on an interlingual pivot. Any *SenseAxis* element groups together monolingual synsets that correspond one to another by means of a particular type of relation, for instance a *synonymy* or *near_synonymy* relation.

The following is an illustration of how the SenseAxis element represents the information that three different synsets are all corresponding to the same English synset through a synonymy relation:

```
<SenseAxis      id="sa_ita16-spa30-zho30-eng30_001"
relType="eq_synonym">
<Target ID="ita-16-1251-n"/>
<Target ID="spa-30-09686541-n"/>
<Target ID="zho-30-05231501-n"/>
<Target ID="eng-30-13480848-n"/>
</SenseAxis>
```

Additional and custom components
As it should be clear from the previous section, Wordnet-LMF fully complies with standard LMF as for its major lexical objects and general framework. Expression of WordNet-related types of information (such as names of synset relations, name and values of external sources linked to wordnets) fall into the realm of LMF Data Categories, which are by definition either selectable from pre-defined standard registries or custom-defined. The Wordnet-LMF format, accordingly, has defined a number of specific information, or Data Categories, necessary to fully represent the various wordnets to be integrated in KYOTO[9]. Examples of custom Data Categories are values for describing synset relations, inter-lingual relations, for identifying external resources and their associated nodes, etc.

Wordnet-LMF wordnet format deviates from standard LMF only regarding the way data categories are instantiated: in LMF, these are represented by means of attribute-value pairs that, in an informative annex to LMF specifications, are instantiated as separate XML elements. In Wordnet-LMF wordnet format we decided to represent the same information by means of XML attributes and values instead of nested elements. This decision was motivated on the basis of better parsing efficiency. By explicitly naming the attributes, we also make a stronger claim about the features and properties of the structure of a wordnet. This will

---

[8] A preliminary description of Wordnet-LMF is available in [11].

[9] While the set of skeletal objects is fully determined, the definition of the custom data categories is still in progress.

enforce better compatibility and interoperability across the many wordnets for different languages that are available.

In this respect, the Wordnet-LMF DTD or XML Schema implementation has to be seen as dialectal variant of the LMF DTD, which, according to the specifications, is only one possible translation of the LMF model into a mark-up language ([11], p. 82).

Comparing LMF and Wordnet-LMF

For the purposes of comparison, we illustrate below an LMF and a Wordnet-LMF representation of the same Princeton WordNet 3.0 synset {footprint_1}.

```
<Synset id="eng-30-06645039-n" baseConcept="1">
<Definition gloss="mark of a foot or shoe on a surface">
<Statement example="the police made casts of the footprints in the soft earth outside the window" />
</Definition>
<SynsetRelations>
<SynsetRelation        target="eng-30-06798750-n" relType="has_hyperonym" >
<Meta        author="AH"        date="2008-07-01" source="Wordnet3.0"              status="yes" confidenceScore="1.0"/>
</SynsetRelation>
<SynsetRelation        target="eng-30-06645266-n" relType="has_hyponym" >
<Meta author="AH2" date="2008-07-01" source="eng-Wordnet3.0" status="yes" confidenceScore="1.0"/>
</SynsetRelation>
</SynsetRelations>
<MonolingualExternalRefs>
<MonolingualExternalRef externalSystem="Wordnet1.6" externalReference="eng-16-01234567-n"/>
<MonolingualExternalRef        externalSystem="SUMO" externalReference="superficialPart" relType="at"/>
</MonolingualExternalRefs>
</Synset>
```

Example 1. Wordnet-LMF format.

```
<Synset id="eng-30-06645039-n">
<feat att="baseConcept" val="1"/>
<Definition>
<feat att="gloss" val="mark of a foot or shoe on a surface"/>
<Statement>
<feat att="example" val="the police made casts of the footprints in the soft earth outside the window"/>
</Statement>
</Definition>
<SynsetRelation targets="eng-30-06798750-n">
<feat att="relType" val="has_hyperonym"/>
```

```
<feat att="confidenceScore" val="1.0"/>
<feat att="status" val="yes"/>
<feat att="source" val="Wordnet3.0"/>
<feat att="author" val="AH"/>
<feat att="date" val="2008-07-01"/>
</SynsetRelation>
<SynsetRelation targets="eng-30-06645266-n">
<feat att="relType" val="has_hyponym"/>
<feat att="confidenceScore" val="1.0"/>
<feat att="status" val="yes"/>
<feat att="source" val="Wordnet3.0"/>
<feat att="author" val="AH"/>
<feat att="date" val="2008-07-01"/>
</SynsetRelation>
<MonolingualExternalRef>
<feat att="externalSystem" val="SUMO"/>
<feat                     att="externalReference" val="superficialPart"/>
<feat att="relType" val="at"/>
<feat att="externalSystem" val="Wordnet1.6"/>
<feat     att="externalReference"     val="eng-16-01234567-n"/>
</MonolingualExternalRef>
</Synset>
```

Example 2. LMF format.

CONCLUSIONS

The work presented here is work in progress. While being the result of a long debate on how to best represent lexical resources on the one hand and wordnets on the other, it still awaits further testing, most notably an evaluation of its resilience to be used as a working encoding format for storing and access of lexical information in a dedicated database.

Some considerations, however, are allowed even at this preliminary stage, especially regarding usability of LMF.

LMF is, admittedly, a "high-level" specification, that is, an abstract model that needs to be further developed, adapted and specified by the lexicon encoder. LMF does not provide any off-the-shelf representation for a lexical resource; instead, it gives the basic structural components of a lexicon, leaving full freedom for moulding the model to suit the particular features of lexical resources. The drawback of this is that one is left with a specification manual and a few examples. Specifications are by no means instructions, exactly as XML specifications are by no means instructions on how to represent a particular type of data.

Going from LMF specifications to true instantiation of an LMF compliant lexicon is a long way, and the need is felt for comprehensive, illustrative and detailed examples for doing this. In a painstaking search for guidelines, LMF is

often mistakenly taken as a prescriptive description, and the examples contained therein as pre-defined normative examples to be used as coding guidelines. Controlled and careful examples of conversion to LMF compliant formats is also needed to avoid too subjective interpretations of the standard (for similar considerations, see also [1]).

We further believe that the development of Wordnet-LMF paves the way to a number of expected results, both from the point of view of LMF and from the point of view of the WordNet community.

From the point of view of LMF, Wordnet-LMF will:

- demonstrate adaptability of LMF to representation of wordnets;
- promote adoption of LMF to a wider community;
- be one of the first testbed for LMF (as one of its drawback being that it has not been tested on a wide variety of lexicons), particularly relevant since it is related to both Western and Eastern language wordnets;
- specify an LMF-compliant XML format, tested for representative and parsing efficiency;
- provide guidelines for the implementation of an LMF compliant format, thus contributing to the reduction of subjectivity in interpretation of standards.
- work as a crash-test of the multilingual notation package, where the explosion of links among the various monolingual wordnets will allow to assess the viability of this type of representation on a real scale.

From the point of view of wordnets:

- it will provide a format for exchange of information across wordnets and between WordNet-like and differently conceived lexicons. The WordNet model is probably the most widespread model of representation of lexical knowledge, at least in the NLP community, but also outside. WordNet-like resources can thus be endowed with a standardized format representation for relating them to other lexical models, in a rigorous and linguistically controlled way. This seems an important and promising achievement in order to move the sector forward.
- Conversion of Wordnet to LMF was straight-forward. No major problems have been encountered representing the data of a whole range of wordnets.
- Once tested at the relatively local level of the KYOTO grid, Wordnet-LMF will be a candidate format for adoption inside the Global WordNet Grid initiative[10] (see [2]);
- Another lexical grid is being built for Asian languages, integrating lexical resources different from the WordNet model but still interlinked through an interlingual pivot approach. In this grid, developed in a

---

[10] www.globalwordnet.org/gwa/gwa_grid.htm

project under the NEDO International Joint Research Grant Program (NEDO Grant, [12]), the lexical resources are encoded by means of an LMF compliant format. The Wordnet-LMF format will serve as representational bridge to evaluate the needs and problems posed by making two lexical grids interoperable.

- Compose web services for lexicon access functions. These services which are especially tailored for wordnet-like lexicons, since grounded on LMF, can be seen as atomic pieces able to be combined and integrated into the grid of composite lexicon services based on the LMF metamodel [7] to be made available in the global language infrastructure of the Language Grid project [8].

In the near future, we will further investigate more complex ways in which the wordnets in LMF can be related to ontologies. We will investigate the current proposal for LexInfo (reference to Paul Buitelaar's proposal). We will also integrate the proposals made in the Dutch Cornetto project [13], in which the mapping of the Dutch Wordnet to SUMO has been developed and which represents an extension to the way SUMO is now related to the English WordNet.

REFERENCES

1. Bel N. and S. Bel. 2008. "Measuring Standards in Lexical Resources". In *Proc. LREC 2008 Workshop on Uses and usage of language resource-related standards*, ELRA (2008), 15-19.

2. Fellbaum C., Vossen P. Connecting the Universal to the Specific: Towards the Global Grid. In *Proc. IWIC 2007* (2007), 2-16.

3. Francopoulo G., George M., Calzolari N., Monachini M., Bel N., Pet M., Soria C. Lexical Markup Framework (LMF). In *Proc. LREC 2006*, ELRA (2006), 233-236.

4. Francopoulo G., Bel N., George M., Calzolari N., Monachini M., Pet M., Soria C. Lexical Markup Framework: an ISO Standard for Semantic Information in NLP Lexicons. In *Proc. Workshop on Lexical-Semantic and Ontological Resources of the GLDV Working Group on Lexicography*, (2007).

5. Francopoulo G., Bel N., George M., Calzolari N., Monachini M., Pet M., Soria C. Lexical Markup Framework (LMF) for NLP multilingual resources. In *Proc. COLING-ACL Workshop on Multilingual Lexical Resources and Interoperability*, ACL (2006), 1-8.

6. Francopoulo G., Monachini M., Declerck T., Romary L. Morphosyntactic The relevance of Standards for Research Infrastructure. In *Proc.LREC 2006 Workshop Towards Research Infrastructures for Language Resources*, ELRA (2006), 19-22.

7. Hayashi, Y., Narawa, C., Monachini, M., Soria, C., and Calzolari, N. Ontologizing Lexicon Access Functions based on a LMF-based Lexicon Taxonomy. In *Proc. LREC 2008*, ELRA (2008).

8. Ishida, T. 2006. Language Grid: An Infrastructure for Intercultural Collaboration. In *Proc. IEEE/IPSJ Symposium on Applications and the Internet* (2006), 96-100.

9. ISO 24613:2008 Language Resource Management – Lexical Markup Framework, ISO Geneva, 2008.

10. Pease A., Fellbaum, C., Vossen, P. 2008 Building the Global WordNet Grid. In Proc. CIL18 (2008).

11. Soria C. and M. Monachini. 2008. "Kyoto-LMF. Wordnet representation format". KYOTO Working Paper WP02_TR002_V03.

12. Tokunaga T. et al. "Infrastructure for standardization of Asian language resources". In *Proc. COLING/ACL 2006 Main Conference Poster Sessions*, ACL (2006), 827–834.

13. Vossen P., Maks I., Segers R., VanderVliet H. Integrating lexical units, synsets and ontology in the Cornetto Database. In *Proc. LREC 2008*, ELRA (2008).