Predicate Matrix: extending SemLink through WordNet mappings

Maddalen Lopez de Lacalle, Egoitz Laparra, German Rigau
IXA group UPV/EHU, IXA group UPV/EHU, IXA group UPV/EHU
Donostia,Spain, Donostia,Spain, Donostia,Spain
{maddalen.lopezdelacalle, egoitz.laparra, german.rigau}@ehu.com

Abstract
This paper presents the Predicate Matrix v1.1, a new lexical resource resulting from the integration of multiple sources of predicate information including FrameNet (Baker et al., 1997), VerbNet (Kipper, 2005), PropBank (Palmer et al., 2005) and WordNet (Fellbaum, 1998). We start from the basis of SemLink. Then, we use advanced graph-based algorithms to further extend the mapping coverage of SemLink. Second, we also exploit the current content of SemLink to infer new role mappings among the different predicate schemas. As a result, we have obtained a new version of the Predicate Matrix which largely extends the current coverage of SemLink and the previous version of the Predicate Matrix.

Keywords: Verbal Lexicon, WordNet, VerbNet, FrameNet, PropBank

1. Introduction
Predicate models such as FrameNet (Baker et al., 1997), VerbNet (Kipper, 2005) or PropBank (Palmer et al., 2005) are core resources in most advanced NLP tasks, such as Question Answering, Textual Entailment or Information Extraction. Most of the systems with Natural Language Understanding capabilities require a large and precise amount of semantic knowledge at the predicate-argument level. This type of knowledge allows to identify the underlying typical participants of a particular event independently of its realization in the text. Thus, using these models, different linguistic phenomena expressing the same event, such as active/passive transformations, verb alternations, nominalizations, implicit realizations can be harmonized into a common semantic representation. Lately, several systems have been developed for shallow semantic parsing an explicit and implicit semantic role labeling (SRL) using these resources (Erk and Pado, 2004), (Shi and Mihalcea, 2005), (Giuglea and Moschitti, 2006), (Laparra and Rigau, 2013).

However, building large and rich enough predicate models for broad-coverage semantic processing takes a great deal of expensive manual effort. Furthermore, the same effort should be invested for each different language (Subirats and Petruck, 2003).

Most previous research efforts on the integration of lexical resources targeted at knowledge about nouns and named entities rather than predicate knowledge. Well known examples are YAGO (Suchanek et al., 2007), Freebase (Bollacker et al., 2008), DBPedia (Bizer et al., 2009) or BabelNet (Navigli and Ponzetto, 2010).

Following the line of previous works (Shi and Mihalcea, 2005), (Burchardt et al., 2005), (Crouch and King, 2005), (Johansson and Nugues, 2007), (Pennacchiotti et al., 2008), (Cao et al., 2008), (Tonelli and Pianta, 2009), (Laparra et al., 2010), (Necsulescu et al., 2011), (Gurevych et al., 2012) we will also focus on the integration of predicate information. We start from the basis of SemLink (Palmer, 2009) despite its coverage is still far from being complete (López de Lacalle et al., 2014). First, we use advanced graph-based algorithms to further extend the mapping coverage of SemLink. Second, we also exploit the current content of SemLink to infer new role mappings among the different predicate schemas. As a result, we have obtained a new version of the Predicate Matrix which largely extends the current coverage of SemLink. For example, SemLink provides 6,201 mappings between VerbNet and FrameNet roles while the current version of the Predicate Matrix contains 25,688 additional mappings. This paper is organized as follows. Section 2. presents the set of sources of predicate information used for developing the current version of the Predicate Matrix. Section 3. summarizes the mapping coverage of those resources integrated in SemLink. Section 4. details the process for building the current version of the Predicate Matrix, the particular case of how the mappings with PropBank are affected is showed in 5.. Section 6. provides further details of the new predicate result. Finally, Section 7. presents some concluding remarks and our current plans for future work.

2. Sources of Predicate information
We used the following resources to create the first version of the Predicate Matrix. SemLink 1 (Palmer, 2009) is a project whose aim is to link together different predicate resources via set of mappings. Currently, SemLink provides partial mappings between FrameNet (Baker et al., 1997), VerbNet (Kipper, 2005), PropBank (Palmer et al., 2005) and WordNet (Fellbaum, 1998). These mappings make it possible to combine their information for tasks such as inferencing, consistency checking, interoperable SRL, or automatic extending its current overlapping coverage. VerbNet 2 (Kipper, 2005) hierarchical domain-independent broad-coverage verb lexicon for English. VerbNet is organized into verb classes. Each verbal class in VerbNet is completely described by thematic-roles, selectional restrictions on the arguments, and frames consisting of a syntactic description and semantic predicates.

1http://verbs.colorado.edu/semlink/
2http://verbs.colorado.edu/~mpalmer/projects/verbnet.html
FrameNet \(^3\) (Baker et al., 1997) is a very rich semantic resource that contains descriptions and corpus annotations of English words following the paradigm of Frame Semantics (Fillmore, 1976). In frame semantics, a Frame corresponds to a scenario that involves the interaction of a set of typical participants, playing a particular role in the scenario. FrameNet groups words or lexical units (LUs hereinafter) into coherent semantic classes or frames, and each frame is further characterized by a list of participants or frame-elements (FEs hereinafter). Different senses for a word are represented in FrameNet by assigning different frames. PropBank \(^4\) (Palmer et al., 2005) aims to provide a wide corpus annotated with information about semantic propositions, including relations between the predicates and their arguments. PropBank also contains a description of the frame structures, called framesets, of each sense of every verb that belong to its lexicon. Unlike other similar resources, PropBank defines the arguments, or roles, of each verb individually. In consequence, it becomes a hard task obtaining a generalization of the frame structures over the verbs.

WordNet \(^5\) (Fellbaum, 1998) is by far the most widely-used lexical knowledge base. In fact, WordNet is being used worldwide for anchoring different types of semantic knowledge including wordnets for languages other than English (Gonzalez-Agirre et al., 2012). It contains information about English nouns, verbs, adjectives and adverbs and is organized around the notion of a synset. A synset is a set of words with the same part-of-speech that can be interchanged in a certain context. For example, `<learn, study, read, take>` form a synset because they can be used to refer to the same concept. A synset is often further described by a gloss, in this case: "be a student of a certain subject" and by explicit semantic relations to other synsets. Each synset represents a concept that are related with an large number of semantic relations, including hypernym/hyponymy, meronymy/holonymy, antonymy, entailment, etc. As in (López de Lacalle et al., 2014), we use these resources to create a new version of the Predicate Matrix. Our current goal is to exploit all these resources to extend their current coverage, to discover inherent inconsistencies among these resources, to enrich WordNet with predicate information, and possibly to extend this predicate information to languages other than English (by exploiting the local wordnets aligned to the English WordNet).

3. Incomplete coverage of SemLink
(López de Lacalle et al., 2014) studies the coverage of the different mappings encoded in SemLink. For instance, the alignment between FrameNet and VerbNet proves to be the least complete one. Only 1,730 LUs of FrameNet are aligned to, at least, one VerbNet predicate. This number represents only the 16\% of the total 10,195 LUs of FrameNet. Moreover, not only the lexicon but the role sets of both resources are weakly connected. For instance, just 825 of the 7,124 existing FEs of FrameNet are linked to a VerbNet thematic-role. That is, 88\% of the FEs of FrameNet are not aligned to any VerbNet thematic-role. The lexicon mapping between PropBank and VerbNet is also incomplete. From the 6,181 different PropBank predicates, 2,623 have no connection to VerbNet. Regarding the PropBank arguments and the VerbNet thematic-roles, around a half of the total PropBank arguments (7,915 out of 15,871 arguments) are mapped to a thematic-role from VerbNet. Moreover, SemLink does not provide a complete alignment between WordNet and VerbNet. Specifically there are 18,559 verbal senses of WordNet, corresponding to 9,995 different lemmas, that have not been assigned to any VerbNet predicate.

4. Creating the Predicate Matrix
In this work we present a new method to complete and extend the coverage of the mappings between the resources included in SemLink. Particularly, we have focused on the links that connect VerbNet, FrameNet and WordNet. The process, explained in this section, starts from SemLink. Then, a set of sequential steps try to complete the alignments. The whole process can be devided into three consecutive steps:

1. Complete and extend the mappings between the lexicons of WordNet, VerbNet and FrameNet. Following (Laparra and Rigau, 2009; Laparra et al., 2010; Laparra and Rigau, 2013), we apply knowledge-based Word Sense Disambiguation (WSD) algorithms that use a large-scale graph of concepts derived from WordNet to disambiguate the verbs (and also nouns, adjectives and adverbs corresponding to the FrameNet LUs) from both lexicons. Then, for each WordNet verb sense, we collect the desambiguations (and alignments) to FrameNet frames and VerbNet classes. This is explained in section 4.1.

2. Complete the mappings between VerbNet thematic-roles and FrameNet frame elements. For many cases, although there will be a mapping between lemmas, the corresponding links between the roles will be missing. We apply some methods that use the existing mappings and knowledge from the resources to complete these gaps. A detailled explanation of these methods is shown in section 4.2.

3. Following (López de Lacalle et al., 2014) we finally extend the mappings via WordNet synonyms. Section 4.3. describes the results of this process.

4.1. Completing the lexicon mappings
The first step for extending SemLink is completing the mapping between the lexicon of VerbNet and the lexicon of FrameNet. Following (Laparra and Rigau, 2009; Laparra et al., 2010; Laparra and Rigau, 2013), we apply knowledge-based Word Sense Disambiguation (WSD) algorithms that use a large-scale graph of concepts derived from WordNet to disambiguate the verbs (and also nouns, adjectives and adverbs corresponding to the FrameNet LUs) from both

---

\(^3\)http://framenet.icsi.berkeley.edu/
\(^4\)http://verbs.colorado.edu/-mpalmer/projects/ace.html
\(^5\)http://wordnet.princeton.edu/
lexicons that are not already linked to WordNet in Sem-
Link. The WSD algorithms are applied to coherent group-
ings of words belonging to the same FrameNet frame or
VerbNet class. The disambiguation provides new links be-
tween those verbs and the WordNet senses. Thus, we can
connect verbs from different resources that are connected
to the same WordNet sense.

Although FrameNet covers more than 10,000 LUs and 795
frames, only 721 frames have associated a LU. From those,
10,086 LUs (word-frame pairs) where recognized by Word-
Net (out of 92%) corresponding to 708 frames and 2,867
verbs. In FrameNet, the LUs of a frame can be nouns,
verbs, adjectives and adverbs representing a coherent and
 closely related set of meanings that can be viewed as a small
semantic field. For example, the frame Education contains LUs referring to the educational activity and their
participants. It is evoked by LUs like cram, instruction, instruc-
tor, learn, lecturer, study, etc. The frame also defines core semantic roles (or FEs) such as STUDENT or SUBJECT that are semantic participants of the frame and their corresponding LUs.

VerbNet also groups semantically related verbs. VerbNet
contains 4,403 verbs in 386 classes and subclasses. From
those, 6,078 verbal senses (verb-class pairs) where recog-
nized by WordNet (out of 97%). For instance, the Verb-
Net class learn-14 groups together the verbs like assimilate, cram, glean, learn, memorize or read. This VerbNet
class also defines a set of thematic-roles: Agent, Source and
Topic.

We tested two different graph-based Word Sense Disam-
biguation algorithms. An advanced version of the Struc-
tural Semantic Interconnections algorithm (SSI) (Navigli
and Velardi, 2005) called SSI-Dijkstra+ (SSID+) (Cuadros
and Rigau, 2008; Laparra and Rigau, 2013) and UKB
(Agirre and Soroa, 2009). As SemLink also includes
some manual assignments of WordNet senses with VerbNet
and FrameNet, we have been able to evaluate the accu-
tracy of the automatic mapping. For the evaluation, we
used as gold-standard 272 VerbNet classes and their associ-
ated verbs and 214 FrameNet frames having at at least one
WordNet sense manually assigned to a verb. The average
length of the contexts is 23.30 verbs for VerbNet and 19.38
LUs for FrameNet.

Table 1 presents the precision (P), recall (R) and F1 mea-
sure (harmonic mean of recall and precision) of the differ-
ent methods and knowledge resources when mapping
WordNet to VerbNet and FrameNet. WN stands for a Lex-
ical Knowledge Base (LKB) built using only the relations
from WordNet while WN+G refers to the LKB also inte-
grating the relations from the semantically tagged glosses6.

Verbal lemmas of VerbNet and FrameNet (VN-FN) in-
crease in almost 1,500 new cases. Around 1,000 new map-
pings connect WordNet senses with VerbNet lemmas (VN-
WN) while there are 2,000 new alignment with FrameNet
verbs (FN-WN). Finally, the full connections between the
three resources (VN-WN-FN) increase in more than 1,500
new alignments.

4.2. New role alignments

The second step focuses on obtaining the missing corre-
spondances between the semantic roles from VerbNet and
FrameNet. The missing links can belong to verbs that were
previously included in SemLink or can belong to the new
verb senses that have been included in the previous step.

Table 2: Number of alignments discovered by
each method. The table shows how each step increments
the number of cases covered by the previous method and
also includes the individual evaluation of the methods ap-
plied.

<table>
<thead>
<tr>
<th>Method</th>
<th>New</th>
<th>Total</th>
<th>P</th>
<th>R</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>SemLink</td>
<td>6,201</td>
<td>6,201</td>
<td>88.2</td>
<td>88.2</td>
<td>88.2</td>
</tr>
<tr>
<td>Method 1</td>
<td>6,686</td>
<td>12,887</td>
<td>88.2</td>
<td>88.2</td>
<td>88.2</td>
</tr>
<tr>
<td>Method 2</td>
<td>1,088</td>
<td>13,975</td>
<td>76.0</td>
<td>48.6</td>
<td>59.3</td>
</tr>
<tr>
<td>Method 3</td>
<td>1,193</td>
<td>15,168</td>
<td>80.6</td>
<td>80.6</td>
<td>80.6</td>
</tr>
</tbody>
</table>

Table 3: New links added to the mapping between the lexicons.

| New links | 4,712 | 8,504 | 6,338 | 6,745 |

Table 3: Results of the disambiguation process.

<table>
<thead>
<tr>
<th>Source</th>
<th>WN+G</th>
<th>WN</th>
<th>WN+G</th>
<th>WN</th>
</tr>
</thead>
<tbody>
<tr>
<td>VerbNet</td>
<td>85.3</td>
<td>85.3</td>
<td>85.3</td>
<td>85.3</td>
</tr>
<tr>
<td>FrameNet</td>
<td>83.8</td>
<td>83.7</td>
<td>83.7</td>
<td>83.7</td>
</tr>
</tbody>
</table>

6 http://wordnet.princeton.edu/glosstag.shtm

VerbNet Method LKB P R F1
UKB WN 84.2 84.2 84.2
UKB WN+G 85.3 85.3 85.3
SSID+ WN 83.8 83.5 83.7
SSID+ WN+G 83.8 83.5 83.7

FrameNet Method LKB P R F1
UKB WN 79.0 79.0 79.0
UKB WN+G 79.4 79.4 79.4
SSID+ WN 82.5 81.3 81.9
SSID+ WN+G 82.9 81.8 82.4

Table 2: New links added to the mapping between the lexicons.

tween verb lemmas of VerbNet and FrameNet (VN-FN) in-
crease in almost 1,500 new cases. Around 1,000 new map-
pings connect WordNet senses with VerbNet lemmas (VN-
WN) while there are 2,000 new alignment with FrameNet
verbs (FN-WN). Finally, the full connections between the
three resources (VN-WN-FN) increase in more than 1,500
new alignments.

4.2. New role alignments

The second step focuses on obtaining the missing corre-
spondances between the semantic roles from VerbNet and
FrameNet. The missing links can belong to verbs that were
previously included in SemLink or can belong to the new
verb senses that have been included in the previous step.

Table 3 contains the number of alignments discovered by
each method. The table shows how each step increments
the number of cases covered by the previous method and
also includes the individual evaluation of the methods ap-
plied.

Method 1: The first method is based on learning from Sem-
Link which alignments between VerbNet thematic-roles
and FrameNet frame-elements are more frequent. For every
verb of VerbNet aligned to a frame of FrameNet we obtain
the thematic-roles that have not been assigned to any frame-
element. Then, we link each of these roles with the FE of
the frame that is most frequently aligned in other cases. For
example, the verb paddle of the VerbNet class spank-18.3
is mapped to the frame Corporal_punishment of FrameNet.
However, the role Location of this verb is not linked to any FE. The frame Corporal.punishment contains frame-elements like Agent, Evaluate, Reason, Instrument, Degree and Body_part. According to the data showed in table 4, Body_part is the FE of the frame Corporal.punishment that is mapped to the thematic-role Location in a greater number of times. In table 5 we include this new mapping and some other cases obtained by this method.

<table>
<thead>
<tr>
<th>Thematic-Role</th>
<th>FrameElement</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Area</td>
<td>285</td>
</tr>
<tr>
<td>Location</td>
<td>Goal</td>
<td>228</td>
</tr>
<tr>
<td>Location</td>
<td>Path</td>
<td>99</td>
</tr>
<tr>
<td>Location</td>
<td>Sound_source</td>
<td>73</td>
</tr>
<tr>
<td>Location</td>
<td>Ground</td>
<td>54</td>
</tr>
<tr>
<td>Location</td>
<td>Source</td>
<td>49</td>
</tr>
<tr>
<td>Location</td>
<td>Location</td>
<td>23</td>
</tr>
<tr>
<td>Location</td>
<td>Body_part</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 4: Frequency of frame-elements mapped to the thematic-role Location.

Method 2: Although the Method 1 obtains very reliable outcomes, it leaves out cases not included in SemLink. For instance, the verb feel of the class see-30.1 is mapped to the frame Seeking, but none of its thematic-roles, Experiencer and Stimulus, has been previously linked to any of the FEs that are part of the frame Seeking. For these kind of cases, this method obtains from the class files of VerbNet the patterns of the examples included, and it does the same for each frame with the annotation examples included in the FrameNet files. After comparing the most frequent ones, the method aligns the thematic-roles and the FEs that share the same positions. Following the example of the verb feel, for the class see-30.1 the method just finds examples that follow the pattern Experiencer - verb - Stimulus as it is shown in table 6.

According to the table, the most frequent pattern in FrameNet for the frame Seeking is Cognizer_agent - verb - Sought_entity - Ground. Thus, as table 7 shows, the method links the thematic-role Experiencer with the FE Cognizer_agent and Stimulus with Sought_entity because they tend to appear in the same position.

Method 3: For the last step we re-implement the same strategy that for Method 1, but this time the learning of the frequency of the mappings includes the new links obtained by the previous two methods. As can be seen in table 8, in this case, the frequency of the mappings between frame-elements and thematic-roles are different to those showed in table 4. For example, the FrameElement Place did not seem to be very frequently linked to the thematic-role Location in the original version of SemLink. But, after applying Method 2, the number of cases where Place and Location have been related increases remarkably, making it more likely to find new cases of this mapping.

For the evaluations showed in table 3 we have used as testing set the existing 6,201 SemLink role alignments. The evaluation process has been the same as the one used for the lexicon mappings. For each role mapping we apply a leave-one-out evaluation process. We learn the frequencies from the whole SemLink except the one we are evaluating. This process, applied for the three methods explained above, allows to use the full set of role mappings from SemLink as a gold-standard.

4.3. Adding WordNet synonyms

(López de Lacalle et al., 2014) shows how the alignments to WordNet offer a very interesting source of information to be systematically exploited. A simple automatic method to extend SemLink by exploiting properties from WordNet consists on including the synonyms of already aligned WordNet senses. Obviously, this method expects that WordNet synonyms share the same predicate information. For instance, the predicate desert_v, member of the VerbNet class leave-51.2-1, with a link to the frame Departing of FrameNet, appears to be assigned to desert%2:31:00 WordNet verbal sense. In WordNet, this word sense also has three synonyms, abandon_v, forsake_v, and desolate_v. These three verbal senses can also be assigned to the same VerbNet class. This simple approach extends the amount of alignments not only between lexicons but also between roles. Table 9 shows how the synonymy extension offers a great improvement on the coverage.

<table>
<thead>
<tr>
<th>Thematic-Role</th>
<th>FrameElement</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Area</td>
<td>341</td>
</tr>
<tr>
<td>Location</td>
<td>Goal</td>
<td>213</td>
</tr>
<tr>
<td>Location</td>
<td>Place</td>
<td>148</td>
</tr>
<tr>
<td>Location</td>
<td>Path</td>
<td>145</td>
</tr>
<tr>
<td>Location</td>
<td>Ground</td>
<td>111</td>
</tr>
<tr>
<td>Location</td>
<td>Source</td>
<td>83</td>
</tr>
<tr>
<td>Location</td>
<td>Sound_source</td>
<td>78</td>
</tr>
<tr>
<td>Location</td>
<td>Location</td>
<td>71</td>
</tr>
</tbody>
</table>

Table 8: Frequency of frame-elements mapped to the thematic-role Location including automatic links from Method 1 and Method 2.

5. Mapping to PropBank

The integration of the new mappings between VerbNet and FrameNet roles described in previous section also helps to complete the mappings to PropBank roles. This work have focused on the mapping between VerbNet and FrameNet, and their connection to WordNet. However, we can extend the resulting new connections for those cases from PropBank that are currently mapped to any VerbNet and FrameNet element affected by the processes described in this paper. Table 10 shows the old and new mappings to PropBank.
lemmas VN-class Thematic-Role FN-frame FrameElement
sit spatial_configuration-47.6 Location Placing Area
spew substance_emission-43.4 Location Excreting Goal
move roll-51.3.1 Location Change_position_on_a_scale Path
paddle spank-18.3 Location Corporal_punishment Body_part

Table 5: Examples of new frame-elements mapped to the thematic-role Location.

<table>
<thead>
<tr>
<th>Source</th>
<th>Class/Frame</th>
<th>Pattern</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VerbNet</td>
<td>see-30.1</td>
<td>Experiencer v Stimulus</td>
<td>100%</td>
</tr>
<tr>
<td>FrameNet</td>
<td>Seeking</td>
<td>Cognizer_agent v Sought_entity Ground</td>
<td>80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sought_entity v Ground Cognizer_agent</td>
<td>20%</td>
</tr>
</tbody>
</table>

Table 6: Frequency of role patterns in VN class see-30.1 and frame Seeking.

| lemma VN-class Thematic-Role FN-frame FrameElement |
|---------|----------------|-----------------|-----|

Table 7: Examples of new mappings between thematic-roles and frame-elements of the frame Seeking.

<table>
<thead>
<tr>
<th>Source</th>
<th>Class/Frame</th>
<th>Pattern</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VerbNet</td>
<td>see-30.1</td>
<td>Experiencer v Stimulus</td>
<td>100%</td>
</tr>
<tr>
<td>FrameNet</td>
<td>Seeking</td>
<td>Cognizer_agent v Sought_entity Ground</td>
<td>80%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sought_entity v Ground Cognizer_agent</td>
<td>20%</td>
</tr>
</tbody>
</table>

Table 10: Number of new lexicon and role alignments for PropBank.

6. Predicate Matrix v1.1

We produced a new version of the Predicate Matrix by exploiting SemLink and applying advanced WSD methods to extend and validate its content. Each row of this Predicate Matrix represents the mapping of a role over the different resources and includes all the aligned knowledge about its corresponding WordNet verb sense. As shown in table 9 the new version of the Predicate Matrix (PM v1.1) contains much more alignments than SemLink.

For instance, as the Predicate Matrix uses the verbal part of WordNet as a backbone, now we also know how much of WordNet is still not covered by VerbNet, FrameNet or PropBank. For instance, from the total number of 25,148 WordNet verbal senses, the new version of the Predicate Matrix only contains 11,629 WordNet verb senses aligned to VerbNet classes. That is, there are 13,997 WordNet verb senses still without mappings to VerbNet classes. Similarly, the Predicate Matrix now only contains 7,573 WordNet senses aligned to FrameNet frames. Thus, there are 18,672 WordNet word senses without mappings to FrameNet frames.

Table 11: WordNet verbal senses not covered by VerbNet classes and FrameNet frames in the Predicate Matrix v1.1.

As an example of the current coverage of the Predicate Matrix v1.1, table 11 shows the distribution according to the

---

5 The Predicate Matrix can be obtained from http://adimen.si.ehu.es/web/PredicateMatrix
lexicographic files from WordNet of the verbal senses not covered by VerbNet classes and FrameNet frames in the Predicate Matrix v1.1. From left to right, the table shows the lexicographic file number, the number of verb senses pertaining to the lexicographic file, the number (and percentage) of verb senses not aligned to a VerbNet class, the number (and percentage) of verb senses not aligned to FrameNet frames and the lexicographic file name. Interestingly, the coverage of both resources are quite different depending on the area of WordNet selected. The VerbNet coverage ranges from emotion verbs (it remains only 29.88% of WordNet verb senses to be complete) up to competition verbs (70.67%) whereas FrameNet coverage ranges from weather (54.79%) up to change (81.35%).

7. Conclusions and future work
This is an ongoing work towards a more complete version of the Predicate Matrix. We current version of the Predicate Matrix exploits SemLink and applies advanced WSD methods to extend its content. Each row of this Predicate Matrix represents the mapping of a role over the different resources and includes all the aligned knowledge about its corresponding verb sense. The current version of the Predicate Matrix also includes ontological knowledge from the Multilingual Central Repository (MCR) (Atserias et al., 2004; Gonzalez-Agirre et al., 2012). With the Predicate Matrix, we expect to provide a more robust interoperable verbal lexicon. We also plan to discover and solve inherent inconsistencies among the integrated resources. Moreover, we plan to extend the coverage of current predicate resources (by including from WordNet morphologically related nominal and verbal concepts, by exploiting also FrameNet information, etc.). We also plan to enrich WordNet with predicate information, and possibly to extend predicate information to languages other than English (by exploiting the local wordnets aligned to the English WordNet) and predicate information from other languages. For instance, AncoraNet (Taulé et al., 2008).

8. Acknowledgment
We are grateful to the anonymous reviewers for their insightful comments. This work has been partially funded by SKaTer (TIN2012-38584-C06-02), OpeNER (FP7-ICT-2011-SME-DCL-296451) and NewsReader (FP7-ICT-2011-SME-DCL-296451) and the READERS project with the multilingual central repository. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data, pages 1247–1250. ACM.


9. References


