New Features for FrameNet – WordNet Mapping

Sara Tonelli and Daniele Pighin
FBK-Irst, Human Language Technologies
Via di Sommarive, 18 I-38100 Povo (TN) Italy
{satonelli,pighin}@fbk.eu

Abstract

Many applications in the context of natural language processing or information retrieval may be largely improved if they were able to fully exploit the rich semantic information annotated in high-quality, publicly available resources such as the FrameNet and the WordNet databases. Nevertheless, the practical use of similar resources is often biased by the limited coverage of semantic phenomena that they provide.

A natural solution to this problem would be to automatically establish anchors between these resources that would allow us 1) to jointly use the encoded information, thus possibly overcoming limitations of the individual corpora, and 2) to extend each resource coverage by exploiting the information encoded in the others.

In this paper, we present a supervised learning framework for the mapping of FrameNet lexical units onto WordNet synsets based on a reduced set of novel and semantically rich features. The automatically learnt mapping, which we call *MapNet*, can be used 1) to extend frame sets in the English FrameNet, 2) to populate frame sets in the Italian FrameNet via MultiWordNet and 3) to add frame labels to the MultiSemCor corpus. Our evaluation on these tasks shows that the proposed approach is viable and can result in accurate automatic annotations.

1 Introduction

In recent years, the integration of manually-built lexical resources into NLP systems has received growing interest. In particular, resources annotated with the surface realization of semantic roles, like FrameNet (Baker et al., 1998) or PropBank (Palmer et al., 2005) have shown to convey an improvement in several NLP tasks, from question answering (Shen and Lapata, 2007) to textual entailment (Burchardt et al., 2007) and shallow semantic parsing (Giuglea and Moschitti, 2006). Nonetheless, the main limitation of such resources is their poor coverage, particularly as regards FrameNet. Indeed, the latest FrameNet release (v. 1.3) contains 10,195 lexical units (LUs), 3,380 of which are described only by a lexicographic definition without any example sentence. In order to cope with this lack of data, it would be useful to map frame information onto other lexical resources with a broader coverage. We believe that WordNet (Fellbaum, 1998), with 210,000 entries in version 3.0, can represent a suitable resource for this task. In fact, both FrameNet and WordNet group together semantically similar words, and provide a hierarchical representation of the lexical knowledge (in WordNet the relations between synsets, in FrameNet between frames, see Ruppenhofer et al. (2006)). On the other hand, WordNet provides a more extensive coverage particularly for adjectives and nouns denoting artifacts and natural kinds, that are mostly neglected in FrameNet.

In this paper, we present an approach using Support Vector Machines (SVM) to map FrameNet lexical units to WordNet synsets. The proposed approach addresses some of the limitations of previous works on the same task (see for example De Cao et al. (2008) and Johansson and Nugues (2007)). Most notably, as we do not train the SVM on a per-

frame basis, our model is able to cope also with those frames that have little or no annotated sentences to support the frame description. After learning a very fast model on a small set of annotated lexical unit-synset pairs, we can automatically establish new mappings in never-seen-before pairs and use them for our applications. We will evaluate the effect of the induced mappings on two tasks: the automatic enrichment of lexical unit sets in the English and Italian FrameNet via MultiWordNet (Pianta et al., 2002), and the annotation of the MultiSemCor corpus (Bentivogli and Pianta, 2005) with frame labels.

The discussion is structured as follows: in Section 2 we review the main characteristics of FrameNet and WordNet; in Section 3 we discuss previous attempts to establish a mapping between them; in Section 4 we describe our supervised approach to map lexical units onto synsets; Section 5 details the dataset that we employed for our experiments; Section 6 describes the novel features that we used to characterize the mapping; Section 7 details the results of our experiments; in Section 8 we apply the mapping to three resource annotation tasks; finally, in Section 9 we draw our conclusions.

2 FrameNet and WordNet

The FrameNet database (Baker et al. (1998), Fillmore et al. (2003)) is an English lexical resource based on the description of some prototypical situations, the *frames*, and the frame-evoking words or expressions associated to them, the *lexical units* (LU). Every frame corresponds to a scenario involving a set of participants, the *frame elements* (FEs), that are typically the semantic arguments shared by all LUs in a frame.

We report in Table 1 the information recorded in FrameNet for the CAUSE_TO_WAKE frame. In the first row there is the frame definition with the relevant frame elements, namely AGENT, CAUSE, SLEEPER and SLEEP_STATE. Then there is the list of all lexical units evoking the frame and the corresponding part of speech. Note that, differently from WordNet synsets, a frame can contain LUs with different PoS as well as antonymous words. In the last row, an example for each frame element is reported. The lexical unit is underlined, while the constituent bearing the FE label is written in italics. The FrameNet resource is corpus-based, i.e. every lexical unit should be instantiated by at least one example sentence. Besides, every lexical unit comes with a manual lexicographic definition. The latest database release contains 795 frame definitions and 10,195 lexical units, instantiated through approximately 140.000 example sentences. Despite this, the database shows coverage problems when exploited for NLP tasks, and is still being extended by the Berkeley group at ICSI.

WordNet (Fellbaum, 1998) is a lexical resource for English based on psycholinguistics principles and developed at Princeton University. It has been conceived as a computational resource aimed at improving some drawbacks of traditional dictionaries such as the circularity of definitions and the ambiguity of sense references. At present, it covers the majority of nouns, verbs, adjectives and adverbs in the English language, organized in synonym sets called *synsets*, which correspond to concepts. WordNet also includes a rich set of semantic relations across concepts, such as hyponymy, entailment, antonymy, similar-to, etc. Each synset is encoded as a set of synomyms having the same part of speech and described by a definition or *gloss*. In some cases, one or more example sentences may also be reported. The Princeton English WordNet has also been augmented with domain labels (Magnini and Cavaglià, 2000) that group synsets into homogeneous clusters in order to reduce polysemy in the database.

We believe that mapping FrameNet LUs to WordNet synsets would have at least three different advantages: 1) for the English FrameNet, it would automatically increase the number of LUs for frame by

Frame: CAUSE_TO_WAKE		
Def.	An AGENT or CAUSE causes a SLEEPER to transition from the SLEEP_STATE to wakeful consciousness.	
LUs	awaken.v, get_up.v, rouse.v, wake.v, wake_up.v singe.v, sizzle.v, stew.v	
FEs	AGENT	*We* tried to <u>rouse</u> Peter.
	CAUSE	*The rain* <u>woke</u> the children.
	SLEEPER	*Neighbors* were <u>awakened</u> by screams.
	SL_STATE	He <u>woke</u> Constance *from her doze*.

Table 1: Frame CAUSE_TO_WAKE

importing all synonyms from the mapped synset(s), and would allow to exploit the semantic and lexical relations in WordNet to enrich the information encoded in FrameNet. This would help coping with coverage problems and disambiguating the LU senses. 2) For WordNet, it would be possible to add a semantic layer between the synset level and the domain level represented by frame relations, and to enrich the synsets with a computational description of the situation they refer to together with the semantic roles involved. 3) Since frames are mostly defined at conceptual level, the FrameNet model is particularly suitable for cross-lingual induction (Boas, 2005). In this framework, the FrameNet-WordNet mapping could help modelling frame-based resources for new languages using minimal supervision. In fact, the availability of multilingual resources like *MultiWordNet* (Pianta et al., 2002) and *EuroWordNet* (Vossen, 1998) allows to easily populate frame sets for new languages with reduced human effort and near-manual quality by importing all lemmas from the mapped synsets.

3   Related work

Several experiments have been carried out to develop a FrameNet-WordNet mapping and test its applications. Shi and Mihalcea (2005) described a semi-automatic approach to exploit VerbNet as a bridge between FrameNet and WordNet for verbs, using synonym and hyponym relations and similarity between Levin's verb classes and FrameNet frames. Their mapping was used to develop a rule-based semantic parser (Shi and Mihalcea, 2004) as well as to detect target words and assign frames for verbs in an open text (Honnibal and Hawker, 2005).

Burchardt et al. (2005) presented a rule-based system for the assignment of FrameNet frames by way of a "detour via WordNet". They applied a WordNet-based WSD system to annotate lexical units in unseen texts with their contextually determined WordNet synsets and then exploited synonyms and hypernyms information to assign the best frame to the lexical units. The system was integrated into the SALSA RTE system for textual entailment (Burchardt et al., 2007) to cope with sparse-data problems in the automatic assignment of frame labels.

Johansson and Nugues (2007) created a feature representation for every WordNet lemma and used it to train an SVM classifier for each frame that tells whether a lemma belongs to the frame or not. The best-performing feature representation was built using the sequence of unique identifiers for each synset in its hypernym tree and weigthing the synsets according to their relative frequency in the SemCor corpus. They used the mapping in the Semeval-2007 task on frame-semantic structure extraction (Baker et al., 2007) in order to find target words in open text and assign frames.

Crespo and Buitelaar (2008) carried out an automatic mapping of medical-oriented frames to WordNet synsets applying a *Statistical Hypothesis Testing* to select synsets attached to a lexical unit that were statistically significant using a given reference corpus. The mapping obtained was used to expand Spanish FrameNet using *EuroWordNet* (Vossen, 1998) and evaluation was carried out on the Spanish lexical units obtained after mapping.

Given a set of lexical units, De Cao et al. (2008) propose a method to detect the set of suitable WordNet senses able to evoke a frame by applying a similarity function that exploits different WordNet information, namely conceptual density for nouns, synonymy and co-hyponymy for verbs and synonymy for adjectives. The mapping approach was applied also to LU induction for the English FrameNet and for Italian frames via MultiWordNet.

4   Problem formulation

Our objective is to be able to assign to every lexical unit l, belonging to a frame F_i defined in the FrameNet database, one or more WordNet senses that best express the meaning of l. More specifically, for every $l \in F_i$, we consider the set of all WordNet senses where l appears, *CandSet*, and then find the best WordNet sense(s) $best_s \subset CandSet$ that express the meaning of l.

For example, the lexical unit *rouse.v* belonging to the CAUSE_TO_WAKE frame, is defined in FrameNet as "bring out of sleep; awaken". Its *CandSet* comprises 4 senses[1]: 1# *bestir, rouse* (become active); 2# *rout_out, drive_out, force_out, rouse* (force or drive_out); #3 *agitate, rouse, turn_on, charge, com-*

---

[1]The gloss is reported between parenthesis

move, excite, charge_up (cause to be agitated, excited or roused); #4 *awaken, wake, waken, rouse, wake_up, arouse* (cause to become awake or conscious). In this example, $best_s = \{\#4\}$ for *rouse.v* in CAUSE_TO_WAKE.

We aim at creating a mapping system that can achieve a good accuracy also with poorly-documented lexical units and frames. In fact, we believe that under real-usage conditions, the automatic induction of LUs is typically required for frames with a smaller LU set, especially for those with only one element. In the FrameNet database (v. 1.3), 33 frames out of 720 are described only by one lexical unit, and 63 are described by two. Furthermore, more than 3,000 lexical units are characterized only by the lexicographic definition and are not provided with example sentences. For this reason, we suggest an approach that makes also use of usually unexploited information in the FrameNet database, namely the *definition* associated to every lexical unit, and disregards example sentences.

This is the main point of difference between our and some previous works, e.g. Johansson and Nugues (2007) and De Cao et al. (2008), where unsupervised approaches are proposed which strongly rely either on the number of lexical units in a frame or on the example sentences available for *l* in the FrameNet corpus. We claim that the relative short time necessary to annotate a small dataset of frame-synset pairs will result in a more reliable mapping system and, as a consequence, in consistent time savings when we actually try to use the mappings for some tasks. The ability to cope with different cases while retaining a good accuracy will allow to bootstrap the mapping process in many cases where other approaches would have failed due to lack of training data.

To this end, we can train a binary classifier that, given *l* and *CandSet*, for each pair $\langle l, s \rangle$, $s \in CandSet$, delivers a positive answer if $s \in best_s$, and a negative one otherwise. To follow on the previous example, for *rouse.v* we would have 4 classifier examples, i.e. the pairs $\langle$rouse.v,#1$\rangle$, $\langle$rouse.v,#2$\rangle$, $\langle$rouse.v,#3$\rangle$ and $\langle$rouse.v,#4$\rangle$. Of these, only the last would be considered a positive instance. As a learning framework, we decided to use SVMs due to their classification accuracy and robustness to noisy data (Vapnik, 1998).

5 Dataset description

In order to train and test the classifier, we created a gold standard by manually annotating 2,158 LU-synset pairs as positive or negative examples. We don't have data about inter-annotator agreement because the dataset was developed only by one annotator, but De Cao et al. (2008) report 0.90 as Cohen's Kappa computed over 192 LU-synset pairs for the same mapping task. This confirms that senses and lexical units are highly correlated and that the mapping is semantically motivated.

The annotation process can be carried out in reasonable time. It took approximately two work days to an expert annotator to manually annotate the 2,158 pairs that make up our gold standard. The lexical units were randomly selected from the FrameNet database regardless of their part of speech or amount of annotated data in the FrameNet database. For each lexical unit, we extracted from WordNet the synsets where the LU appears, and for each of them we assigned a positive label in case the LU-synset pairs share the same meaning, and a negative label otherwise. Statistics about the dataset are reported in Table 2.

N. of LU-synset pairs	2,158
N. of lexical units	617
Verbal lexical units	39%
Nominal lexical units	51%
Adjectival lexical units	9%
Adverbial lexical units	<1%
Targeted frames	386
Pairs annotated as positive	32%
Pairs annotated as negative	68%
Average polysemy	3.49
LUs with one candidate synset	204
LUs with 10 or more cand. synsets	32

Table 2: Statistics on the dataset

The 386 frames that are present in the dataset represent about one half of all lexicalized frames in the FrameNet database. This proves that, despite the limited size of the dataset, it is well representative of FrameNet characteristics. This is confirmed by the distribution of the part of speech. In fact, in the FrameNet database about 41% of the LUs

are nouns, 40% are verbs, 17% are adjectives and <1% are adverbs (the rest are prepositions, which are not included in our experiment because they are not present in WordNet). In our dataset, the percentage of nouns is higher, but the PoS ranking by frequency is the same, with nouns being the most frequent PoS and adverbs the less represented. The average polysemy corresponds to the average number of candidate synsets for every LU in the dataset. Note that the high number of lexical units with only one candidate does not imply a more straightforward mapping, because in some cases the only candidate represents a negative example. In fact, a LU could be encoded in a frame that does not correspond to the sense expressed by the synset.

6 Feature description

For every LU-synset pair in the gold standard, we extracted a set of features that characterize different aspects of the mapping. In the remainder, we detail the meaning as well as the feature extraction procedure of each of them.

Stem overlap Both WordNet glosses and LU definitions in FrameNet are manually written by lexicographers. We noticed that when they share the same sense, they show high similarity, and sometimes are even identical. For example, the definition of *thicken* in the *Change_of_consistency* frame is *"become thick or thicker"*, which is identical to the WordNet gloss of synset n. v#00300319. The *thicken* lemma occurs in three WordNet synsets, and in each of them it is the only lemma available, so no other information could be exploited for the sense disambiguation.

We believe that this information could help in the choice of the best candidate synset, so we stemmed all the words in the synset gloss and in the lexical unit definition and measured their overlap. As features, we use the ratio between the number of overlapping words and the number of words in the definition, both for the gloss and the LU description.

Prevalent Domain and Synset Since a frame represents a prototypical situation evoked by the set of its lexical units, our intuition is that it should be possible to assign it to a WordNet domain, that groups homogeneous clusters of semantically similar synsets (see Section 2).

Given the LU-synset pair $\langle l, s \rangle$, $l \in F_i$, $s \in CandSet$, we extract all the lexical units in F_i and then build a set *AllCandSet* of pairs $\langle s_j, c_j \rangle$, where s_j is a synset in which at least one $l_i \in F_i$ appears, and c_j is the count of lexical units that are found in s_j.

We exploit the information conveyed by *AllCandSet* in two ways: i) if there is a prevalent WordNet domain that characterizes the majority of the synsets in *AllCandSet*, and $s \in CandSet$ belongs to that same domain, we add a boolean feature to the feature vector representing $\langle l, s \rangle$; ii) if s is the synset with the highest count in *AllCandSet*, i.e. if $s = s_j$ and $c_j > c_i \forall \langle s_j, c_j \rangle \in AllCandSet, i \neq j$, then we add another boolean feature to encode this information.

Cross-lingual parallelism Our idea is that, if an English lexical unit and its Italian translation belong to the same frame, they are likely to appear also in the same MultiWordNet synset, and the latter would be a good candidate for mapping. In fact, in Multi-WordNet the Italian WordNet is strictly aligned with the Princeton WordNet 1.6, with synsets having the same id for both languages, and also semantic relations are preserved in the multilingual hierarchy. Since no Italian FrameNet is available yet, we extended the parallel English-Italian corpus annotated on both sides with frame information described in Tonelli and Pianta (2008) by adding and annotating 400 new parallel sentences. The final corpus contains about 1,000 pairs of parallel sentences where the English and the Italian lexical unit belong to the same frame.

Given a pair $\langle l, s \rangle$, we check if l appears also in the corpus with the frame label F_i and extract its Italian translation l_{it}. If l_{it} appears also in the Italian version of synset s in MultiWordNet, we consider s as a good candidate for the mapping of l and encode this information as a binary feature.

Simple synset-frame overlap Intuitively, the more lemmas a frame and a synset have in common, the more semantically similar they are. In order to take into account this similarity in our feature vector, given the pair $\langle l, s \rangle, l \in F_i$, we extract all lexical units in F_i and all lemmas in s and we compute the number of overlapping elements. Then we divide

the value by the number of synsets where the same overlapping element(s) occur.

As an example, the words *tank* and *tank car* in the *Vehicle* frame, occur together only in the fourth synset related to *tank*, which therefore will have a higher value for this feature.

Extended synset-frame overlap   This feature is a generalization of overlapping value described above. In fact, we noticed that the hypernym information in WordNet can help disambiguating the synsets. Therefore, we take into account not only the overlaps according to the previous criterion, but also the number of overlapping words between the lexical units in a frame and the hypernyms of a synset. For example, the *party.n* lexical unit in the AGGREGATE frame has 5 senses in WordNet. According to the previous criterion, there is no overlap between the LUs in the frame and the lemmas in any of the five synsets. Instead, if we look at the direct hypernym relation of *party*, we find that sense #3 is also described as *set, circle, band*, that are also lexical units of AGGREGATE.

In those cases where the hypernym relation is not defined, e.g. adjectives, we used the *similar-to* relation.

7   Experimental setup and evaluation

To evaluate our methodology we carried out a 10-fold cross validation using the available data, splitting them in 10 non-overlapping sets. For each iteration, 70% of the data was used for training, 30% for testing. All the splits were generated so as to maintain a balance between positive and negative examples in the training and test sets.

We used the SVM optimizer SVM-Light[2] (Joachims, 1999), and applied polynomial kernels (*poly*) of different degrees (i.e. 1 through 4) in order to select the configuration with the best generalization capabilities. The accuracy is measured in terms of Precision, Recall and F_1 measure, i.e. the harmonic average between Precision and Recall. For the sake of annotation, it is important that an automatic system be very precise, thus not producing wrong annotations. On the other hand, the higher the recall, the larger the amount of data that the system will be able to annotate.

---

[2]Available at http://svmlight.joachims.org/

The macro-average of the classifier accuracy for the different configurations is shown in Table 3. We report results for linear kernel (i.e. poly 1), maximizing recall and f-measure, and for polynomial kernel of degree 2 (i.e. poly 2), scoring the highest precision. In general , we notice that all our models have a higher precision than recall, but overall are quite balanced. Different polynomial kernels (i.e. conjunction of features) do not produce very relevant differences in the results, suggesting that the features that we employed encode significant information and have a relevance if considered independently.

As a comparison, we also carried out the same evaluation by setting a manual threshold and considering a LU-synset pair as a positive example if the sum of the feature values was above the threshold. We chose two different threshold values, the first (Row 1 in Table 3) selected so as to have comparable precision with the most precise SVM model (i.e. poly2), the second (Row 2) selected to have recall comparable with poly1, i.e. the SVM model with highest recall. In the former case, the model has a recall that is less than half than poly2, i.e. 0.214 vs. 0.569, meaning that such model would establish a half of the mappings while making the same percentage of mistakes. In the latter, the precision of the SVM classifier is 0.114 points higher, i.e. 0.794 vs. 0.680, meaning the SVM can retrieve as many mappings but making 15% less errors.

In order to investigate the impact of different features on the classifier performance, we also considered three different groups of features separately: the ones based on stem overlap, those computed for prevalent domain and synset, and the features for simple and extended frame – synset overlap. We did not take into account cross-lingual parallelism because it is one single feature whose coverage strongly relies on the parallel corpus available. As a consequence, it is not possible to test the feature in isolation due to data sparseness.

Results are shown in Table 3, in the second group of rows. Also in this case, we carried out a 10-fold cross validation using a polynomial kernel of degree 2. The stem overlap features, which to our best knowledge are an original contribution of our approach, score the highest recall among the three groups. This confirms our intuition that LU defini-

tions and WordNet glosses can help extending the number of mapped LUs, including those that are poorly annotated. For instance, if we consider the KNOT_CREATION frame, having only *tie.v* as LU, the features about prevalent domain & synset and about synset-frame overlap would hardly be informative, while stem overlap generally achieves a consistent performance regardless of the LU set. In fact, *tie.v* is correctly mapped to synset v#00095054 based on their similar definition (respectively *"to form a knot"* and *"form a knot or bow in"*). Best precision was scored by the feature group considering prevalent domain & synset, which are also new features introduced by our approach. The positive effect of combining all features is clearly shown by comparing the results obtained with individual feature groups against the figures in the row labeled *poly2*.

	Prec.	Recall	F1
Man. thresh. (P)	0.789	0.214	0.337
Man. thresh. (F1)	0.680	0.662	0.671
Stem Overlap	0.679	0.487	0.567
Prev.Dom.& Syn.	0.756	0.434	0.551
Syn.- Frame Overlap	0.717	0.388	0.504
poly1	0.761	**0.613**	**0.679**
poly2	**0.794**	0.569	0.663

Table 3: Mapping evaluation

8 MapNet and its applications

Since we aim at assigning at least one synset to every lexical unit in FrameNet, we considered all the frames and for every LU in the database we created a list of LU-synset pairs. We re-trained the classifier using the whole annotated gold standard and classified all the candidate pairs. The mapping produced between the two resources, that we call *MapNet*, comprises 5,162 pairs. Statistics on MapNet are reported in table 4.

About one thousand lexical units in FrameNet have no candidate synsets because the lemma is not present in WordNet. The remaining LUs have 3.69 candidate synsets each on average, similarly to the average polysemy reported for the gold standard (see Table 2). This confirms our hypothesis that the data used for training are well representative of the char-

N. of LUs in FrameNet	10,100
N. of LUs with at least one syn.cand.	9,120
N. of LU-synset candidate pairs	33,698
N. of mapped pairs	5,162

Table 4: Statistics on the mapping

acteristics of the whole resource. We expect about 80% of these mappings to be correct, i.e. in line with the precision of the classifier.

8.1 Automatic FrameNet extension

MapNet can be easily exploited to automatically extend FrameNet coverage, in particular to extend the set of lexical units for each frame. In fact, we can assume that all lemmas in the mapped synsets have the same meaning of the LUs in the corresponding frames. We use MapNet to extract from WordNet the lemmas in the mapped synsets and add them to the frames.

For English FrameNet, we can acquire 4,265 new lexical units for 521 frames. In this way, we would extend FrameNet size by almost 42%. In the random evaluation of 100 newly acquired LUs belonging to 100 different frames, we assessed a precision of 78%. For the Italian side, we extract 6,429 lexical units for 561 frames. Since no Italian FrameNet has been developed yet, this would represent a first attempt to create this resource by automatically populating the frames. We evaluate the content of 15 complete frames containing 191 Italian LUs. The assigned LUs are correct in 88% of the considered cases, which represent a promising result w.r.t. the unsupervised creation of Italian FrameNet.

The difference in the evaluation for the two languages most likely lies in the smaller number of synsets on the Italian side of MultiWordNet if compared to the English, which results in less ambiguity. Furthermore, we should consider that the task for Italian is easier than for English, since in the former case we are building a resource from scratch, while in the latter we are extending an already existing resource with lexical units which are most likely peripheral with respect to those already present in the database.

8.2 Frame annotation of MultiSemCor

MultiSemCor (Bentivogli and Pianta, 2005) is an English/Italian parallel corpus, aligned at word level and annotated with PoS, lemma and Word-Net synsets. The parallel corpus was created starting from the SemCor corpus, which is a subset of the English Brown corpus containing about 700,000 running words. The corpus was first manually translated into Italian. Then, the procedure of transferring word sense annotations from English to Italian was carried out automatically.

We apply *MapNet* to enrich the corpus with frame information. We believe that this procedure would be interesting from different point of views. Not only we would enrich the resource with a new annotation layer, but we would also automatically acquire a large set of English and Italian sentences having a lexical unit with a frame label. For the English side, it is a good solution to automatically extract a dataset with frame information and train, for example, a machine learning system for frame identification. For the Italian side, it represents a good starting point for the creation of a large annotated corpus with frame information, the base for a future Italian FrameNet.

MultiSemCor contains 12,843 parallel sentences. If we apply *MapNet* to the corpus, we produce 27,793 annotated instances in English and 23,872 in Italian, i.e. about two lexical units per sentence. The different amount of annotated sentences depends on the fact that in MultiSemCor some synset annotations have not been transferred from English to Italian. From both sides of the resulting corpus, we randomly selected 200 sentences labeled with 200 different frames, and evaluated the annotation quality. As for the English corpus, 75% of the sentences was annotated with the correct frame label, while on the Italian side they were 70%. This result is in line with the expectations, since *MapNet* was developed with 0.79 precision. Besides, synset annotation on the English side of MultiSem-Cor was carried out by hand, while annotation in Italian was automatically acquired by transferring the information from the English corpus (precision 0.86). This explains why the resulting annotation for English is slightly better than for Italian. In some cases, the wrongly annotated frame was strictly connected to the right one, i.e. APPLY_HEAT instead of COOKING_CREATION and ATTACHING instead of INCHOATIVE_ATTACHING.

9 Conclusions

We proposed a new method to map FrameNet LUs to WordNet synsets using SVM with minimal supervision effort.

To our best knowledge, this is the only approach to the task that exploits features based on stem overlap between LU definition and synset gloss and that makes use of information about WordNet domains. Differently from other models, the SVM is not trained on a per-frame basis and we do not rely on the number of the annotated sentences for a LU in the FrameNet corpus, thus our mapping algorithm performs well also with poorly-annotated LUs. After creating *MapNet*, the mapping between FrameNet and WordNet, we applied it to three tasks: the automatic induction of new LUs for English FrameNet, the population of frames for Italian FrameNet and the annotation of the MultiSemCor corpus with frame information. A preliminary evaluation shows that the mapping can significantly reduce the manual effort for the development and the extension of FrameNet-like resources, both in the phase of corpus annotation and of frame population.

In the future, we plan to improve the algorithm by introducing syntactic features for assessing similarity between LU definitions and WordNet glosses. We also want to merge all information extracted and collected for Italian FrameNet and deliver a seed version of the resource to be validated. Finally, we plan to extend the mapping to all languages included in MultiWordNet, i.e. Spanish, Portuguese, Hebrew and Romanian.

References

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet Project. In *Proceedings of the 36th ACL Meeting and 17th ICCL Conference*. Morgan Kaufmann.

Collin F. Baker, Michael Ellsworth, and Katrin Erk. 2007. SemEval-2007 Task 10: Frame Semantic Structure Extraction. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 99–104, Prague, CZ, June.

Luisa Bentivogli and Emanuele Pianta. 2005. Exploiting Parallel Texts in the Creation of Multilingual Semantically Annotated Resources: The MultiSemCor Corpus. *Natural Language Engineering, Special Issue on Parallel Texts*, 11(03):247–261, September.

Hans C. Boas. 2005. Semantic frames as interlingual representations for multilingual lexical databases. *International Journal of Lexicography*, 18(4):445–478.

Aljoscha Burchardt, Katrin Erk, and Annette Frank. 2005. A WordNet detour to FrameNet. In B. Fisseni, H. Schmitz, B. Schröder, and P. Wagner, editors, *Sprachtechnologie, mobile Kommunikation und linguistische Resourcen*, Frankfurt am Main, Germany. Peter Lang.

Aljoscha Burchardt, Nils Reiter, Stefan Thater, and Annette Frank. 2007. A semantic approach to textual entailment: System evaluation and task analysis. In *Proceedings of Pascal RTE-3 Challenge*, Prague, CZ.

Diego De Cao, Danilo Croce, Marco Pennacchiotti, and Roberto Basili. 2008. Combining Word Sense and Usage for modeling Frame Semantics. In *Proceedings of STEP 2008*, Venice, Italy.

Mario Crespo and Paul Buitelaar. 2008. Domain-specific English-to-Spanish Translation of FrameNet. In *Proc. of LREC 2008*, Marrakech.

Christiane Fellbaum, editor. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

C.J. Fillmore, C.R. Johnson, and M. R. L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16:235–250, September.

Ana-Maria Giuglea and Alessandro Moschitti. 2006. Semantic role labeling via FrameNet, VerbNet and PropBank. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual ACL meeting*, pages 929–936, Morristown, NJ, US. Association for Computational Linguistics.

Matthew Honnibal and Tobias Hawker. 2005. Identifying FrameNet frames for verbs from a real-text corpus. In *Proceedings of Australasian Language Technology Workshop 2005*.

Thorsten Joachims. 1999. Making large-scale support vector machine learning practical. In Bernhard Schölkopf, Christopher J. C. Burges, and Alexander J Smola, editors, *Advances in kernel methods: support vector learning*, pages 169–184. MIT Press, Cambridge, MA, USA.

R. Johansson and P. Nugues. 2007. Using WordNet to extend FrameNet coverage. In *Proc. of the Workshop on Building Frame-semantic Resources for Scandinavian and Baltic Languages, at NODALIDA*, Tartu.

Bernardo Magnini and Gabriela Cavaglià. 2000. Integrating Subject Field Codes into WordNet. In *Proceedings of LREC 2000*, pages 1413–1418, Athens, Greece.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31.

Emanuele Pianta, Luisa Bentivogli, and Christian Girardi. 2002. MultiWordNet: developing an aligned multilingual database. In *First International Conference on Global WordNet*, pages 292–302, Mysore, India.

Josef Ruppenhofer, Michael Ellsworth, Miriam R.L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. 2006. *FrameNet II: Extended Theory and Practice*. Available at http://framenet.icsi.berkeley.edu/book/book.html.

Dan Shen and Mirella Lapata. 2007. Using Semantic Roles to Improve Question Answering. In *Proceedings of EMNLP and CONLL*, pages 12–21, Prague, CZ.

Lei Shi and Rada Mihalcea. 2004. Open Text Semantic Parsing Using FrameNet and WordNet. In *Proceedings of HLT-NAACL 2004*.

Lei Shi and Rada Mihalcea. 2005. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. In *Proceedings of CICLing 2005*, pages 100–111. Springer.

Sara Tonelli and Emanuele Pianta. 2008. Frame Information Transfer from English to Italian. In European Language Resources Association (ELRA), editor, *Proceedings of LREC 2008*, Marrakech, Morocco.

Vladimir N. Vapnik. 1998. *Statistical Learning Theory*. Wiley-Interscience.

Piek Vossen, editor. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Springer, October.