The Impact of Grammar Enhancement on Semantic Resources Induction

Luca Dini, Giampaolo Mazzini
CELI
Via San Quintino, 31- 10121-Torino
dini@celi.it, mazzini@celi.it

Abstract

In this paper describes the effects of the evolution of an Italian dependency grammar on a task of multilingual FrameNet acquisition. The task is based on the creation of virtual English/Italian parallel annotation corpora, which are then aligned at dependency level by using two manually encoded grammar based dependency parsers. We show how the evolution of the LAS (Labeled Attachment Score) metric for the considered grammar has a direct impact on the quality of the induced FrameNet, thus proving that the evolution of the quality of syntactic resources is mirrored by an analogous evolution in semantic ones. In particular we show that an improvement of 30% in LAS causes an improvement of precision for the induced resource ranging from 5% to 10%, depending on the type of evaluation.

1. Introduction

Since at least two decades a number of methods have been experimented to acquire lexical resources. In many cases they are used to derive manually coded semantic information from languages with a rich repository of resources (typically English) in favour of languages for which no effort of manual coding was envisaged. The most populated examples in this direction are probably WordNet and FrameNet. A number of such approaches relies on the capability of parsing sentences in one or more languages. While there are relatively well known standards on how to evaluate the results of these experiments, it is extremely rare to find information *about the impact of the quality of the adopted grammars on those same experiments*. In this paper we will describe the effects of the evolution of an Italian dependency grammar on a task of multilingual FrameNet acquisition. We will show how the evolution of the LAS metric for the considered grammar has a direct impact on the quality of the induced FrameNet, thus proving how the evolution of the quality of syntactic resources is mirrored by an analogous evolution in semantic ones.

2. The Approach

Our approach to Italian FrameNet Induction is discussed in details in Dini & Bosca (2009) and briefly described here for the sake of completeness. Its crucial assumption is that frames and frame elements are constant across languages (Lönneker-Rodman, 2007). While that assumption might be questionable for some elements describing very specific, domain-dependent situations, here it is assumed correct for the general situation of a language, and it has been validated in projects such as SALSA (Burchardt et al., 2006).

A second crucial assumption is that machine translation systems are currently able to produce correct translations, if not of entire sentences, at least of parts of them. On the basis of this assumption, it should be the case that FrameNet English examples (annotations) translated into a target language retain *some* syntactically correct frame element realizations.

From a technical point of view our implementation of these assumptions will require the following components:

- A machine Translation system from Source to Target (we used SYSTRAN);
- A dictionary based translation system from Source to Target (we adopted our bilingual dictionaries:310.000 entries for EN<->IT);
- Two dependency grammars, for Source and Target (we made use of Xerox XIP Italian and English grammars (Mokhtar et al., 2001), in the context of the CACAO project);
- A component of word sense disambiguation (see below);
- A component of word sense clustering (see below).

The proposed methodology is based on the following steps:

1. Identify the most probable Italian candidate for a given English lexical unit;
2. Translate all available English examples in Italian;
3. Parse the English examples in order to identify the semantic head of any realized frame element;
4. Parse the Italian translation and identify the head associated to the frame in the Italian example as well as the head of frame elements.
5. Store the relation between the head and the dependant identified in the Italian sentence as valence information.

As it can be noticed, Our work is closely related with the ones of Padò and Pitel (2007), Padò(2007), Tonelli and Pianta (2008), as they all try to automatically produce frames and frame elements information by producing automatic annotations of parallel corpora. In both cases the basic ingredients are the same, i.e. parsed structure for both source and target corpora and a word alignment system that might exploit a bilingual translation component. However, it is worth to point out some differences:

- The goals of those works is to obtain FrameNet annotated corpora in the target languages. Our goal is rather to obtain "abstract" assignments of frames and frame elements (with their syntactic valence) to Italian words, without producing an annotated corpus.
- Our approach does not need parallel bilingual corpora. On the contrary we exploit a reliable Machine Translation system, as well as a bilingual word-to-word translation system.
- Our approach is entirely based on dependency grammars, rather than phrasal grammars.

More similarities can be found with works on Chinese BiFrameNet (Fung and Chen, 2004), which map FrameNet entries to the HowNet resources by using bilingual translation dictionaries. Our methodology differs from such approach in that we do not use an already available semantic net for linking FrameNet entries, but we rather induce a brand new FrameNet. Moreover our specific focus, is, again, the induction of Italian valence elements to be used in conjunction with dependency grammars. Nevertheless, we recognize that the BiFrameNet example gathering technique deserve further attention for the follow up of our work..
Finally, Crespo and Buitelaar (2008) propose a method for performing frame assignment to Spanish words in a medical domain by using domain corpora and mappings via EuroWordNet. Their techniques seem to be quite accurate in detecting frame based correspondences among different words , thus disambiguating among different word senses and providing correct frame assignment. With respect to the work described in this paper, however, their investigation does not address neither the issue of valence induction nor the one of example annotation.

2.1 Translating the lexical unit

Translation of lexical units associated to a frame is achieved via bilingual lexicon access. It is well known that this kind of technique generates many translation alternatives (Curtoni and Dini, 2006) which reflect *at least* the semantic ambiguity of the input word plus, in many cases, the set of synonyms of the target word.
As our methodology is based on translation of examples, we apply a double heuristic in order to perform disambiguation.

- **Intersection**: We prefer the translation of the lexical unit that is found in the biggest number of translated examples.
- **WSD**: We perform disambiguation of translation equivalences by applying the same technique described in Curtoni and Dini (2006) and Bosca and Dini (2008). Shortly, we use the corpus based semantic vectors (Widdows & Ferraro, 2008) associated to each translation and evaluate the degree of semantic coherence with the other words in the translate example.

2.2 Parsing

Once we obtain the candidate target lexical unit as well as all relevant translated examples we proceed to parse both source and target examples in order to identify semantic heads of realized frame elements. Source examples are parsed because we need to identify with precision the semantic head of the constituent triggering the frame element assignment. Once this is identified, it is translated into Italian and a semantically matching element is looked for. Again, a *semantically matching element* is an element which is either identical to one of the available translations or which can be related by the semantic clustering procedure we mentioned in the previous section. The following is an example of a simple matching:

in the resolution of a [foreign **policy***[sv2]* $_{\text{Issue}}$] [**dispute***[sv1]* $_{\text{Quarreling}}$]
nella risoluzione di una **disputa***[sv1]* di **politica***[sv2]* straniera.

with relevant semantic dependencies in a simplified COLLN format:

6	Foreign	8	MOD
7	Policy	8	MOD
8	Dispute	3	MOD

17	Disputa	14	ARG
18	Di	17	MOD
19	Politica	18	ARG

If the matching has success, then we are in a position to compute hypotheses about the valence of the target lexical unit. This is an example of acquired valence for the Italian verb *fornire* as a translation of the lexical unit *supply.v* in the frame **Supply**:

Frame Element	# Ann.	Realizations(s)
Circumstances	2	NOUN[a].VMOD
Purpose_of_recipient	1	NOUN[per].NMOD
Purpose_of_recipient	1	VERB[+inf].TOP
Theme	2	NOUN[con].VMOD
Theme	3	NOUN.OBJ
Theme	1	NOUN.SUBJ_PASSIVE
Co_theme	1	NOUN[senza].VMOD
Means	1	NOUN.AGENT
Supplier	6	NOUN.SUBJ
Recipient	13	NOUN[a].VMOD
Recipient	1	NOUN.VMOD

3. Results

The whole process of FrameNet induction lasted about one week on a bi-processor server. The generated FrameNet contains 5960 Italian lexical units associated to 628 different frames. As for valences, out of 42923 Italian examples we collected 38109 valences (163 valence types)

realizing different frame elements.

3.1 Failures

Table 3 contains a comparison of Italian and English FrameNet data. It is evident that in the shift from the English resource to the Italian one a lot of information was lost.

	EN	IT
Instantiated Frames	721	628
Available units	10,195	5,960
Available examples	139,382	42,923
Available FE instances	70,075	8,426

Table 3 : Comparison between the English and the Italian FrameNet.

While logging the learning process, we monitored the following set of fault types:

- No translation for a lexical unit (7,815);
- Absence of examples in the source FrameNet (4,922);
- No translated example contains the candidate translation(s) of the lexical unit (1,736).
- No head could be identified for English frame element realization (parse error or difficult structure, e.g. coordination) (6,191)
- The translation of the semantic head of the frame element or of the frame bearing head could not be matched in the Italian example. (99,808)
- The semantic heads of both the lexical unit and the frame element are found in the Italian example but the parser could not find any dependency among them. (94,004)

These data were obtained with a version of an Italian grammar which was rather "embryonic" in nature, as confirmed by the amount of errors in the last category. The LAS (Labeled Attachment Score) for such a grammar, as computed on the TUT (Lesmo et al. 2002) corpus was about 40%. The goal of the experiment is to evaluate how parsing errors evolved by using the same grammar after 6 month of manual configuration, with an achieved LAS score of about 70% on the same gold standard.

4. Evaluation of the impact LAS on Semantic Acquisition

The count of parsing anomalies as described in the previous section is not necessarily a quality indicator for the extracted semantic resource: it is a symptom that something went wrong, but nothing authorize to claim that in absence of those errors the quality of the induced FrameNet would increase. In order to prove this, we set up two evaluation experiments, described in the two following sections. As the main goal of this paper is to show the impact of parsing quality over semantic acquisition, we will devote the last section to the description of the grammar development process.

3.1 Semantic Role Labeling Evaluation

This evaluation is meant to measure the precision of the induced Italian FrameNet in a typical operational context i.e. semantic role labeling (Màrquez et al., 2008). The ideal corpus for our evaluation must have semantic head-to-head frame element annotations and manually assigned syntactic dependencies. In order to produce such a corpus we started from the Italian TUT corpus (Lesmo et al., 2002), which, in the newspaper section, contains 1000 Italian sentences annotated according to the dependency grammar theory. The corpus was used in EVALITA (Bosco et al., 2007) and as such can be considered a community-acknowledged resource for Italian. The details of manual annotation to produce a compatible SRL gold standard are described in Dini and Bosca (2009).

Once the gold standard is obtained we are in a position to evaluate the results of the SRL task based on the induced FrameNet. SRL is not the central objective of this work, but only a way of evaluating the quality of the induced Italian FrameNet. Therefore we use as input to our semantic labeller the parses obtained from the TUT corpus, in order eliminate the possibility of syntactic mismatch. Also, as the main focus is not word sense disambiguation but valence induction, we assume that lexical units come to the semantic labeller as already disambiguated.

Fully correct parsing and fully disambiguated lexical units might look like an excessive help to the SRL task, but this way we can focus on the precision and recall of semantic valences, thus on the effect of the evolution of the Italian grammar.

In the SRL task we adopted the F-Measure metrics proposed by Toutanova et al. (2008). By using the induced valence and a very elementary algorithm for frame element assignment we obtained a precision of 0.53, a recall of 0.33 and a consequent precision of 0.41 when using the FrameNet induced with the 40% LAS grammar. The results obtained with the 70% LAS grammar based FrameNet show that precision increases significantly (0,59) while recall stays almost the same (0.34).

3.1 Valence Matching Evaluation

In the second kind of evaluation we wanted to measure *directly* the quality of induced valences, i.e. without passing through a SRL task. In order to do this we first had to build a gold standard of valences. For that we adopted a more "standard" Berkley methodology, in the sense that **first** we manually annotated a set of about 20 sentences containing a certain frame bearing lexical unit (actually we annotated only the corresponding frame elements) and **then** we computed the abstract valence of the lexical unit. In this way, from about 400 manually annotated sentences, we obtained a set of 20 lexical unit, together with their "gold" valence to be compared with the one of the induced FrameNet.

The methodology for the comparison was the same as the one adopted in Bosca & Dini (2009) for the so called

"Monolingual Evaluation" and "Transitive Evaluation", with the difference that now we do not compare our results against some manipulations of the original English FrameNet, but against a specific manually annotated Gold Standard. In short we cycle on all induced lexical units, we retrieve valences and we look for a match in the set of original valences. A match is positive if part-of-speech (including prepositional form), grammatical function and frame element name all match. This allows us to compute precision and recall for valences of each lexical unit.
In general we notice that the adoption of the 70% LAS grammar causes an increment of about 10% in recall and 7 % in precision.

3.1 Grammar development

It is worthwhile to mention the fact that for parsing both English and Italian sentences we adopted the Xerox parser XIP (Mokhtar et al., 2001), which is a parsing engine completely based on manually coded grammars. The English grammar underwent no modification, so the increase in accuracy is to be attributed only to the changes of the Italian grammar, described in Testa & al. (2009). In particular the reported increase of LAS was achieved via a grammar coding activity which lasted approximately six months. Such an activity was completely independent from the experiment described in this paper, and it was mainly based on the TUT corpus.

5. Conclusions

In the present work we have shown how the quality of a grammar impacts on a task of semantic induction heavily based on parsing results. In general we noticed that a 30% improvement in LAS causes the induction of a semantic resource which performs about 5% better in a SRL task and about 10% better in a direct evaluation of the quality of the mapping from syntax to semantics (induced valences).

6. Acknowledgements

This work has been carried on in the context of ICT4Law (ICT Converging on Law: Next Generation Services for Citizens, Enterprises, Public Administration and Policymakers) funded by Regione Piemonte. We are also grateful to the members of the IFRAME project (http://sag.art.uniroma2.it/iframe/doku.php) for ideas and discussions.

7. References

Aït-Mokhtar S., Chanod J.-P. et Roux C. 2001. *A multi-input dependency parser*. In Actes d' IWPT'01.

Bosca A. and L. Dini, *CACAO Project at the TEL@CLEF 2008 Task* in CLEF 2008 Working Notes.

Bosco C., A. Mazzei, V. Lombardo. 2007. *Evalita parsing task: an analysis of the first parsing system contest for Italian.* In Intelligenza artificiale, anno IV, num 2 .

Burchardt A., K. Erk, A. Frank, A. Kowalski, S. Padó, M. Pinkal. 2006. *The SALSA corpus: a German corpus resource for lexical semantics*. In Proceedings of the 5th LREC, Genoa, Italy.

Crespo M. and Buitelaar P. 2008. *Domain-Specific English-To-Spanish Translation of FrameNet*. Proc. of LREC 2008, Marrakech, Morocco.

Curtoni P. and Dini L. 2006. Celi participation at CLEF 2006: Cross language delegated search. In CLEF2006 Working notes.

Dini L & Bosca A.., Dependency Based Valence Induction for an Italian FrameNet, Proc. GL09.

Fung. P. and B. Chen. 2004. *BiFrameNet: bilingual frame semantics resource construction by cross-lingual induction*. COLING '04: Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics.Morristown, NJ, USA

Lesmo L., Lombardo V.,Bosco C. 2002. *Treebank Development: the TUT Approach*. In Proceedings of ICON 2002, Mumbay, India.

Lönneker-Rodman B. 2007. *Multilinguality and FrameNet*. ICSI Technical Report TR-07-001, Berkeley, CA, March 2007

Padó S. 2007. *Translational Equivalence and Cross-lingual Parallelism: The Case of FrameNet Frames*. Proc. of the NODALIDA Workshop on Building Frame Semantics Resources for Scandinavian and Baltic Languages, Tartu, Estonia.

Padó S. and G. Pitel. 2007. *Annotation précise du français en sémantique de rôles par projection cross-linguistique*. Proceedings of TALN 2007, Toulouse, France.

Tonelli S. and E. Pianta. 2008. *Frame Information Transfer from English to Italian*. Proceedings of the Sixth International Language Resources and Evaluation (LREC'08). pp 28-30.

Toutanova K, A. Haghighi, Manning C. *A Global Joint Model for Semantic Role Labeling* in "Computational Linguistics" vol. 34 n. 2.

Widdows D. and Ferraro K., *Semantic Vectors: a Scalable Open Source Package and Online Technology Management Application*. In Proceedings of the Sixth International Language Resources and Evaluation (LREC'08).

Testa M., Bolioli A, Dini L., and Mazzini G.: *Evaluation of a Semantically Oriented Dependency Grammar for Italian at EVALITA 2009* . In Proceedings of the 11th Conference of the Italian Association for Artificial Intelligence, 12th December 2009, Reggio Emilia, Italy, ISBN 978-88-903581-1-1 (2009)

.