VENSES++: Adapting a deep semantic processing system to the identification of null instantiations

Sara Tonelli
Fondazione Bruno Kessler
Trento, Italy.
satonelli@fbk.eu

Rodolfo Delmonte
Università Ca' Foscari
Venezia, Italy.
delmont@unive.it

Abstract

The system to spot INIs, DNIs and their antecedents is an adaptation of VENSES, a system for semantic evaluation that has been used for RTE challenges in the last 6 years. In the following we will briefly describe the system and then the additions we made to cope with the new task. In particular, we will discuss how we mapped the VENSES analysis to the representation of frame information in order to identify null instantiations in the text.

1 Introduction

The SemEval-2010 task for linking events and their participants in discourse (Ruppenhofer et al., 2009) introduced a new issue w.r.t. the SemEval-2007 task "Frame Semantic Structure Extraction" (Baker et al., 2007), in that it focused on linking local semantic argument structures across sentence boundaries. Specifically, the task included first the identification of frames and frame elements in a text following the FrameNet paradigm (Baker et al., 1998), then the identification of locally uninstantiated roles (NIs). If these roles are indefinite (INI), they have to be marked as such and no antecedent has to be found. On the contrary, if they are definite (DNI), their coreferents have to be found in the wider discourse context. The challenge comprised two tasks, namely the *full task* (semantic role recognition and labelling + NI linking) and the *NIs only task*, i.e. the identification of null instantiations and their referents given a test set with gold standard local semantic argument structure.

We took part to the NIs only task by modifying the VENSES system for deep semantic processing and entailment recognition (Delmonte et al., 2005). In our approach, we assume that the identification of null instantiations is a complex task requiring different levels of semantic knowledge and several processing steps. For this rea-

son, we believe that the rich analysis performed by the pipeline architecture of VENSES is particularly suitable for the task, also due to the small amount of training data available and the heterogeneity of NI phenomena.

2 The VENSES system

VENSES is a reduced version of GETARUNS (Delmonte, 2008), a complete system for text understanding, whose backbone is LFG theory in its original version (Bresnan, 1982 and 2000). The system produces different levels of analysis, from syntax to discourse. However, three of them contribute most to the NI identification task: the lexico-semantic, the anaphora resolution and the deep semantic module.

2.1 The syntactic and lexico-semantic module

The system produces a c(onstituent)-structure representation by means of a cascade of augmented FSA, then it uses this output to map lexical information from a number of different lexica which however contain similar information related to verb/adjective and noun subcategorization. The mapping is done by splitting sentences into main and subordinate clauses. Other clauses are computed in their embedded position and can be either complement or relative clauses.

The system output is an Augmented Head Dependent Structure (AHDS), which is a fully indexed logical form, with Grammatical Relations and Semantic Roles. The inventory of semantic roles we use is however very small – 35, even though it is partly overlapping the one proposed in the first FrameNet project. We prefer to use generic roles rather than specific Frame Elements (FEs) because sense disambiguation at this stage of computation may not be effective.

2.2 The anaphora resolution module

The AHDS structure is passed to and used by a full-fledged module for pronominal and anaphora resolution, which is in turn split into *two submodules*. The resolution procedure takes care only of third person pronouns of all kinds – reciprocals, reflexives, possessive and personal. Its mechanisms are quite complex, as described in (Delmonte et al., 2006). The *first submodule* basically treats all pronouns at sentence level – that is, taking into account their position – and if they are left free, they receive the annotation "external". If they are bound, they are associated to an antecedent's index; else they might also be interpreted as expletives, i.e. they receive a label that prevents the following submodule to consider them for further computation.

The *second submodule* receives as input the external pronouns, and tries to find an antecedent in the previous stretch of text or discourse. To do that, the systems computes a *topic hierarchy* that is built following suggestions by (Sidner and Grosz, 1986) and is used in a centering-like manner.

2.3 The semantic module

The output of the anaphora resolution module is used by the semantic module to substitute the pronoun's head with the antecedent's head. After this operation, the module produces Predicate-Argument Structures or PAS on the basis of a previously produced Logical Form. PAS are produced for each clause and they separate obligatory from non-obligatory arguments, and these from adjuncts and modifiers. Some adjuncts, like spatiotemporal locations, are only bound at propositional level.

3 From VENSES output to NIs identification and binding

After computing PAS information for each sentence, we first map the test set gold standard annotation of frame information to VENSES output. Starting from the PAS with frames and FE labels attached to the predicates and the arguments, we run a module for DNI/INI spotting and DNI binding. It is composed by two different submodules, one for *verbal* predicates and one for *nominal* ones.

3.1 NIs identification and binding with verbal predicates

As pointed out in (Ruppenhofer et al., 2009), the identification of DNI/INIs includes three main steps: i) recognizing that a core role is missing ii) ascertaining if it has a definite interpretation and iii) if yes, finding a role filler for it.

For verbal predicates, the two first steps are accomplished starting from the PAS structure produced by VENSES and trying to map them with the valence patterns in FrameNet. To this purpose, we take into account the list of all valence patterns extracted for every LU and every frame from FrameNet 1.4 and from the training data, in which all possible sequences of FEs (both overtly expressed and null instantiated) are listed with their grammatical functions, coreness status and frequencies. For example, the predicate "barbecue.v" in the APPLY_HEAT frame is characterized by two patterns, both occurring once. In the first, *Food* is the subject (ext) and *Cook* is constructionally not instantiated (cni). In the second, the peripheral FE *Time* is also present:

```
ssr(barbecue-v,apply_heat,[[[[food-
c,np,ext],[cook-c,cni,null]],1],[[[time-
p,pp,dep],[food-c,np,ext],[cook-
c,cni,null]],1]]).
```

The first step in our computation is selecting for the current predicate those patterns or templates that contain the same number of core arguments of the clause under analysis plus one. This is due to the fact that NIs are always core FEs. For example, if a test sentence contains the "barbecue.v" lexical unit labelled with the APPLY_HEAT frame and only the *Food* FE is overtly annotated, we look in the template list for all patterns in which "barbecue.v" appears with the *Food* FE and another implicit core FE (either INI or DNI). If "barbecue.v" is not present in the template list, we consider the templates of the other verbal lexical units in the same frame.

The second step is assessing the licensor of the omission, whether lexical or constructional. Here we only distinguish complement governing predicates and passive constructions. For example, if "barbecue.v" is attested in the template list both with an indefinite and with a definite instantiation of the *Cook* FE, we check if it occurs in the passive form in the test sentence. If yes, we infer that *Cook* has to be labelled as an indefinite null instantiation (INI). Another licensor of the omission could be the imperative form of the verb, which however has not been considered yet by our system.

If we assess that the null instantiation is not indefinite, we look for an antecedent of the NI and, if we find it, we label it as a DNI. Otherwise, we don't encode any information about

omitted roles. The strategy devised for searching for possible referential expressions is as follows:

1. Given the current PAS (with frame labels), look in the previous sentence(s) for comparable PAS. *Comparable* means that the predicate is the same or semantically related based on WordNet synsets.
2. If a comparable PAS is found, check if they share at least one argument slot – typically they should share the subject role.
3. If yes, look for the best head available in that PAS by semantic matching with the FE label as a referent for the DNI label in the current sentence. In case that does not produce any matching, we look into the list of all heads in FrameNet associated to the FE label and select the one present in the PAS that matches.

3.2 NIs identification and binding with nominal predicates

In order to identify DNI/INIs of *nominal* predicates, we take into account the History List produced by VENSES in the AHDS analysis, where all nominal heads describing Events, Spatial and Temporal Locations and Body Parts in the document are collected together with their current sentence ID. Such list is derived from WordNet general nouns.

Based on a computational lexicon of Common Sense Reasoning relations made available with ConceptNet 2.0 by MIT AI Lab (Liu and Singh, 2004), we first process the history list in order to identify the relations between nominal heads in different sentences. Such relations include inheritance and inferences. For instance, if the current sentence contains the nominal heads "door" or "window", they are connected to the "house" head, if it is present in the History List as a spatial location occurring in a previous sentence. For instance, sentence 42 of the test document n. 13 contains the noun "wall" as lexical unit of the ARCHITECTURAL_PART frame. In the History List, it is classified as a place. Also the noun "house" in sentence n. 7 (token 7) is classified as a place in the History List. Since ConceptNet allows us to infer a meronymy relation between "wall" and "house", we can derive the following information, saying that "place" in sent. 45, token 25, is related to "house", in sent. 7, token 7:

```
loc(42-25, place, wall, house-[7-7]).
```

Starting from this information, we then check which core FEs are overtly expressed in the test sentence for the "wall" lexical unit. As encoded in the FrameNet database, the ARCHITEC-

TURAL_PART frame has two core FEs, namely *Part* and *Whole*. Since *Part* is already present in sentence n. 45, we assume that *Whole* could be a candidate DNI. After looking up the relations between nominal heads identified in the previous step, we make the hypothesis that "house" be the antecedent of the *Whole* DNI. We then check if "house" appears as a head of the *Whole* FE either in the FrameNet database or in the training data of the SemEval task in order to perform some semantic verification. If this hypothesis is confirmed, we finally take the syntactic node headed by the antecedent as the best DNI referent. In our example, "house" is the head of the node 501, so we generate the following output, in which the Whole FE is identified with the node 501 (headed by "house") in sentence 7:

```
<fe id="s42_f5_e2" name="Whole">
<fenode idref="s7_501"/>
<flag name="Definite_Interpretation">
```

Note that, in case the antecedent does not appear as the head of the candidate FE, it is discarded and no information about NIs is generated. This is clearly a limit of our approach, because nominal predicates are never assigned an INI label.

4 System output and evaluation

The SemEval test data comprise two annotated documents extracted from Conan Doyle's novels. We report some statistics about the test data with gold standard annotation and a comparison with our system output in Table 1.

	Text 1	Text 2
N. of sentences	249	276
Gold standard data		
N. of DNIs	158	191
N. of INIs	115	245
System output		
N. of DNIs	35	30
N. of INIs	16	20
F-score	0.0121	

Table 1: Comparison between gold standard and system output

The amount of NIs detected by our system is much lower than the gold standard one, particularly for INIs. This depends partly on the fact that no specific strategy for INI detection with nominal predicates has been devised so far, as described in Section 3.2. Another problem is that a lot of DNIs in the gold standard don't get resolved, while our system always looks for a re-

ferent in case of DNIs and if it is not found, the procedure fails.

The issue of detecting which DNIs are liable not to have an explicit antecedent remains an open problem. In general, Ruppenhofer et al. (2009) suggest to treat the DNI identification and binding as a coreference resolution task. However, the only information available is in fact the label of the missing FE. The authors propose to obtain information about the likely fillers of a missing FE from annotated data sets, but the task showed that this procedure could be successful only in case all FE labels are semantically well identifiable: in fact many FE labels are devoid of any specific associated meaning. Furthermore, lexical fillers of a given semantic role in the FrameNet data sets can be as diverse as possible. For example, a complete search in the FrameNet database for the FE *Charges* will reveal heads like "possession, innocent, actions", where the significant portion of text addressed by the FE would be in the specification - i.e. "possession of a gun" etc. Only in case of highly specialized FEs there will be some help in the semantic characterization of a possible antecedent. Another open issue is the notion of context where the antecedent should be searched for, which is lacking an appropriate definition.

If we take into account our system results on Text 1, we notice that only 3 DNIs have been identified and linked to the correct antecedent, while the overall amount of exact matches including INIs is 7. However, in 21 other cases the system correctly identifies a null instantiated role and assigns the right FE label, but it either detects an INI instead of a DNI (and vice-versa), or it finds the wrong antecedent for the DNI. A similar performance is achieved on Text 2: no DNI has been linked to the correct antecedent, and in only 8 cases there is an exact match between the INIs identified by the system and those in the gold standard. However, in 18 cases a null instantiation is detected and assigned the correct FE label, even if either the referent or the definiteness label is wrong. Some evaluation metrics taking into account the different information layers conveyed by the system would help highlighting such differences and pointing out the NI identification steps that need to be consolidated.

5 Conclusions

In this paper, we have introduced VENSES++, a modified version of the VENSES system for deep semantic processing and entailment detection.

We described two strategies for the identification of null instantiations in a text, depending on the predicate class (either nominal or verbal).

The system took part to the SemEval task for NIs identification and binding. Even if the preliminary results are far from satisfactory, we were able to devise a general strategy for dealing with the task. Only 2 teams took part to the competition, and the first ranked system achieved F1 = 0.0140. This confirms that NI identification is a very challenging issue which can be hardly modeled. Anyway, it deserves further efforts, as various NLP applications could benefit from the effective identification of null instantiated roles, from SRL to coreference resolution and information extraction.

References

Baker, C., Ellsworth, M. and Erk, K. 2007. *Frame Semantic Structure Extraction*. In Proceedings of the 4th International Workshop on Semantic Evaluations. Prague, Czech Republic.

Baker, C. F., Fillmore, C. J., & Lowe, J. B. 1998. *The Berkeley FrameNet project*. In Proceedings of COLING-ACL-98, Montreal, Canada.

Bresnan, J. 2000. *Lexical-functional syntax*. Oxford: Blackwell.

Bresnan, J. (ed.). 1982. *The mental representation of grammatical relations*, The MIT Press, Cambridge.

Delmonte R., 2008. *Computational Linguistic Text Processing – Lexicon, Grammar, Parsing and Anaphora Resolution*, Nova Science, New York.

Delmonte, R., Tonelli, S., Piccolino Boniforti, M. A., Bristot, A. 2006. VENSES – A Linguistically-based System for Semantic Evaluation. In Proc. of the 1st PASCAL RTE Workshop.

Delmonte, R., Bristot, A., Piccolino Boniforti, M.A., and Tonelli, S. 2006. *Another Evaluation of Anaphora Resolution Algorithms and a Comparison with GETARUNS' Knowledge Rich Approach*, In Proc. of ROMAND 2006, Trento, pp. 3-10.

Grosz, B., and Sidner, C. 1986. *Attention, intentions and the structure of discourse*. Computational Linguistics, 12, 175–204.

Liu, H., and Singh, P. 2004. ConceptNet: a practical commonsense reasoning toolkit. At http://web.media.mit.edu/~push/ConceptNet.pdf.

Ruppenhofer, J., Sporleder, C., Morante, R., Baker, C. and Palmer, M. 2009. *SemEval-2010 Task 10: Linking Events and Their Participants in Discourse*. In Proc. of the HLT-NAACL Workshop on Semantic Evaluations: Recent Achievements and Future Directions. Boulder, Colorado.