Mapping between Compositional Semantic Representations and Lexical Semantic Resources: Towards Accurate Deep Semantic Parsing

Sergio Roa†‡, **Valia Kordoni**† and **Yi Zhang**†
Dept. of Computational Linguistics, Saarland University, Germany†
German Research Center for Artificial Intelligence (DFKI GmbH)†
Dept. of Computer Science, University of Freiburg, Germany‡
{sergior,kordoni,yzhang}@coli.uni-sb.de

Abstract

This paper introduces a machine learning method based on bayesian networks which is applied to the mapping between deep semantic representations and lexical semantic resources. A probabilistic model comprising Minimal Recursion Semantics (MRS) structures and lexicalist oriented semantic features is acquired. Lexical semantic roles enriching the MRS structures are inferred, which are useful to improve the accuracy of deep semantic parsing. Verb classes inference was also investigated, which, together with lexical semantic information provided by VerbNet and PropBank resources, can be substantially beneficial to the parse disambiguation task.

1 Introduction

Recent studies of natural language parsing have shown a clear and steady shift of focus from pure syntactic analyses to more semantically informed structures. As a result, we have seen an emerging interest in parser evaluation based on more theory-neutral and semantically informed representations, such as dependency structures. Some approaches have even tried to acquire semantic representations without full syntactic analyses. The so-called shallow semantic parsers build basic predicate-argument structures or label semantic roles that reveal the partial meaning of sentences (Carreras and Màrquez, 2005). Manually annotated lexical semantic resources like PropBank (Palmer et al., 2005), VerbNet (Kipper-Schuler, 2005), or FrameNet (Baker et al., 1998) are usually used as gold standards for training and evaluation of such systems. In the meantime, various existing parsing systems are also adapted to provide semantic information in their outputs. The obvious advantage in such an approach is that one can derive more fine-grained representations which are not typically available from shallow semantic parsers (e.g., modality and negation, quantifiers and scopes, etc.). To this effect, various semantic representations have been proposed and used in different parsing systems. Generally speaking, such semantic representations should be capable of embedding shallow semantic information (i.e., predicate-argument or semantic roles). However, it is non-trivial to map even the basic predicate-arguments between different representations. This becomes a barrier to both sides, making the cross-fertilization of systems and resources using different semantic representations very difficult.

In this paper, we present a machine learning approach towards mapping between deep and shallow semantic representations. More specifically, we use Bayesian networks to acquire a statistical model that enriches the Minimal Recursion Semantics structures produced by the English Resource Grammar (ERG) (Flickinger, 2002) with VerbNet-like semantic roles. Evaluation results show that the mapping from MRS to semantic roles is reliable and beneficial to deep parsing.

2 Minimal Recursion Semantics

The semantic representation we are interested in in this paper is the Minimal Recursion Semantics (MRS). Because of its underspecifiability, it has been widely used in many deep and shallow processing systems. The main assumption behind MRS is that the interesting linguistic units for computational semantics are the *elementary predications* (EPs), which are single relations with associated arguments (Copestake et al., 2006). In this paper, the MRS structures are created with the English Resource Grammar (ERG), a HPSG-based broad coverage precision grammar for English. The seman-

tic predicates and their linguistic behaviour (including the set of semantic roles, indication of optional arguments, and their possible value constraints are specified by the grammar as its semantic interface (SEM-I) (Flickinger et al., 2005).

3 Relating MRS structures to lexical semantic resources

3.1 Feature extraction from linguistic resources

The first set of features used to find corresponding lexical semantic roles for the MRS predicate arguments are taken from Robust MRS (RMRS) structures (Copestake, 2006). The general idea of the process is to traverse the bag of elementary predications looking for the verbs in the parsed sentence. When a verb is found, then its arguments are taken from the **rarg** tags and alternatively from the **in-g** conjunctions related to the verb. So, given the sentence:

(1) Yields on money-market mutual funds continued to slide, amid signs that portfolio managers expect further declines in interest rates.

the obtained features for *expect* are shown in Table 1.

SEM-I roles	Features	Words
ARG1	`manager_n_of`	managers
ARG2	`propositional_m_rel`	further declines

Table 1: RMRS features for the verb *expect*

The SEM-I role labels are based mainly on syntactic characteristics of the verb. We employed the data provided by the PropBank and VerbNet projects to extract lexical semantic information. For PropBank, the argument labels are named ARG1,..., ARGN and additionally ARGM for adjuncts. In the case of VerbNet, 31 different thematic roles are provided, e.g. *Actor, Agent, Patient, Proposition, Predicate, Theme, Topic*. A treebank of RMRS structures and derivations was generated by using the PropBank corpus. The process of RMRS feature extraction was applied and a new verb dependency trees dataset was created.

To obtain a correspondence between the SEM-I role labels and the PropBank (or VerbNet) role labels, a procedure which maps these labellings for each utterance and verb found in the corpus was implemented. Due to the possible semantic roles that subjects and objects in a sentence could bear, the mapping between SEM-I roles and VerbNet role labels is not one-to-one. The general idea of this alignment process is to use the words in a given utterance which are selected by a given role label, both a SEM-I and a PropBank one. With these words, a naive assumption was applied that allows a reasonable comparison and alignment of these two sources of information. The naive assumption considers that if all the words selected by some SEM-I label are found in a given PropBank (VerbNet) role label, then we can deduce that these labels can be aligned. An important constraint is that all the SEM-I labels must be exhausted. An additional constraint is that ARG1, ARG2 or ARG3 SEM-I labels cannot be mapped to ARGM PropBank labels. When an alignment between a SEM-I role and a corresponding lexical semantic role is found, no more mappings for these labels are allowed. For instance, given the example in Table 1, with the following Propbank (VerbNet) labelling:

(2) $[_{\text{Arg}_0(\text{Experiencer})}$ Portfolio managers] *expect* $[_{\text{Arg}_1(\text{Theme})}$ further declines in interest rates.]

the alignment shown in Table 2 is obtained.

SEM-I roles	Mapped roles	Features
ARG1	Experiencer	`manager_n_of`
ARG2	Theme	`propositional_m_rel`

Table 2: Alignment instance obtained for the verb *expect*

Since the use of fine-grained features can make the learning process very complex, the WordNet semantic network (Fellbaum, 1998) was also employed to obtain generalisations of nouns. The algorithm described in (Pedersen et al., 2004) was used to disambiguate the sense, given the heads of the verb arguments and the verb itself (by using the mapping from VerbNet senses to WordNet verb senses (Kipper-Schuler, 2005)). Alternatively, a *naive* model has also been proposed, in which these features are simply generalized as nouns. For prepositions, the ontology provided by the SEM-I was used. Other words like adjectives or verbs in arguments were simply generalised as their corresponding type (e.g., *adjectival_rel* or *verbal_rel*).

190

3.2 Inference of semantic roles with Bayesian Networks

The inference of semantic roles is based on training of BNs by presenting instances of the features extracted, during the learning process. Thus, a training example corresponding to the features shown in Table 2 might be represented as Figure 1 shows, using a first-order approach. After training, the network can infer a proper PropBank (VerbNet) semantic role, given some RMRS role corresponding to some verb. The use of some of these features can be relaxed to test different alternatives.
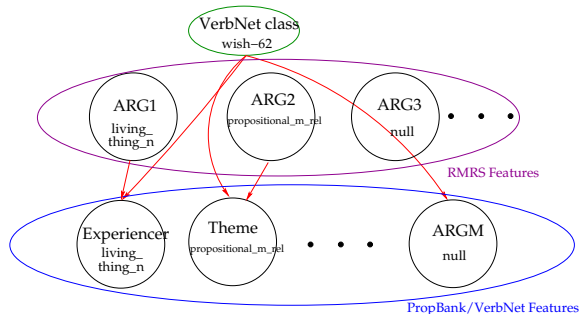

Figure 1: A priori structure of the BN for lexical semantic roles inference.

Two algorithms are used to train the BNs. The Maximum Likelihood (ML) estimation procedure is used when the structure of the model is known. In our experiments, the a priori structure shown in Figure 1 was employed. In the case of the Structural Expectation Maximization (SEM) Algorithm, the initial structure assumed for the ML algorithm serves as an initial state for the network and then the learning phase is executed in order to learn other conditional dependencies and parameters as well. The training procedure is described in Figure 2.

procedure *Train (Model)*
1: **for all** Verbs **do**
2:     **for all** Sentences and Parsings which include the current verb **do**
3:         Initialize vertices of the network with SEM-I labels and features.
4:         Initialize optionally vertices with the corresponding VerbNet class.
5:         Initialize edges connecting corresponding features.
6:         Append the current features as evidence for the network.
7:     **end for**
8:     Start Training Model for the current Verb, where Model is ML or SEM.
9: **end for**

Figure 2: Algorithm for training Bayesian Networks for inference of lexical semantic roles

After the training phase, a testing procedure using the Markov Chain Monte Carlo (MCMC) inference engine can be used to infer role labels. Since it is reasonable to think that in some cases the VerbNet class is not known, the presentation of this feature as evidence can be left as optional. Thus, after presenting as evidence the SEM-I related features, a role label with highest probability is obtained after using the MCMC with the current evidence.

4 Experimental results

The experiment uses 10370 sentences from the PropBank corpus which have a mapping to VerbNet (Loper et al., 2007) and are successfully parsed by the ERG (December 2006 version). Up to 10 best parses are recorded for each sentence. The total number of instances, considering that each sentence contains zero or more verbs, is 13589. The algorithm described in section 3.1 managed to find at least one mapping for 10960 of these instances (1020 different verb lexemes). If the number of parsing results is increased to 25 the results are improved (1460 different verb lexemes were found). In the second experiment, the sentences without VerbNet mappings were also included.

The results for the probabilistic models for infering lexical semantic roles are shown in Table 3, where the term *naive* means that no WordNet features were included in the training of the models, but only simple features like *noun_rel* for nouns. On the contrary, when mode is *complete*, WordNet hypernyms up to the 5th level in the hierarchy were used. In this set of experiments the VerbNet class was also included (in the marked cases) during the learning and inference phases.

Corpus	Nr. iter. MCMC	Mode	Model	Verb classes	Accuracy %
PropBank with VerbNet labels	1000	ML	naive		78.41
	10000	ML	naive		84.48
	10000	ML	naive	×	87.92
	1000	ML	complete		84.74
	10000	ML	complete		86.79
	10000	ML	complete	×	87.76
	1000	SEM	naive		84.25
	1000	SEM	complete		87.26
PropBank with PropBank labels	1000	ML	naive		87.46
	1000	SEM	naive		90.27

Table 3: Results of role mapping with probabilistic model

In Table 3, the errors are due to the problems introduced by the alternation behaviour of the verbs, which are not encoded in the SEM-I labelling and

also some contradictory annotations in the mapping between PropBank and VerbNet. Furthermore, the use of the WordNet features may also generate a more complex model or problems derived from the disambiguation process and hence produce errors in the inference phase. In addition, it is reasonable to use the VerbNet class information in the learning and inference phases, which in fact improves slightly the results. The outcomes also show that the use of the SEM algorithm improves accuracy slightly, meaning that the conditional dependency assumptions were reasonable, but still not perfect.

The model can be slightly modified for verb class inference, by adding conditional dependencies between the VerbNet class and SEM-I features, which can potentially improve the parse disambiguation task, in a similar way of thinking to (Fujita et al., 2007). For instance, for the following sentence, we derive an incorrect mapping for the verb *stay* to the VerbNet class EXIST-47.1-1 with the (falsely) favored parse where the PP *"in one place"* is treated as an adjunct/modifier. For the correct reading where the PP is a complement to *stay*, the mapping to the correct VerbNet class LODGE-46 is derived, and the correct LOCATION role is identified for the PP.

(3) Regardless of whether [$_{Theme}$ you] hike from lodge to lodge or **stay**$_{LODGE-46}$ [$_{Location}$ *in one place*] and take day trips, there are plenty of choices.

5 Conclusions and Future Work

In this paper, we have presented a study of mapping between the HPSG parser semantic outputs in form of MRS structures and lexical semantic resources. The experiment result shows that the Bayesian network reliably maps MRS predicate-argument structures to semantic roles. The automatic mapping enables us to enrich the deep parser output with semantic role information. Preliminary experiments have also shown that verb class inference can potentially improve the parse disambiguation task. Although we have been focusing on improving the deep parsing system with the mapping to annotated semantic resources, it is important to realise that the mapping also enables us to enrich the shallow semantic annotations with more fine-grained analyses from the deep grammars. Such analyses can eventually be helpful for applications like question answering, for instance, and will be investigated in the future.

References

Collin Baker, Charles Fillmore, and John Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of the 36th Annual Meeting of the ACL and 17th International Conference on Computational Linguistics*, pages 86–90, San Francisco, CA.

Xavier Carreras and Lluís Màrquez. 2005. Introduction to the CoNLL-2005 shared task: Semantic role labeling. In *Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005)*, pages 152–164, Ann Arbor, Michigan.

Ann Copestake, Dan P. Flickinger, and Ivan A. Sag. 2006. Minimal recursion semantics: An introduction. *Research on Language and Computation*, 3(4):281–332.

Ann Copestake. 2006. Robust minimal recursion semantics. Working Paper.

Christiane D. Fellbaum. 1998. *WordNet – An Electronic Lexical Database*. MIT Press.

Dan Flickinger, Jan T. Lønning, Helge Dyvik, Stephan Oepen, and Francis Bond. 2005. SEM-I rational MT. Enriching deep grammars with a semantic interface for scalable machine translation. In *Proceedings of the 10th Machine Translation Summit*, pages 165 – 172, Phuket, Thailand.

Dan Flickinger. 2002. On building a more efficient grammar by exploiting types. In Stephan Oepen, Dan Flickinger, Jun'ichi Tsujii, and Hans Uszkoreit, editors, *Collaborative Language Engineering*, pages 1–17. CSLI Publications.

Sanae Fujita, Francis Bond, Stephan Oepen, and Takaaki Tanaka. 2007. Exploiting semantic information for hpsg parse selection. In *ACL 2007 Workshop on Deep Linguistic Processing*, pages 25–32, Prague, Czech Republic.

Karin Kipper-Schuler. 2005. *VerbNet: A broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, University of Pennsylvania.

Edward Loper, Szu ting Yi, and Martha Palmer. 2007. Combining lexical resources: Mapping between Propbank and Verbnet. In *Proceedings of the 7th International Workshop on Computational Linguistics*, Tilburg, the Netherlands.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106.

Ted Pedersen, Siddharth Patwardhan, and Jason Michelizzi. 2004. WordNet::Similarity - Measuring the Relatedness of Concepts. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*.