Exploiting Event Semantics to Parse the Rhetorical Structure of Natural Language Text

Rajen Subba
Department of Computer Science
University of Illinois at Chicago
Chicago, IL 60613
rsubba@cs.uic.edu

Abstract

Previous work on discourse parsing has mostly relied on surface syntactic and lexical features; the use of semantics is limited to shallow semantics. The goal of this thesis is to exploit event semantics in order to build discourse parse trees (DPT) based on informational rhetorical relations. Our work employs an Inductive Logic Programming (ILP) based rhetorical relation classifier, a Neural Network based discourse segmenter, a bottom-up sentence level discourse parser and a shift-reduce document level discourse parser.

1 Introduction

Discourse is a structurally organized set of *coherent* text segments. The minimal unit of discourse is called an elementary discourse unit (EDU). An EDU or a span of EDUs constitute a segment. When we read text, we automatically assign *rhetorical* (coherence) relations to segments of text that we deem to be related. Consider the segmented text below:

(Example 1) [Clean the walls thoroughly$_{(1a)}$] [and allow them to dry.$_{(1b)}$] [If the walls are a dark color,$_{(2a)}$] [apply primer.$_{(2b)}$] [Put a small amount of paste in the paint tray;$_{(3a)}$] [add enough water$_{(4a)}$] [to thin the paste to about the consistency of cream soup.$_{(4b)}$]

It is plausible to state that the rhetorical relation between $(1a)$ and $(1b)$ is *preparation:act*. We can also posit that the relation *act:goal* holds between $(4a)$ and $(4b)$. Figure 1 shows the complete annotation of the full text. Now, if we were to reorder these segments as $[(1b), (4a), (2a), (4b), (3a), (2b), (1a)]$, the text would not make much sense. Therefore, it is imperative that the contiguous spans of discourse be coherent for comprehension. Rhetorical relations help make the text coherent.

Rhetorical relations based on the subject matter of the segments are called informational relations. A common understanding in discourse study is that informational relations are based on the underlying content of the text segments. However, previous work (Marcu, 2000; Polanyi et al., 2004; Soricut and Marcu, 2005; Sporleder and Lascarides, 2005) in discourse parsing has relied on syntactic and lexical information, and *shallow* semantics only.

The goal of this thesis is to build a computational model for parsing the informational structure of instructional text that exploits "deeper semantics", namely event semantics. Such discourse structures can be useful for applications such as information extraction, question answering and intelligent tutoring systems. Our approach makes use of a neural network discourse segmenter, a rhetorical relation classifier based on ILP and a discourse parsing model that builds sentence level DPTs bottom-up and document level DPTs using a shift-reduce parser.

In section 2, we describe how we collected our data. In section 3, we present our automatic discourse segmenter. Section 4 details our discourse parsing model based on event semantics followed by the conclusion in section 5.
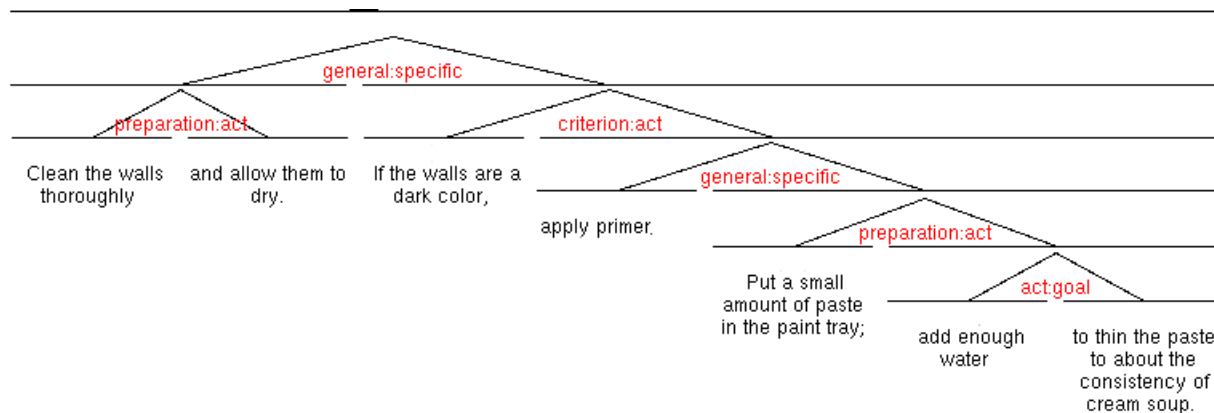
Figure 1: Discourse Annotation for Example 1

2 Data Collection

Our work calls for the use of a supervised machine learning approach. Therefore, we have manually annotated a corpus of instructional text with rhetorical relations and event semantic information. We used an existing corpus on home repair manuals (5Mb).[1]

2.1 Manual Discourse Annotation

In order to carry out the manual discourse annotation, a coding scheme was developed based on Marcu (1999) and RDA (Moser et al., 1996). The annotated data consists of 5744 EDUs and 5131 relations with a kappa value of 0.66 on about 26% of the corpus. We analyzed a total of 1217 examples to determine whether a cue phrase was present or not. Only 523 examples (43%) were judged to be signalled. Furthermore, discourse cues can be ambiguous with regard to which relation they signal. In order to account for cases where discourse cues are not present and to resolve such ambiguities, we intend to exploit event semantics.

2.2 Semi-Automatic Event Semantic Annotation

Informational relations describe how the content of two text segments are related. Therefore, it makes intuitive sense that verb semantics can be useful in determining these relations.[2] In Subba et al. (2006),

we integrated LCFLEX (Rose and Lavie, 2000) with VerbNet (Kipper et al., 2000) and CoreLex (Buitelaar, 1998) to compositionally build verb based event semantic representations of our EDUs.

VerbNet groups together verbs that undergo the same syntactic alternations and share similar semantics. It accounts for about 4962 distinct verbs classified into 237 main classes. The semantic information is described in terms of an event that is decomposed into four stages, namely *start*, *during*, *end* and *result*. Semantic predicates like *motion* and *together* describe the participants of an event at various stages. CoreLex provides meaning representations for about 40,000 nouns that are compatible with VerbNet.

The parser was used to semi-automatically annotate both our training and test data. Since the output of the parser can be ambiguous with respect to the verb sense, we manually pick the correct sense.[3]

3 Automatic Discourse Segmentation

The task of the discourse segmenter is to segment sentences into EDUs. In the past, the problem of sentence level discourse segmentation has been tackled using both symbolic methods (Polanyi et al., 2004; Huong et al., 2004) as well as statistical models (Soricut and Marcu, 2003; Marcu, 2000) that have exploited syntactic and lexical features.

We have implemented a Neural Network model

---

[1] The corpus was collected opportunistically off the internet and from other sources, and originally assembled at the Information Technology Research Institute, University of Brighton.

[2] Especially in instructional manuals where the meaning of most sentences is centered on verbs.

[3] In addition, the parser generates semantic representations for fragments of the sentence to handle ungrammatical sentences, etc.

for sentence level discourse segmentation that uses syntactic features and discourse cues. Our model was trained and tested on RST-DT (2002) and achieves a performance of up to 86.12% F-Score, which is comparable to Soricut and Marcu (2003). We plan to use this model on our corpus as well.

4 Discourse Parsing

Once the EDUs have been identified by the discourse segmenter, the entire discourse structure of text needs to be constructed. This concerns determining which text segments are related and what relation to assign to those segments. Our discourse parsing model consists of a rhetorical relation classifier, a sentence level discourse parser and a document level discourse parser.

4.1 Rhetorical Relation Classifier

In a preliminary investigation (Subba et al., 2006), we modeled the problem of identifying rhetorical relations as a classification problem using rich verb semantics only.

Most of the work in NLP that involves learning has used more traditional machine learning paradigms like decision-tree algorithms and SVMs. However, we did not find them suitable for our data which is represented in first order logic (FOL). We found Progol (Muggleton, 1995), an ILP system, appropriate for our needs. The general problem specification for Progol (ILP) is given by the following posterior sufficiency property:

$$B \wedge H \models E$$

Given the background knowledge B and the examples E, Progol finds the simplest consistent hypothesis H, such that B and H entails E. The rich verb semantic representation of pairs of EDUs form the background knowledge and the manually annotated rhetorical relations between the pairs of EDUs serve as the positive examples.[4] An A*-like search is used to search for the most probable hypothesis. Given our model, we are able to learn rules such as the ones given in Figure 2. Due to the lack of space we only explain RULE1 here. RULE1 states that

---

RULE1:

relation(EDU1,EDU2,'before:after') :- motion(EDU1,event0,during,C),

       location(EDU2,event0,start,C,D).

RULE2:

relation(EDU1,EDU2,'act:goal') :- cause(EDU1,C,event0),

       together(EDU1,event0,end,physical,F,G),cause(EDU2,C,event0).

Figure 2: Examples of Rules learned by Progol

there is a theme (C) in motion during the event in EDU1 (the first EDU) and that C is located in location D at the start of the event in EDU2 (the second EDU).

We trained our classifier on 423 examples and tested it on 85 examples.[5] A majority function baseline performs at a 51.7 F-Score. Our model outperforms this baseline with an F-Score of 60.24.

Relation	Precision	Recall	F-Score
goal:act	31.57	26.08	28.57
step1:step2	**75**	**75**	**75**
before:after	54.5	54.5	54.5
criterion:act	71.4	71.4	71.4
Total	**61.7**	**58.8**	**60.24**

Table 1: Rhetorical Relation Classifier Result

This study has shown that it is possible to learn rules from FOL semantic representations using Inductive Logic Programming to classify rhetorical relations. However, it is not yet clear how useful event semantics is for discourse parsing. In the future, we intend to extend our model to incorporate syntactic and lexical information as well. Such an extension will allow us to assess the contribution of event semantics.

4.2 Building Discourse Parse Trees

In addition to extending the rhetorical relation classifier, our future work will involve building the discourse parse tree at the sentence level and at the document level. At the document level, the input will be the sentence level discourse parse trees and the output will be the discourse structure of the entire

---

[4]The output from the parser was further processed into definite clauses. Positive examples are represented as ground unit clauses.

[5]For this preliminary experiment, we decided to use only those relation sets that had more than 50 examples and those that were classified as *goal:act*, *step1:step2*, *criterion:act* or *before:after*

document.

When combining two text segments, promotion sets that approximate the most important EDUs of the text segments will be used. As a starting point, we propose to build sentence level DPTs bottom-up. EDUs that are subsumed by the same syntactic constituent (usually an S, S-Bar, VP) will be combined together into a larger text segment recursively until the the DPT at the root level has been constructed. At the document level, the DPT will be built using a shift-reduce parser as in Marcu (2000). However, unlike Marcu (2000), there will only be one shift and one reduce operation. The reduce operation will be determined by the rhetorical relation classifier and an additional module that will determine all the possible attachment points for an incoming sentence level DPT. An incoming sentence level DPT may be attached to any node on the right frontier of the left DPT. Lexical cohesion will be used to rank the possible attachment points. For both sentence level discourse parsing and document level discourse parsing, the rhetorical relation classifier will be used to determine the informational relation between the text segments.

5 Conclusion

In conclusion, this thesis will provide a computational model for parsing the discourse structure of text based on informational relations. Our approach exploits event semantic information of the EDUs. Hence, it will provide a measurement of how helpful event semantics can be in uncovering the discourse structure of text. As a consequence, it will also shed some light on the coverage of the lexical resources we are using. Other contributions of our work include a parser that builds event semantic representations of sentences based on rich verb semantics and noun semantics and a data driven automatic discourse segmenter that determines the minimal units of discourse.

References

Buitelaar, P.: CoreLex: Systematic Polysemy and Underspecification. Ph.D. thesis, Computer Science, Brandeis University, February 1998.

Huong Le Thanh, G. A. and Huyck., C.: Automated discourse segmentation by syntactic information and cue phrases. International Conference on Artificial Intelligence and Applications, 2004.

Kipper, K., H. T. D. and Palmer., M.: Class-based construction of a verb lexicon. AAAI-2000, Proceedings of the Seventeenth National Conference on Artificial Intelligence, 2000.

Livia Polanyi, Christopher Culy, M. H. v. d. B. G. L. T. and Ahn., D.: Sentential structure and discourse parsing. ACL 2004, Workshop on Discourse Annotation, 2004.

Marcu, D.: Instructions for Manually Annotating the Discourse Structures of Texts. Technical Report, University of Southern California, 1999.

Marcu, D.: The theory and practice of discourse parsing and summarization. Cambridge, Massachusetts, London, England, MIT Press, 2000.

Moser, M. G., Moore, J. D., and Glendening, E.: Instructions for Coding Explanations: Identifying Segments, Relations and Minimal Units. University of Pittsburgh, Department of Computer Science, 1996.

Muggleton., S. H.: Inverse entailment and progol. In New Generation Computing Journal, 13:245–286, 1995.

Rosé, C. P. and Lavie., A.: Balancing robustness and efficiency in unification-augmented context-free parsers for large practical applications. In Jean-Clause Junqua and Gertjan van Noord, editors, Robustness in Language and Speech Technology, 2000.

RST-DT.: Rst discourse treebank. Linguistic Data Consortium., 2002.

Sporleder, C. and Lascarides., A.: Exploiting linguistic cues to classify rhetorical relations. Recent Advances in Natural Language Processing, 2005.

Soricut, R. and Marcu., D.: Sentence level discourse parsing using syntactic and lexical information. Proceedings of the HLT and NAACL Conference, 2003.

Subba, R., Di Eugenio, B., E. T.: Building lexical resources for princpar, a large coverage parser that generates principled semantic representations. LREC 2006, 2006.

Subba, R., Di Eugenio, B., S. N. K.: Learning FOL rules based on rich verb semantic representations to automatically label rhetorical relations. EACL 2006, Workshop on Learning Structured Information in Natural Language Applications, 2006.

Wellner, B., Pustejovsky, J., C. H. R. S. and Rumshisky., A.: Classification of discourse coherence relations: An exploratory study using multiple knowledge sources. SIGDIAL Workshop on Discourse and Dialogue, 2006.