Finding Discourse Relations in Student Essays

Rohana Mahmud and Allan Ramsay

School of Informatics, University of Manchester
R.Mahmud@postgrad.umist.ac.uk,
Allan.Ramsay@manchester.ac.uk

Abstract. The aim of the work reported here is to provide a tool to help secondary school (high school) age students to reflect on the structure of their essays. Numerous tools are available to help students check their spelling and grammar. Very little, however, has been done to help them with higher level problems in their texts. In order to do this, we need to be able to analyse the discourse relations within their texts. This is particularly problematic for texts of this kind, since they contain few instances of explicit discourse markers such as '*however*', '*moreover*', '*therefore*'. The situation is made worse by the fact that many texts produced by such students contain large numbers of spelling and grammatical errors, thus making linguistic analysis extremely challenging. The current paper reports on a number of experiments in classification of the discourse relations in such essays. The work explores the use of machine learning techniques to identify such relations in unseen essays, using a corpus of manually annotated essays as a training set.

1 Introduction

Secondary school students have numerous problems when trying to compose extended texts. They have low-level errors with spelling and grammar, for which a range of support tools exists. But they also have problems with organising their texts into coherent well-structured discourses, and they have particular problems using the devices that the language places at their disposal for indicating the structure to the reader (e.g. lexical cohesion relations and careful construction of referential chains). The underlying aim of the work reported here, then, is to provide a tool which will reveal the discourse cues that are present in such essays, and hence to allow students to reflect on what they have written.

There is, clearly, no such thing as the 'right' way to structure an essay. There are correct ways to spell things (though no extant spell-checker gets them all right), and there are correct and incorrect grammatical forms (though no extant parser can be relied on to pass all grammatically correct constructions and flag all grammatically incorrect ones). So it is, at least in principle, possible to produce a tool which will tell you whether you have spelt all the words in some document correctly, and whether all your sentences are grammatically acceptable. But there is no right or wrong structure for an extended text, so it simply makes no sense to talk of showing students where they have made 'errors' in the organisation of their texts. The best we can hope for is to show them where they have used discourse structuring markers, and to show them

A. Gelbukh (Ed.): CICLing 2005, LNCS 3406, pp. 116 – 119, 2005.

the structure that the clues they have used impose on their texts. The hope is that by making these things manifest we can help students see what the choices are and what their consequences are. If they can at least come to appreciate the importance of discourse clues, we will have provided them with a useful tool.

The problem is that identifying discourse relations is a very complex task and to date there is still no robust discourse parser (Marcu and Echihabi 2002). This process is more difficult when the texts under analysis contain large numbers of grammatical and spelling errors. However, students at a lower secondary school level (age approximately 12-14) frequently make such mistakes, and we have to be able to cope with texts containing low-level errors. To make matters even worse, such students seldom use explicit cue words such as '*however*' and '*even*' in their essays, so that algorithms that depend heavily on such terms will not work in this context.

Despite the absence of explicit cues, essays by students at this level clearly do display structure. Some students make appropriate use of lexical cohesion relations and of appropriate referential chains in organising their essays. Others, however, are less successful, and would clearly benefit from feedback in this area, thus providing motivation for the current study (Mahmud 2004), see also (Burstein, Marcu et al. 2003), (White and Arndt 1991).

2 The Experiment

A number of essays were collected from a school in North West England The essays were segmented into independent sentences simply using the standard delimiters '.', '!' and '?' (paying due attention to the use of '.' for marking the end of standard abbreviations, as in '*Mr.*'). The essays were then parsed using the PARASITE robust parser (Ramsay 2001) and a certain amount of linguistic information was recorded. This information was used as attributes for finding discourse relations. The key features are shown in table 1.0. A human annotator then determined which sentences were related and classified the relations that were found. This annotation did not require the discourse to be structured as a well-formed tree. The assumption underlying our work is that student essays do not always take the shape of well-formed trees (if they did, the tool we are building would not be needed!), so the links proposed during annotation were allowed to cross (which they did occasionally), and sentences were allowed to be unconnected to the remainder of the discourse (which happened quite frequently).

We then used the WEKA machine learning tool (Witten and Frank 2000) to acquire rules for classifying relations between sentences, based on the manual annotation. We used the following small set of relations, since it seemed very unlikely that the texts contained enough information to make learning a finer-grained classification possible:

i. Narrative (a sequence relation)
ii. Elaboration (gives more explanation of the other sentence)
iii. Contrast (if the pair sentences are contrast to each other)
iv. otherRelations (any other types of rhetorical relations (other than the above three))
v. noRelations (if the sentence is not related to any other sentences)

Table 1. Recorded Linguistic Information

Nucleus attributes (n)	Satellite attributes (s)
– n-position id	– s-position id
– n-cue words	– s-cue words
– referential status of n-subject NP	– referential status of s-subject NP
– n-mood	– s-mood
– n-head Verb	– s-head Verb
Pair-sentences attributes (p)	
– p-distance: distances between the nucleus and the satellite, can be negative or positive	
– p-cohesive: superordinate, subordinate, same or none based on the semantic relations of the head Verb; these relations were obtained from WordNet (A. Miller, Beckwith et al. 1993)	
– p-centers: the referential connection between the subject NP of the nucleus and satellite; cf. Centering Theory (Grosz, Joshi et al. 1995), but looking only at the subjects of the two sentences	

3 Results

In this experiment, we were trying to find other possible attributes that can help in identifying discourse relations if the corpus is contains few cue words. Using the RandomForest and RandomTree algorithms and 155 instances of pair-sentences produced 88.4% accuracy in classifying the discourse relations compared to the human annotator (the annotation was not carried out all that rigorously, so all we have actually shown is that the algorithms can learn this annotator's intuitions. However, if at least one annotator intuitions can be learnt then it is likely that the common intuitions of a wider group can also be learnt). The most important attributes turn out to be the s-subject and the p-distances. It is likely that lexical relations, particularly between the main verbs and between the head nouns of the subjects, are also significant, but the mechanisms we had for detecting such relations simply were not powerful enough to capture them.

4 Conclusion

In this paper, we have presented an experiment of finding discourse relations from noisy corpus. Although most of the literature on discourse relations uses cue words as the main attributes in finding discourse relations (Hutchinson 2003), (Marcu 2000), (Corston-Oliver 1998), (Knott 1996), we found other features like sentences-distances and the transitions of the subject NP can also be used as clue in developing a better discourse parser.

References

1. A.Miller, G., R. Beckwith, et al. *Introduction to WordNet: An On-line Lexical Database*, University of Princeton**:** 85, 1993
2. Burstein, J., D. Marcu, et al. Finding the WRITE Stuff: Automatic Identification of Discourse Structure in Student Essays. IEEE Intelligent Systems, 18 (January-February): 32-39, 2003
3. Corston-Oliver, S. H. *Computing Representations of the Structure of Written Discourse*. Santa Barbara, University of California**:** 256, 1998
4. Grosz, B. J., A. K. Joshi, et al. "Centering: A Framework for Modelling the Local Coherence of Discourse." *Computational Linguistics* 21(2): 203-225, 1995
5. Hutchinson, B. Automatic classification of discourse markers by their co-occurrences. *Proceedings of the ESSLLI'03 workshop on The Meaning and Implementation of Discourse Particles*, Vienna, Austria**:** 65-72, 2003
6. Knott, A. A Data-Driven Methodology for Motivating a Set of Coherence Relations. Department of Artificial Intelligence. Edinburgh, University of Edinburgh, 1996
7. Mahmud, R. *Possible attributes for identifying discourse relations structure*. CLUK2004, Birmingham, 2004
8. Marcu, D. *The Theory and Practice of Discourse Parsing and Summarization*. London England, The MIT Press, 2000
9. Marcu, D. and A. Echihabi. An Unsupervised Approach to Recognizing Discourse Relations. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (ACL-2002), Philadelphia, 2002
10. Ramsay, A. PrAgmatics = ReAsoning about the Speaker's InTEnsions—The Parasite Manual. Manchester, Computation Department, UMIST**:** 48, 2001
11. White, R. and V. Arndt. *Process Writing*. Essex, Addison Wesley Longman Limited, 1991
12. Witten, I. H. and E. Frank. *Data Mining - Practical machine Learning Tools and Techniques with Java Implementation*. San Francisco, Morgan Kaufmann, 2000