Using Cohesive Devices to Recognize Rhetorical Relations in Text

Huong Le Thanh, Geetha Abeysinghe, and Christian Huyck

School of Computing Science, Middlesex University,
The Burroughs, London NW4 4BT, UK.
{H.Le, G.Abeysinghe, C.Huyck}@mdx.ac.uk

Abstract

This paper investigates factors that can be used in discourse analysis, specifically, cohesive devices. The paper shows that cohesive devices such as cue phrases can provide information about the linkages inside a text. We propose three types of cue phrases (the ordinary cue phrases, noun-phrase cues, and verb-phrase cues). An algorithm to compute rhetorical relations between two elementary discourse units is also presented.

1 Introduction

Rhetorical Structure Theory (RST) (Mann and Thompson, 1988) offers an explanation of the coherence of texts. It models the discourse structure of a text by a hierarchical tree diagram that labels relations between text spans (typically clauses or larger linguistic units). There are two kinds of relations: nucleus-satellite relation and multinuclear relations. A nucleus-satellite relation involves two nodes in which one node has a specific role relative to the other. The more important node between them in realising the writer's communicative goals is called a *nucleus*; the less important one is called a *satellite*. A multinuclear relation involve two or more nodes, each of which is equally important in realising the writer's communicative goals. RST can be applied in many fields, such as automatic summarisation, text generation, and text indexing.

Analysing textual rhetorical structures is difficult because discourse can be complex and vague. Many approaches in this area use cue phrases (such as "but", "however") to recognise rhetorical relations (e.g. Marcu, 1997; Corston-Oliver, 1998; Webber, 2001) because of their efficiency and simplicity. Cue phrases show a great potential in discourse analysis because most cue phrases have a specific discourse role. They indicate a rhetorical relation between different parts of a text. However, these approaches have problems when no cue phrases are found, which frequently happens.

This research carries out a study on textual coherence devices in order to solve this problem. The discourse parser we proposed involves the following three steps. Firstly, we split text into elementary discourse units (EDUs)[1]. Secondly, after defining EDUs, all potential rhetorical relations between these units are discovered. Finally, based on this relation set, all rhetorical structures will be produced using a discourse parser to combine small texts into larger ones.

This paper discusses step 2 of our proposed system. Different factors that can be used in identifying relations among discourse units are analyzed in Section 2. Section 3 describes the relation set and the method for recognizing relations. We present our conclusions in Section 4.

2 Factors Used in Recognizing Relations

2.1 Cohesive Devices

Cohesive devices are not the unique way to make text coherence. However, they are chosen in our research because of their efficiency and simplicity. Salkie (1995) presented different types of cohesive devices. We have considered a few of them to be implemented in our system. They are categorised into four groups: reiterative devices, reference words, ellipsis, and cue phrases.

The reiterative devices include synonyms (employer/boss), superordinates/hyponyms (country/Mexico), co-hyponyms (United Kingdom/Mexico), and antonyms (simple/complex). These devices are good factors in recognizing rhetorical relations. For example, antonyms often express a CONTRAST relation.

Reference words include personal pronouns (he,

[1] For further information on "EDUs", see (Marcu, 1997).

she, it, etc.), demonstratives (this, that, these, those), and comparative constructions (the same thing, a different person, etc.). Reference words need help from their environment to determine their full meaning. Thus, they create links between texts.

To benefit from reiterative and reference words, we extract the main noun phrases from the actor and the object of sentences. A thesaurus is then used to find the semantic relation between these noun phrases.[2] This cohension participates in deciding rhetorical relations of text.

Another important cohesive device is ellipsis. This is a special form of substitution, where only a part of a sentence is omitted. Ellipsis can be found by analyzing the syntax of the sentence. Ellipsis often occurs in question/answer sequences. Therefore, ellipsis can be used to recognize a SOLUTIONHOOD relation (see Section 3).

2.2 Cue Phrases

Cue phrases (such as "however", "as a result"), sometime called connectives or conjunctions, are used to indicate a specific connection between different parts of a text. This is the strongest cohesive device due to two reasons. Firstly, most cue phrases have a rhetorical meaning. If two text spans are connected by a cue phrase, their relation will be determined by the cue phrase's rhetorical meaning. Secondly, identifying cue phrases is quite simple because it is essentially based on pattern matching. Meanwhile, syntactic information is needed in order to explore other text devices such as synonyms and antonyms. Because of its strength and simplicity, there are many approaches that use cue phrases to recognize rhetorical relations (e.g., Knott and Dale, 1995; Marcu, 1997). However, as mentioned before, these approaches have problems when no cue phrase is found.

Our solution to this problem is to further expand the cue phrase concept. We propose three kinds of cue phrases:

1. Ordinary cue phrase (called cue phrase).
2. Special words or phrases in the main noun phrase (subject or object) of a sentence (called noun-phrase cue or NP cue).
3. Special words or phrases in the verb phrase of a sentence (called verb-phrase cue or VP cue).

[2] We have chosen WordNet (WordNet, 2002), a machine-readable thesaurus and semantic network, for this purpose.

Cue phrases must match exactly, whereas noun phrases and verb phrases are stemmed before being compared with NP/VP cue. Examples of NP and VP cues are shown in example (1) and example (2) respectively, below.

(1) [New York style pizza meets California ingredients,] [and the ***result*** is the pizza from this Church Street pizzeria.]

(2) [Chairman Silas Cathcart retires to his Lake Forest.] [That ***means*** Michael Carpenter will take complete control of Kidder.]

The noun "***result***" indicates a RESULT relation in example (1); meanwhile the verb "***means***" determines an INTERPRETATION relation in example (2).

A word/phrase can be a cue word/phrase in some cases, but not in the others. For example, the word "*and*" is a cue word in example (3), but not in example (4) as shown below.

(3) [Mary borrowed that book from our library last Monday,] [***and*** she returned it this morning.]

SEQUENCE

(4) Mary has a cat ***and*** a dog.

Some phrases (e.g., "*in spite of*") have a discourse meaning in all of their occurrences. Thus, each cue phrase has a different effect in deciding rhetorical relations. To control their strength, scores are assigned to different cue phrases.

If a word/phrase always has a discourse meaning and represents only one rhetorical relation, it will get the highest score, 1. If a word/phrase always has a discourse meaning and represents N relations (e.g., the cue phrase *"although"* expresses an ANTITHESIS relation or a CONCESSION relation), the score of that cue phrase for each type of relation will be 1/N. If a cue phrase only has a discourse meaning in some cases (e.g., *"and"*), its maximum score will be lower than 1.

Examples (3) and (4) show that the word's position is also important in deciding the word's discourse role. Therefore, if a word or a phrase has a discourse meaning in only some special positions inside a sentence, the information about its position will be given to the word/ phrase. If a word/phrase has a discourse role irrespective of its position in the sentence, no information will be provided about its position.

For example, the word "*second*" only has a discourse meaning when it stands at the beginning of a clause/sentence (indicated by the letter "B"). It has 50% certainty to be a LIST relation (hence given a

score of 0.5). Then it will be stored in the cue phrases' set for the LIST relation as "*second*(B, 0.5)".

Similarly, NP cues and VP cues also have scores depending on their strength in deciding rhetorical relations. Information involving ordinary cue phrases, NP cues, and VP cues (such as the relations that the cue represents for, and relation's score) are stored in text files for further use.

3 Relation Set and Relation Recognition

To generate a rhetorical structure from text, we need to decide which rhetorical relations,[3] and how many relations are enough. If we define just a few relations, the rhetorical trees will be easy to construct, but they will not be very informative. On the other hand, if we have a large relation set, the trees will be very informative; but they will be difficult to construct.

The RST discourse corpus consists of 78 rhetorical relation types. It is difficult to automatically construct RST trees based on such a large relation set. Therefore, we define a smaller set but sufficient to characterize relations by grouping similar relations into one. Based on the rhetorical relations that have been proposed in the literature, e.g., (Mann and Thompson, 1988), and (Hovy, 1990), the following set of 22 relations has been chosen to be used in our system:

LIST, SEQUENCE, CONDITION, OTHERWISE, HYPOTHETICAL, ANTITHESIS, CONTRAST, CONCESSION, CAUSE, RESULT, CAUSE-RESULT, PURPOSE, SOLUTIONHOOD, CIRCUMSTANCE, MANNER, MEANS, INTERPRETATION, EVALUATION, SUMMARY, ELABORATION, EXPLANATION, and JOINT.

3.1 Relation Recognition

Similar to (Corston-Oliver, 1998), we divide the features that help us to recognize a rhetorical relation into two parts:

(1) the conditions that two text spans must satisfy in order *to accept* a specific relation between them;

(2) and, the tokens used for *predicting* a relation.

We call the features in part (1) the necessary conditions and the features in part (2) the cue set. A cue set consists of heuristic rules involving cue phrases, NP cues, VP cues, and cohesive devices. The necessary conditions ensure that the two text spans have no conflict with the definition of the relation being tested. The necessary conditions may not consist of any token to realise a specific relation. The system can only recognise a rhetorical relation between two units if all necessary conditions and at least one cue are satisfied.

3.2 Scoring Heuristic Rules

Cue phrases, NP cues, VP cues, and cohesive devices have different effects in deciding rhetorical relations. Therefore, it is necessary to assign a score to each heuristic rule. The cue phrase's rule has the highest score of 1, as cue phrases are the strongest signal. NP cues and VP cues are the extension cases of cue phrases. They are also strong cues, but weaker than normal cue phrases. Thus, the heuristic rules involving NP cues and VP cues have the score of 0.9. The cohesive devices have lower scores than NP cues and VP cues. Depending on their certainty, the heuristic rules corresponding to these devices receive the scores of 0.2 to 0.8. It is of interest to note that each score can be understood as the percentage of cases in which the cue recognises a correct rhetorical relation. These scores are first assigned to heuristic rules according to human linguistic intuitions. After building the whole system, different sets of scores will be tested in order to find the optimal scores for the system.

As mentioned is Section 2.2, each cue phrase, NP cue or VP cue has its own score. It follows that the actual score for those cues is:

Actual Score = Score(heuristic rule) * Score(cue phrase, or NP cue, or VP cue).

The final score of a relation is equal to the sum of all heuristic rules contributing to that relation. The system will test the necessary conditions of that relation if its final score is more than or equal to a threshold θ.[4]

In the following section, we analyze the LIST relation to illustrate the usage of necessary conditions, cue set, and scores in recognizing relations between two EDUs.

3.3 Algorithm for recognising relations between two EDUs

As mentioned in Section 3.1, the heuristics rules in the cue set provide a suggestion of relations between

[3] For further information on "rhetorical relation", see (Mann and Thompson, 1988).

[4] Threshold θ is selected as 0.5.

two text spans. Thus, we start detecting relations between two text spans by testing the cue set, from the highest score rule to the lowest one. If several relations are recommended, the necessary conditions of these relations are checked in order to find the appropriate relations. Due to lack of space, a detailed description of this process is not presented in this paper. The pseudo-code for recognising relations between two EDUs is shown below:

Input: Two EDUs U_1 and U_2, list of ordinary cue phrases (CPs), list of VP cues, and list of NP cues.
Output: Relation set {R} between U_1 and U_2.

1. Find all CPs of U_1 and U_2.
2. If CPs are found, compute actual score of the relations suggested by CPs.
3. Check necessary conditions (NCs) of the relations suggested by CPs whose actual score > θ.
4. Add the relations that satisfy NCs to {R}.
5. If no relation satisfies, go to step 6. Otherwise, Return.
6. Find the main VP of each unit and stem them.
7. If one of these stemmed VPs consists of a VP cue, compute actual score of the stemmed VP and total score.[5]
8. Check NCs of the relations corresponding to the VP cue whose total score > θ.
9. Add the relations that satisfy NCs to {R}.
10. If no relation satisfies, go to step 11. Otherwise, Return.
11. Find the subject of each unit and stem these NPs.
12. If one of these stemmed NPs consists of a NP cue, compute actual score of the stemmed NP and total score.
13. Check NCs of the relations corresponding to the NP cue whose total score > θ.
14. Add the relations that satisfy NCs to {R}.
15. If no relation satisfies, go to step 16. Otherwise, Return.
16. For each of 22 relations in the proposed relation set:
    16.1. Check the remaining cues of the current relation (the cues that do not involve ordinary CP, VP cues, and NP cues).
    16.2. Compute total score of the relations suggested by cues.
    16.3. Check NCs of the relations whose total score > θ.
    16.4. Add the relations that satisfy NCs to {R}.
17. Return.

In the following section, we analyze the LIST relation to illustrate the usage of necessary conditions, cue set, scores, and the algorithm for recognizing relations between two EDUs.

3.4 LIST Relation

A LIST is a multinuclear relation whose elements can be listed, but not in a CONTRAST or other stronger types of multinuclear relation (Carlson and Marcu, 2001). A LIST relation is often considered as a SEQUENCE relation if there is an explicit indication of temporal sequence.

The necessary conditions for a LIST relation between two units, $Unit_1$ and $Unit_2$, are shown below:

1. Two units are syntactically co-ordinates.
2. If both units have subjects and do not follow the reported style, then these subjects need to meet the following requirement: they must either be identical or be synonym, co-hyponym, or superordinate/hyponym; or the subject of $Unit_2$ is a pronoun or a noun phrase that can replace the subject of $Unit_1$.
3. There is no explicit indication that the event expressed by $Unit_1$ temporally precedes the event expressed by $Unit_2$.
4. The CONTRAST relation is not satisfied.

The first condition is based on syntactic information to guarantee that the two units are syntactically independent. The second condition checks the linkage between the two units by using reiterative and co-reference devices. The third condition distinguishes a LIST relation from a SEQUENCE relation. The last condition ensures that the stronger relation, CONTRAST, is not present in that context. In order to check this condition, the CONTRAST relation is always examined before the LIST relation.

The cue set of the LIST relation is shown below:

1. $Unit_2$ contains a LIST cue phrase. Score: 1
2. Both units contain enumeration conjunctions (*first*, *second*, *third*…). Score: 1
3. Both subjects of $Unit_1$ and $Unit_2$ contain NP cues. Score: 0.9
4. If both units are reported sentences, they mention the same object. Score: 0.8
5. If the subjects of two units are co-hyponyms, then the verb phrase of $Unit_2$ must be the same as

[5] Total score is the accumulated scores of heuristic rules up to the current time.

the verb phrase of $Unit_1$, or $Unit_2$ should have the structure "*so + auxiliary + sbj*". Score: 0.8

6. Both units are clauses in which verb phrases agree in tense (e.g., past, present). Score: 0.5

For example, the cue "*also*" in sentence (5.2) suggests a LIST relation between unit (5.1) and unit (5.2) in the following case: [6]

(5) [Mr. Cathcart is credited with bringing some basic budgeting to traditionally freewheeling Kidder.[5.1]] [He ***also*** improved the firm's compliance procedures for trading.[5.2]]

The actual score of cue 1, with the cue word "*also*", is equal to Score(cue 1) * Score("*also*"). The cue word "*also*" has the score of 1 for the LIST relation, so the actual score is $1*1=1>\theta$. Therefore, the necessary conditions of the LIST relation are checked. Text spans (5.1) and (5.2) are two sentences, thus they syntactically coordinate (condition 1). In addition, the subject of the text span (5.2), "*he*", is a pronoun, which replaces the subject of the text span (5.1), "*Mr. Cathcart*" (condition 2). There is no evidence of an increasingly temporal sequence (condition 3), and also no signal of a CONTRAST relation (condition 4). Therefore, a LIST relation is recognized between text spans (5.1) and (5.2).

The cue word "*and*" is found in example (6):

(6)[But the Reagan administration thought otherwise,[6.1]] [***and*** so may the Bush administration.[6.2]]

"A*nd*" is considered as a cue word because it stands at the beginning of the clause (6.2) (cue 1). It can be used in a LIST relation, a SEQUENCE relation, or an ELABORATION relation. With the score of 0.3 for the cue word "*and*" in the LIST relation, the actual score of cue 1 = Score(cue 1)*Score("*and*") = $1*0.3 = 0.3 < \theta$. Also, another cue of the LIST relation is found between clause (6.1) and clause (6.2). The subjects of two text spans, "*the Reagan administration*" and "*the Bush administration*", are co-hyponyms. In addition, clause (6.2) has the structure "*so + auxiliary + sbj*". With the satisfaction of cue 1 and cue 5, the total score is:

Total score = Actual Score(cue 1) + Score(cue 5)
$= 0.3 + 0.8 = 1.1 > \theta$.

As in the previous example, the necessary conditions of the LIST relation are checked and then a LIST relation is recognized between clause (6.1) and clause (6.2).

4 Conclusion

In this paper, we have explored several variants of cue phrases, and exploring combining with other feasible cohesive devices to recognise relations between two text spans. It was shown that NP cues, and VP cues are good predictors for discovering rhetorical relations. In the case where cue phrases are not available, other text cohesive devices (e.g., synonyms, and antonyms) can be a reasonable substitution.

The algorithm for recognising relations between two text spans is being implemented. The evaluation will be done by using documents from the RST Discourse Treebank after the completion of the implementation. Future work will focus on improving this algorithm's performance by refining the conditions to recognise relations mentioned in Section 3.1.

References

Carlson, L. and Marcu, D. (2001) *Discourse Tagging Manual*. ISI Tech Report ISI-TR-545.

Corston-Oliver, S. (1998). *Computing Representations of the Structure of Written Discourse*. PhD Thesis, University of California, U.S.A.

Hovy, E. H. (1990) *Parsimonious and profligate approaches to the question of discourse structure relations.* Proceedings of the 5th International Workshop on Natural Language Generation, Pittsburgh, 128-136.

Knott, A., Dale, R. (1995) *Using linguistic phenomena to motivate a set of coherence relations.* Discourse Processes 18:35-62.

Mann, W. C. and Thompson, S. A. (1988) *Rhetorical Structure Theory: Toward a Functional Theory of Text Organization.* Text, vol. 8(3), 243-281.

Marcu, D. (1997) *The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts.* PhD Thesis, Department of Computer Science, University of Toronto.

RST Discourse Treebank - http://www.ldc.upenn.edu/Catalog/LDC2002T07.html

Salkie, R. (1995) *Text and discourse analysis.* London: Routledge.

Webber, B. et al. (2001) *D-LTAG System - Discourse Parsing with a Lexicalized Tree Adjoining Grammar*. In ESSLLI 2001 Workshop on Information structure, Discourse structure and Discourse Semantics.

WordNet (2002) - http://www.cogsci.princeton.edu/~wn/index.shtml

[6] The superscripts such as 5.1 and 5.2 are used to distinguish different discourse units focussed on in each example.