Restricting the rhetorical input for the non-hierarchical planning of document structures

Nadjet Bouayad-Agha
Technology Department
University Pompeu Fabra
nadjet.bouayad@tecn.upf.es

Abstract

Whereas *rhetorical (or discourse) structuring* is the process of constructing rhetorical (or discourse) structures (RSs), *document structuring* is introduced by Power et al (in press) as the process of building *document structures* (DSs) from a discourse representation, where a DS represents the division of the text into abstract textual units such as paragraphs or orthographic sentences. In the context of the non-hierarchical planning of DSs, the problem that is addressed in the present paper is how to ensure that the DSs produced express all the information (i.e., rhetorical assertions) that has to be conveyed.

1 Introduction

Dynamic approaches to discourse planning have used Rhetorical Structure Theory or RST (Mann and Thompson, 1987) successfully as the basis for their plan representation. In hierarchical planning (Moore and Paris, 1994), RST relations have been translated into plan operators whose definition specifies which relations or propositions can realise their arguments. In non-hierarchical planning (Marcu, 1997a; Mellish et al., 1998a), a valid text plan is built from a set of rhetorical assertions using a general principle of strong compositionality, also called the nuclearity principle, and defined as follows: "a relation R holds between two spans of a text plan if it holds between the most important units (i.e., the nuclei) of the constituent spans". The best RS tree can be constructed using additional constraints based on ordering, adjacency, size of substructures, etc.

Whereas *rhetorical (or discourse) structuring* is the process of constructing rhetorical (or discourse) structures (RSs), *document structuring* is introduced by Power et al (in press) as the process of building *document structures* (DSs) from a discourse representation. A DS represents the division of the text into abstract textual units such as paragraphs or orthographic sentences. This representation is independent from discourse/rhetorical theories. For example, figure 1 illustrates the RS and DS of the text [1] below. The RS's leaves are the four clauses A to D cascaded into three elaboration relations. The DS's leaves are four textphrases grouped into two text (i.e.,orthographic) sentences within the same paragraph. This example shows that the RS and DS need not be homomorphic structures. Indeed, the grouping of C and D in the RS is lost in the DS. This non homomorphism is due to the linearisation constraints (i.e., how to divide the message into syntactic and document units) which are applied to the message when generating a text, and is especially likely as the message gets bigger. Without a distinction between RS and DS the version [1] of the message could not be generated.

[1] Ciproxin may cause a problem with your kidneys^{A} called crystalluria^{B} which results in tiny crystals forming in the urine^{C}. These crystals cannot be seen by the naked eye.^{D} *Ciproxin tablets 250mg notice*

Power et al (in press) show how DSs with some

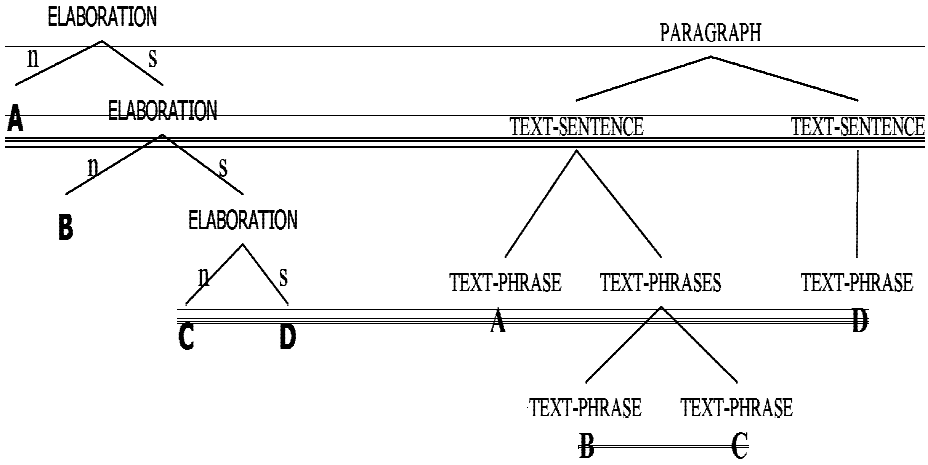

Figure 1: RS (left) and DS (right) for text [1]

degree of homomorphism with their input RS tree can be built systematically. Bouayad-Agha (2001) abolishes homomorphism altogether and presents an approach in which non-hierarchical document structuring is performed. The planning starts from a set of rhetorical assertions similar to Marcu(1997a) and builds valid DSs using a set of locally applied rhetorically motivated constraints on the DS constituents. These constraints are of two types: hard and soft. Constraints concerning the syntactic relations between the constituents are hard and thus *cannot* be violated. For example, when two rhetorically related propositions are in adjacent document constituents, say in two text-sentences, the linearly second of the two clauses must be a main clause (i.e., non subordinate). This is illustrated in the example below, where [2a] violates this constraint (i.e., B is not a main clause) and conveys the wrong message, whilst [2b] satisfies the constraint.

```
concession(sat:C,nuc:B)
explanation(sat:B,nuc:A)
```

[2*a*] #John is a good student^{C}. Nevertheless, because he failed his exam^{B}, he looks very upset^{A}.

[2*b*] John is a good student^{C}. Nevertheless, he failed his exam^{B}. As a result, he looks very upset^{A}.

On the other hand, constraints concerning the non-syntactic structural relations between the constituents of the DS are soft and *can* be violated. The main constraint of this kind is the one that requires that the groupings of non-syntactic DS constituents reflect rhetorical groupings. Its violation is not fatal as illustrated by example [1] above. Nevertheless, as the size of the message increases and the number of soft violations increases too, the resulting text may be infelicitous.

The DSs thus produced can be felicitous in the sense that the original message can be recovered. However, they are not necessarily homomorphic with their corresponding input RS. What has been built is not a rhetorical tree but a tree that respects the document grammar (i.e., grammar for deriving document structures) and the lexical realisation of the relations.

The problem that is addressed in the present paper is how to ensure that the document structures produced express all the information (i.e., rhetorical assertions) that has to be conveyed. No assumption is made about the nature of the rhetorical input. Since we do not impose a homomorphism between the document and the rhetorical structures, we cannot use the nuclearity principle for selecting our assertions for building valid DSs. Nevertheless, we show that the rhetorical input can still be restricted following some basic principles (section 2). The document structurer implicitly enforces those constraints whilst permitting the construction of non-homomorphic DSs, thus allowing a certain freedom of the input with respect to RST (section 3). This confirms that document structure is a useful distinction from discourse representation (section 4).

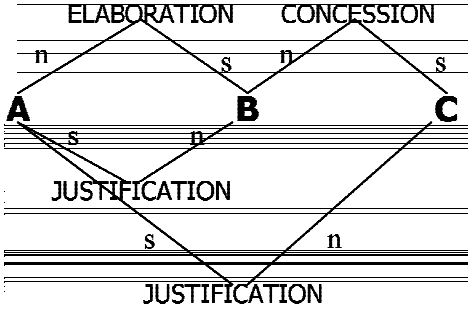

Figure 2: Graph corresponding to [1]

[1c] With the next music day scheduled for July 21 (Saturday), noon-midnight^A – I'll post more details later^B – this is a good time to reserve the place on your calendar^C.

[1d] With the next music day scheduled for July 21 (Saturday), noon-midnight^A, although I'll post more details later^B, this is a good time to reserve the place on your calendar^C.

[1e] With the next music day scheduled for July 21 (Saturday), noon-midnight^A:
- This is a good time to reserve the place on your calendar^C.
- I'll post more details later^B.

2 A rhetorical graph structure

The type of input we start from is a set of rhetorical assertions, a subset of which has to be selected to construct a valid text. This type of input has been used in NLG projects such as ILEX (Mellish et al., 1998b) and could be derived from a domain's knowledge base or from the input plan operators. The following input [1] is an illustrative example. It was introduced by Mann and Thompson (1987) and analysed in terms of rhetorical assertions following RST relations by Marcu (1997b). It is given together with output texts, each of which illustrating the realization of a possible structural configuration:

- [1a] is the original text and expresses R3 and R2;
- [1b] expresses R4 (the marker *with* is taken to give an interpretation of justification) and R2;
- [1c] expresses R1 and R3;
- [1d] expresses R1 and R2;
- [1e] expresses R1 and R4.

[1]
```
R1: justification(sat:A,nuc:C)
R2: concession(sat:B,nuc:C)
R3: elaboration(sat:B,nuc:A)
R4: justification(sat:A,nuc:B)
```

[1a] The next music day is scheduled for July 21 (Saturday), noon-midnight^A. I'll post more details later^B but this is a good time to reserve the place on your calendar^C.

[1b] With the next music day scheduled for July 21 (Saturday), noon-midnight^A, I'll post more details later^B. However, this is a good time to reserve the place on your calendar^C.

Apart from [1e] (which may have been obtained following some rhetorical aggregation), all the texts obtained correspond to a valid RS. Thus, during rhetorical structuring, the appropriate subset of assertions is selected that forms a valid RS. Each text expresses all the propositions in the input. Note however that a constraint that all the propositions in the input are expressed in the output is not sufficient. For example, given an input with four propositions A, B, C, D, we may select the two assertions R1(sat:A,nuc:B) and R2(sat:C,nuc:D) which realise all the propositions but are not connected in any way.

In our approach, the subset of assertions expressed in the final document structure respects a couple of basic principles. We represent the initial set of rhetorical assertions as a graph as in figure 2 with the following properties:

Labeled. The relations and propositions are represented by nodes linked by edges labelled either as satellite or nucleus. This allows a relation to be the argument of another relation.

Connected. The input must be built into a single connected graph. This is a simplifying assumption for the document structurer as potentially, it would be possible to realise two unconnected graphs in two separate paragraphs or sections.

A valid subset of assertions corresponds to a connected subgraph with no rhetorical circuits, where a rhetorical circuit is a closed path of successive rhetorical assertions. Given these restrictions, there are five subgraphs (figure 3) that can

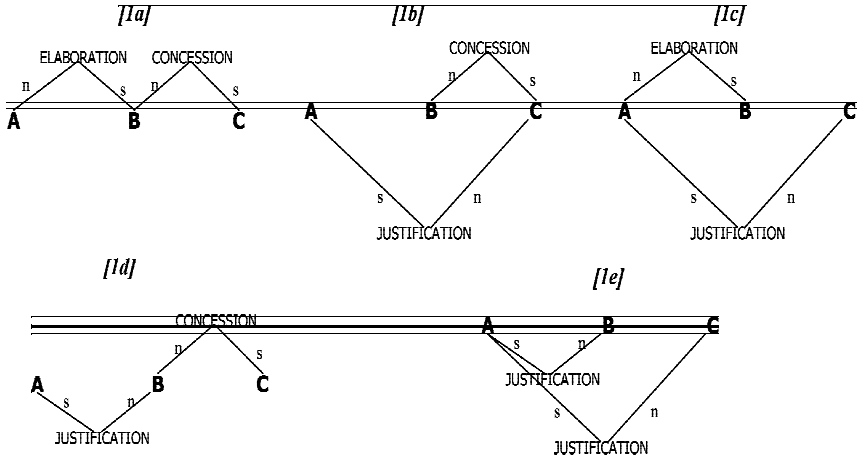

Figure 3: Five graphs derived from graph in figure 2, with no rhetorical circuit

be extracted from the main graph (figure 2). These correspond to all the open hamiltonian paths that can be found in the main graph, which is NP-complete problem. Each subgraph can be used by the document structurer, which, given the appropriate lexical and document resources, will realise all the assertions in that subgraph and produce a valid text.

A document structurer could in principle, in addition to consider non-connected graphs, consider all the possible subgraphs of the graph in figure 2. This would include subgraphs with (1) direct circuits, that is, multiple relations over the same pair of propositions as discussed in (Webber et al., 1999), and (2) indirect circuits, licensing texts like the one below, which expresses R1, R2 and R4:

[1f] With the next music day scheduled for July 21 (Saturday), noon-midnight^A:
- I'll post more details later^B.
- But this is a good time to reserve the place on your calendar^C.

3 Relation to RST

3.1 RST Properties

The properties of an RST structure defines it as a tree. They are (Mann and Thompson, 1987) :

Connectedness: except for the root, each text span in the analysis is either a minimal unit or a constituent of another schema application of the analysis.

Uniqueness: each schema application involves a different set of text spans.

Adjacency: the text spans of each schema application constitute one contiguous text spans.

Completeness: one schema application spans the entire text.

As described in the previous section, connectedness is enforced between the rhetorical assertions. In the current implementation, this happens because the non-hierarchical document structurer will only combine two (sub-)document structures together if each contains a proposition which is an argument of a rhetorical assertion to realise. Also, uniqueness is enforced because once a pair of sub-DSs has been used to realise an assertion, it cannot be re-used to realise another assertion. This enforces the no-circuit requirement described in the previous section. Adjacency is enforced by the document structurer which cannot interleave a DS inside another one. For example, given the following rhetorical input that satisfies our graph requirement:

```
R1(sat:A,nuc:B)
R2(sat:C,nuc:D)
R3(sat:E,nuc:F)
R4(sat:E,nuc:C)
R5(sat:F,nuc:B)
```

There are 6!=720 possible linear orderings for this rhetorical input, some of which are altogether

invalid since they can't be grouped into a single span. For example, ACBDFE is not possible because AB and CD are interleaved. On the other hand, ABFECD is possible because it can be linearly spanned into rhetorically related propositions: ((AB)((FE)(CD))).

Finally, completeness is enforced too given that the resulting DS is a tree. It should be obvious now that the DSs currently generated are RS-like structures. However and as explained in section 1, they need not be strongly compositional. This aspect allows a certain freedom of the input with respect to RST.

3.2 non-RST Properties

The rhetorical input of the following texts respects our graph requirement and can be built into a valid DS.[1] However, their RST analysis poses a problem.

[2] You should continue taking the tablets for as long as your doctor has asked^{A}, *unless* you develop any problems^{B}, *in which case*, consult your doctor^{C}.

condition(sat:B,nuc:C)
unless(sat:B,nuc:A)

[3] Avoid driving^{A} because dizziness can occur with the first dose^{B}. You may also feel dizzy if you stand up quickly after lying down^{C}.

cause(sat:B,nuc:A)
list(nuc1:B,nuc2:C)

The rhetorical assertions for [2] cannot be built into a tree because RST, given the uniqueness principle, does not allow a span to be satellite of two spans. However, a DS can be built given this input since the DS's constraint on uniqueness applies on DS spans, not rhetorical spans.

The rhetorical assertions for [3] can be built into an RS-tree but that RS-tree does not satisfy the nuclearity principle. On the other hand, a DS can be built given this input since we do not rely on this principle for DS composition. Other examples might involve relations whose arguments are relations, or relations with a wide scope (i.e., spanning more than one proposition, such as a condition or a circumstance relation) which cannot be represented in a strongly compositional RS-tree but can be represented by a valid rhetorical graph and could be made into a text via a DS.

[1] Both examples are taken from ABPI (1997). [2] is taken from Innozide (Merck-Sharp Dohme) whilst [3] is a simplified excerpt from Hytrin BPH (Abbott).

4 Conclusion

The absence of a homomorphic requirement between the discourse structure and the document structure returns to the rhetorical input some of RST's original descriptive power it was deprived of for the purposes of NLG. In particular, the rhetorical input does not have to be built into a rhetorical structure following a strong compositionality principle, whose limitations to account for real texts has been pointed out. In particular, Knott et al (2001) have discussed the conflicts between the hierarchical discourse structure of a text and the focus relations that may exist between segments of the text.

By distinguishing between a rhetorical graph input structure and a document structure output, we avoid the difficulties associated with the definition of RST as a theory of text analysis, where the hierarchical structure of the surface text is intermingled with the non-hierarchical structure of the input message (Power et al., in press). In effect, this means that document structuring is not dependent on a particular discourse theory.

References

ABPI, editor. 1997. *Compendium of Patient Information Leaflets*. Association of British Pharmaceutical Industry.

N. Bouayad-Agha. 2001. *The Role of Document Structure in Text Generation*. Ph.D. thesis, Information Technology Research Institute (ITRI), University of Brighton. Also availaible as a Technical Report ITRI-01-24.

A. Knott, J. Oberlander, M. O'Donnell, and C. Mellish. 2001. Beyond elaboration: the interaction of relations and focus in coherent text. In T.Sanders, J.Schilperoord, and W.Spooren, editors, *Text Representation: Linguistic and Psycholinguistics Aspects*, pages 181–196. Benjamins, Amsterdam.

W.C. Mann and S.A. Thompson. 1987. Rhetorical structure theory: A theory of text organization. Technical Report ISI/RS-87-190, Information Sciences Institute, University of Southern California.

D. Marcu. 1997a. From local to global coherence: A bottom-up approach to text planning. In *The Proceedings of the Fourteenth National Conference on*

Artificial Intelligence, pages 629–635, Providence, Rhode Island, July. AAAI.

D. Marcu. 1997b. *The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts*. Ph.D. thesis, Department of Computer Science, University of Toronto. Available at: http://www.isi.edu/ marcu/papers.html.

C. Mellish, A. Knott, J. Oberlander, and M. O'Donnell. 1998a. Experiments using stochastic search for text planning. In *Proceedings of the 9th International Workshop on Natural Language Generation*, pages 98–107, Niagara-on-the-Lake, Ontario, Canada, 5-7 August.

C. Mellish, M. O'Donnell, J. Oberlander, and A. Knott. 1998b. An architecture for opportunistic text generation. In *Proceedings of the 9th International Workshop on Natural Language Generation*, pages 28–37, Niagara-on-the-Lake, Ontario, Canada, 5-7 August.

J.D. Moore and C.L. Paris. 1994. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational linguistics*, 19(4):651–693.

R. Power, D. Scott, and N. Bouayad-Agha. (in press). Document structure. *Computational Linguistics*.

B.L Webber, A.Knott, and A.Joshi. 1999. Multiple discourse connectives in a lexicalized grammar for discourse. In *Third International Workshop on Computational Semantics*, Tilburg, The Netherlands.