Building Up Rhetorical Structure Trees

Daniel Marcu
Department of Computer Science
University of Toronto
Toronto, Ontario
Canada M5S 3G4
marcu@cs.toronto.edu

Abstract

I use the distinction between the nuclei and the satellites that pertain to discourse relations to introduce a compositionality criterion for discourse trees. I provide a first-order formalization of rhetorical structure trees and, on its basis, I derive an algorithm that constructs all the valid rhetorical trees that can be associated with a given discourse.

Motivation

Driven mostly by research in natural language generation, rhetorical structure theory (RST) (Mann & Thompson 1988) has become one of the most widely applied discourse theories. Despite its popularity, RST still lacks both a formal specification that would allow one to distinguish between well- and ill-formed rhetorical structure trees (RS-trees) and algorithms that would enable one to determine all the possible rhetorical analyses of a given discourse. For example, consider the following text (in which each textual unit[1] is labeled for reference):

(1) [No matter how much one wants to stay a non-smoker,^{A_1}] [the truth is that the pressure to smoke in junior high is greater than it will be any other time of one's life.^{B_1}] [We know that 3,000 teens start smoking each day,^{C_1}] [although it is a fact that 90% of them once thought that smoking was something that they'd never do.^{D_1}]

According to Mann and Thompson's definitions (1988), the rhetorical relations given in (2) below hold between the individual text units,[2] because the understanding of both A_1 and D_1 will increase the reader's readiness to accept the writer's right to present B_1; the understanding of C_1 will increase the reader's belief of B_1; the recognition of C_1 as something compatible with the situation presented in D_1 will increase the reader's positive regard for the situation presented in D_1; and the situation presented in D_1 is a restatement of the situation presented in A_1.

$$\textbf{(2)}\quad RR = \left\{ \begin{array}{l} \mathit{rhet\_rel}(\text{JUSTIFICATION}, A_1, B_1) \\ \mathit{rhet\_rel}(\text{JUSTIFICATION}, D_1, B_1) \\ \mathit{rhet\_rel}(\text{EVIDENCE}, C_1, B_1) \\ \mathit{rhet\_rel}(\text{CONCESSION}, C_1, D_1) \\ \mathit{rhet\_rel}(\text{RESTATEMENT}, D_1, A_1) \end{array} \right.$$

Assume now that one is given the task of building an RS-tree for text (1) and that one produces the candidates in figure 1.[3] Any student in RST would notice from the beginning that the tree in figure 1.d is illegal with respect to the requirements specified by Mann and Thompson (1988) because C_1 belongs to more than one text span, namely A_1–C_1 and C_1–D_1. However, even a specialist in RST will have trouble determining whether the trees in figure 1.a–c represent *all* the possible ways in which a rhetorical structure could be assigned to text (1), and moreover, in determining if these trees are *correct* with respect to the requirements of RST.

In this paper, I provide a formalization of the structure of RS-trees and show how one can use it to find answers to the questions given above. Section 2 reviews the elements of RST that are relevant for this paper, provides an explanation for the ambiguity of RS-trees, and proposes an informal mechanism that would enable one to alleviate the problems that are associated with this ambiguity. Section 3 creates the setting for the full formalization of RS-trees, which is presented in section 4. The last section is dedicated to an algorithmic perspective of the formalization and a discussion of its relevance to discourse processing.

RS-trees: informal intuitions

A critical analysis of RST

I believe that the explanation for the current lack of algorithms capable of automatically building the RS-trees that pertain to a given discourse can be found not only in the ambiguous definition of the rhetorical relations, but also in the incomplete description of RS-trees that is provided in the original theory. A careful analysis of the con-

[1] Throughout this paper, I use interchangeably the terms *textual unit* and *minimal unit* to refer to clauses.

[2] Throughout this paper, I use the convention that rhetorical relations are represented as sorted, first-order predicates having the form *rhet_rel*(*name*, *satellite*, *nucleus*) where *name*, *satellite* and *nucleus* represent the name, satellite, and nucleus of a rhetorical relation, respectively. Multinuclear relations are represented as predicates having the form *rhet_rel*(*name*, $nucleus_1$, $nucleus_2$).

[3] Throughout this paper, I use the graphical representation for RS-trees that is described by Mann and Thompson (1988).

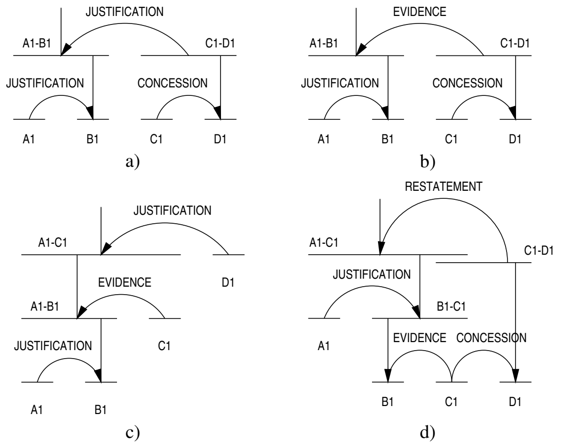

Figure 1: A set of possible rhetorical analyses of text (1).

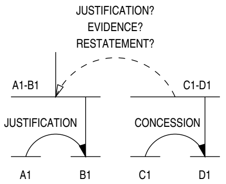

Figure 2: An example of the ambiguity that pertains to the construction of RS-trees.

straints provided by Mann and Thompson (1988, p. 248) shows that their specification for RS-trees is not complete with respect to some compositionality requirements, which would be necessary in order to formulate precisely the conditions that have to be satisfied if two adjacent spans are to be put together. Assume, for example, that an analyst is given text (1) and the set of rhetorical relations that pertain to the minimal units (2), and that that analyst takes the reasonable decision to build the spans A_1–B_1 and C_1–D_1, as shown in figure 2. To complete the construction of the RS-tree, the analyst will have to decide what the best relation is that could span over A_1–B_1 and C_1–D_1. If she considers the elementary relations (2) that hold across the two spans, she has three choices, which correspond to the relations *rhet_rel*(JUSTIFICATION, D_1, B_1), *rhet_rel*(EVIDENCE, C_1, B_1), and *rhet_rel*(RESTATEMENT, D_1, A_1). Which is the correct one to choose?

More generally, suppose that the analyst has already built two partial RS-trees on the top of two adjacent spans that consist of ten and twenty minimal units, respectively. Is it correct to join the two partial RS-trees in order to create a bigger tree just because there is a rhetorical relation that holds between two arbitrary minimal units that belong to those spans? A possible answer is to say that rhetorical relations are defined over spans that are larger than one unit too; therefore, in our case, it is correct to put the two partial RS-trees together if there is a rhetorical relation that holds between the two spans that we have considered. But if this is the case, how can one determine the precise boundaries of the spans over which that relation holds? And how do the rhetorical relations that hold between minimal units relate to the relations that hold between larger text spans? Mann and Thompson (1988) provide no precise answer for these questions.

Nuclearity and RS-trees

Despite the lack of a formal specification of the conditions that must hold in order to join two adjacent text spans, I believe that RST contains an implicit specification, which can be derived from Mann and Thompson's (1988) and Matthiessen and Thompson's (1988) discussion of nuclearity. During the development of RST, these researchers noticed that which is expressed by the nucleus of a rhetorical relation is more essential to the writer's purpose than the satellite; and that the satellite of a rhetorical relation is incomprehensible independent of the nucleus, but not vice-versa. Consequently, deleting the nuclei of the rhetorical relations that hold among all textual units in a text yields an incomprehensible text, while deleting the satellites of the rhetorical relations that hold among all textual units in a text yields a text that is still comprehensible. In fact, as Matthiessen and Thompson put it, "the nucleus-satellite relations are pervasive in texts independently of the grammar of clause combining" (1988, p. 290).

A careful analysis of the RS-trees that Mann, Thompson, and many others built shows that whenever two large text spans are connected through a rhetorical relation, that rhetorical relation holds also between the most important parts of the constituent spans. For example, in figure 1.a, the justification relation that holds between text spans C_1–D_1 and A_1–B_1 holds between their most salient parts as well, i.e., between the nuclei D_1 and B_1.

I propose that this observation can constitute the foundation for a formal treatment of compositionality in RST. More specifically, I will formalize the idea that two adjacent spans can be joined in a larger span by a given rhetorical relation if and only if that relation holds also between the most salient units of those spans. Obviously, such a formalization will also specify the rules for determining the most salient units of the spans.

A precise formulation of the RST problem

Formally, the problem that I want to solve is the following: given a sequence of textual units $U = u_1, u_2, \ldots, u_N$ and a set *RR* of rhetorical relations that hold among these units, find all legal discourse structures (trees) that could be built on the top of the linear sequence $u_1, u_2, \ldots, u_N$. Throughout this paper, I use the predicates $position(u_i, i)$ and *rhet_rel*(*name*, *satellite*, *nucleus*) with the following semantics: predicate $position(u_i, i)$ is true for a textual unit u_i in sequence U if and only if u_i is the i-th element in the sequence; predicate *rhet_rel*($name, u_i, u_j$) is true for textual units u_i and u_j with respect to rhetorical relation *name*, if and only if the definition provided by Mann and Thompson (1988) for rhetorical relation *name* applies for textual

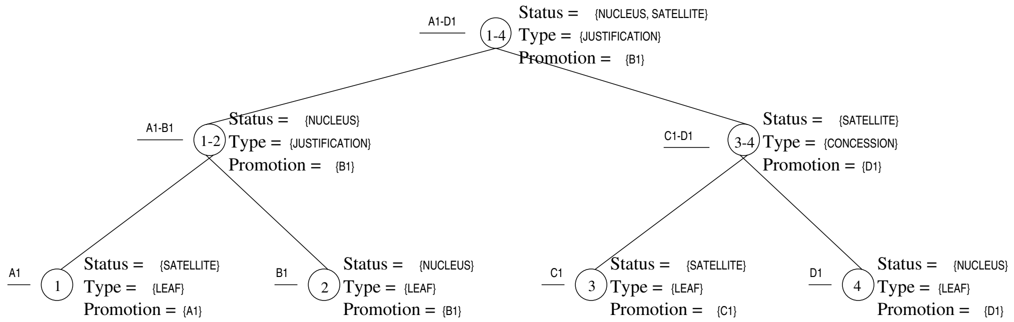

Figure 3: An isomorphic representation of tree in figure 1.a according to the status, type, and promotion features that characterize every node. The numbers associated with each node denote the limits of the text span that that node characterizes. The horizontal segments that pertain to each node underline the limits of the span that that node spans over.

---

units u_i, in most cases a satellite, and u_j, a nucleus. For example, from a rhetorical perspective, text (1) is completely described at the minimal unit level by the relations given in (2) and the relations given below in (3).

(3) $$\left\{ \begin{array}{l} position(\mathrm{A}_1, 1), position(\mathrm{B}_1, 2), \\ \qquad position(\mathrm{C}_1, 3), position(\mathrm{D}_1, 4) \end{array} \right.$$

The formalization that I propose here is built on the following features:

- An RS-tree is a binary tree whose leaves denote elementary textual units.
- Each node has associated a *status* (nucleus or satellite), a *type* (the rhetorical relation that holds between the text spans that that node spans over), and a *salience* or *promotion set* (the set of units that constitute the most "important" part of the text that is spanned by that node). By convention, for each leaf node, the type is LEAF and the promotion set is the textual unit to which it corresponds.

A representation for the tree in figure 1.a, which reflects these characteristics, is given in figure 3. The status, type, and salience unit that are associated with each leaf follow directly from the convention that I have given above. The status and the type of each internal node is a one-to-one mapping of the status and rhetorical relation that are associated with each non-minimal text span from the original representation. The status of the root reflects the fact that text span A_1–D_1 could play either a NUCLEUS or a SATELLITE role in any larger span that contains it.

The most significant differences between the tree in figure 3 and the tree in figure 1.a pertain to the promotion sets that are associated with every internal node. Consider, for example, the JUSTIFICATION relation that holds between units A_1 and B_1: according to the discussion of nuclearity in section 2, the nucleus of the relation, i.e., unit B_1, is the one that expresses what is more essential to the writer's purpose than the satellite A_1. Therefore, it makes sense that if span A_1–B_1 is to be related through other rhetorical relations to another part of the text, then it should do so through its most important or most salient part, i.e., B_1. Similarly, the nucleus D_1 of the rhetorical relation CONCESSION that holds between units C_1 and D_1 is the most salient unit for text span C_1–D_1. The intuition that the tree in figure 3 captures is that spans A_1–B_1 and C_1–D_1 could be assembled in a larger span A_1–D_1, because there is some rhetorical relation, in this case JUSTIFICATION, that holds between their most salient parts, i.e., D_1 and B_1.

The status, type, and promotion set that are associated with each node in an RS-tree provide sufficient information for a full description of an instance of a discourse structure. Given the linear nature of text and the fact that one cannot predict in advance where the boundaries between various text spans will be drawn, I will provide a methodology that permits one to quantify over all possible ways in which a tree could be build on the top of a linear sequence of textual units. The solution that I propose relies on the same intuition that constitutes the foundation of chart parsing: just as a chart parser is capable of quantifying over all possible ways in which different words in a sentence could be clustered into higher-order grammatical units, so my formalization would be capable of quantifying over all the possible ways in which different text spans could be joined into larger spans.

Let $span_{i,j}$, or simply $[i,j]$, denote a text span that includes all the textual units between position i and j. Then, if we consider a sequence of textual units $u_1, u_2, \ldots, u_\mathrm{N}$, there are N ways in which spans of length one could be built, $span_{1,1}, span_{2,2}, \ldots, span_{\mathrm{N},\mathrm{N}}$; $\mathrm{N}-1$ ways in which spans of length two could be built, $span_{1,2}, span_{2,3}, \ldots, span_{\mathrm{N}-1,\mathrm{N}}$; $\mathrm{N}-2$ ways in which spans of length three could be built, $span_{1,3}, span_{2,4}, \ldots, span_{\mathrm{N}-2,\mathrm{N}}; \ldots$; and one way in which a span of length N could be built, $span_{1,\mathrm{N}}$. Since it is impossible to determine a priori the text spans that will be used to make up a RS-tree, I will associate with each text span that could possibly become part of an RS-tree a status, a type, and a promotion relation and let the constraints described by Mann and Thompson (1988, p. 248) and the nuclearity constraints that I have described in section 2 generate the correct RS-trees. In fact, my intent is to determine from the set of $\mathrm{N}(\mathrm{N}+1)/2$ $(= \mathrm{N} + (\mathrm{N}-1) + (\mathrm{N}-2) + \ldots + 1)$ potential text spans that pertain to a sequence of N textual units, the subset that adheres to the constraints that I have mentioned above. For example, for text 1, there are 10 (= 4+3+2+1) potential spans, i.e., $span_{1,1}, span_{2,2}, span_{3,3}, span_{4,4}, span_{1,2}, span_{2,3}, span_{3,4}, span_{1,3}, span_{2,4}$, and $span_{1,4}$, but only seven of them play an active role in the representation given in figure 3, i.e., $span_{1,1}, span_{2,2}, span_{3,3}, span_{4,4}, span_{1,2}, span_{3,4}$, and $span_{1,4}$.

In formalizing the constraints that pertain to an RS-tree, I assume that each possible text span, $span_{l,h}$,[4] which will or will not eventually become a node in the final discourse tree, is characterized by the following relations:

- $S(l,h,status)$ denotes the status of $span_{l,h}$, i.e., the text span that contains units l to h; *status* can take one of the values NUCLEUS, SATELLITE, or NONE according to the role played by that span in the final RS-tree. For example, for the RS-tree depicted in figure 3, some of the relations that hold are: $S(1,2,\text{NUCLEUS}), S(3,4,\text{SATELLITE}), S(1,3,\text{NONE})$.
- $T(l,h,relation\_name)$ denotes the name of the rhetorical relation that holds between the text spans that are immediate subordinates of $span_{l,h}$ in the RS-tree. If the text span is not used in the construction of the final RS-tree, the type assigned by convention is NONE. For example, for the RS-tree in figure 3, some of the relations that hold are: $T(1,1,\text{LEAF}), T(1,2,\text{JUSTIFICATION}), T(3,4,\text{CONCESSION}), T(1,3,\text{NONE})$.
- $P(l,h,unit\_name)$ denotes the set of units that are salient for $span_{l,h}$ and that can be used to connect this text span with adjacent text spans in the final RS-tree. If $span_{l,h}$ is not used in the final RS-tree, by convention, the set of salient units is NONE. For example, for the RS-tree in figure 3, some of the relations that hold are: $P(1,1,\text{A}_1), P(1,2,\text{B}_1), P(1,3,\text{NONE}), P(3,4,\text{D}_1)$.

A complete formalization of RS-trees

Using the ideas that I have discussed in the previous section, I present now a complete first-order formalization of RS-trees. In this formalization, I assume a universe that consists of the set of natural numbers from 1 to N, where N represents the number of textual units in the text that is considered; the set of names that were defined by Mann and Thompson for each rhetorical relation; the set of unit names that are associated with each textual unit; and four extra constants: NUCLEUS, SATELLITE, NONE, and LEAF. The only function symbols that operate over this domain are the traditional + and − functions that are associated with the set of natural numbers. The formalization uses the traditional predicate symbols that pertain to the set of natural numbers ($<, \leq, >, \geq, =, \neq$) and five other predicate symbols: S, T, and P to account for the status, type, and salient units that are associated with every text span; *rhet_rel* to account for the rhetorical relations that hold between different textual units; and *position* to account for the index of the textual units in the text that one considers.

[4] In what follows, l and h always denote the left and right boundaries of a text span.

Throughout the paper, I apply the convention that all unbound variables are universally quantified and that variables are represented in *lower-case letters* while constants in SMALL CAPITALS. I also make use of two extra relations (*relevant_rel* and *relevant_unit*), which I define here as follows: for every text span $span_{l,h}$, $relevant\_rel(l,h,name)$ (4) describes the set of rhetorical relations that are relevant to that text span, i.e., the set of rhetorical relations that span over text spans that have their boundaries within the interval $[l,h]$. For every text span $span_{l,h}$, $relevant\_unit(l,h,u)$ (5) describes the set of textual units that are relevant for that text span, i.e., the units whose positions in the initial sequence are numbers in the interval $[l,h]$:

(4)
$$\begin{aligned} & relevant\_rel(l,h,name) \equiv (\exists s,n,sp,np)[\\ & \quad position(s,sp) \wedge position(n,np) \wedge \\ & \quad (l \leq sp \leq h) \wedge (l \leq np \leq h) \wedge \\ & \quad rhet\_rel(name,s,n)] \end{aligned}$$

(5)
$$\begin{aligned} & relevant\_unit(l,h,u) \equiv \\ & \quad (\exists x)[position(u,x) \wedge (l \leq x \leq h)] \end{aligned}$$

For example, for text (1), which is described formally in (2) and (3), the following is the set of all *relevant_rel* and *relevant_unit* relations that hold with respect to text segment $[1,3]$: $\{relevant\_rel(1,3,\text{JUSTIFICATION}), relevant\_rel(1,3,\text{EVIDENCE}), relevant\_unit(1,3,\text{A}_1), relevant\_unit(1,3,\text{B}_1), relevant\_unit(1,3,\text{C}_1)\}$

The constraints that pertain to the structure of an RS-tree can be partitioned into constraints related to the range of objects over which each predicate ranges and constraints related to the structure of the tree. I describe each set of constraints in turn.

Constraints that concern the objects over which the predicates that describe every span $[l,h]$ of an RS-tree range

• For every span $[l,h]$, the set of objects over which predicate S ranges is the set $\{\text{NUCLEUS}, \text{SATELLITE}, \text{NONE}\}$. Since every textual unit has to be part of the final RS-tree, the elementary text spans, i.e., those spans for which $l = h$, constitute an exception to this rule, i.e., they could play only a NUCLEUS or SATELLITE role.

(6)
$$\begin{aligned} & [(1 \leq h \leq \text{N}) \wedge (1 \leq l \leq h)] \rightarrow \\ & \quad \{[l = h \rightarrow \\ & \qquad (S(l,h,\text{NUCLEUS}) \vee S(l,h,\text{SATELLITE}))] \wedge \\ & \quad [l \neq h \rightarrow \\ & \qquad (S(l,h,\text{NUCLEUS}) \vee S(l,h,\text{SATELLITE}) \vee \\ & \qquad S(l,h,\text{NONE}))]\} \end{aligned}$$

• The status of any text span is unique

(7)
$$\begin{aligned} & [(1 \leq h \leq \text{N}) \wedge (1 \leq l \leq h)] \rightarrow \\ & \quad [(S(l,h,status_1) \wedge S(l,h,status_2)) \rightarrow \\ & \qquad status_1 = status_2] \end{aligned}$$

• **For every span $[l, h]$, the set of objects over which predicate T ranges is the set of rhetorical relations that are relevant to that span.** By convention, the rhetorical relation associated with a leaf is LEAF.

$$\begin{array}{l}[(1 \le h \le \text{N}) \wedge (1 \le l \le h)] \rightarrow \\ \quad \{[l = h \rightarrow T(l, h, \text{LEAF})] \wedge \\ \quad [l \ne h \rightarrow (T(l, h, \text{NONE}) \vee \\ \qquad (T(l, h, name) \rightarrow \\ \qquad \quad relevant\_rel(l, h, name)))]\}\end{array} \tag{8}$$

• **At most one rhetorical relation can connect two adjacent text spans**

$$\begin{array}{l}[(1 \le h \le \text{N}) \wedge (1 \le l < h)] \rightarrow \\ \quad [(T(l, h, name_1) \wedge T(l, h, name_2)) \rightarrow \\ \qquad name_1 = name_2]\end{array} \tag{9}$$

• **For every span $[l, h]$, the set of objects over which predicate P ranges is the set of units that make up that span.**

$$\begin{array}{l}[(1 \le h \le \text{N}) \wedge (1 \le l \le h)] \rightarrow \\ \quad [P(l, h, \text{NONE}) \vee \\ \quad (P(l, h, u) \rightarrow relevant\_unit(l, h, u))]\end{array} \tag{10}$$

Constraints that concern the structure of the RS-trees

The following constraints are derived from Mann and Thompson's formulation of RS-trees and from the nuclearity constraints that I have described in section 2.

• **Text spans do not overlap**

$$\begin{array}{l}[(1 \le h_1 \le \text{N}) \wedge (1 \le l_1 \le h_1) \wedge (1 \le h_2 \le \text{N}) \wedge \\ \quad (1 \le l_2 \le h_2) \wedge (l_1 < l_2) \wedge \\ \quad (h_1 < h_2) \wedge (l_2 \le h_1)] \\ \quad \rightarrow [\neg S(l_1, h_1, \text{NONE}) \rightarrow S(l_2, h_2, \text{NONE})]\end{array} \tag{11}$$

• **A text span with status NONE does not participate in the tree at all**

$$\begin{array}{l}[(1 \le h \le \text{N}) \wedge (1 \le l < h)] \rightarrow \\ \quad [(S(l, h, \text{NONE}) \wedge P(l, h, \text{NONE}) \wedge \\ \qquad T(l, h, \text{NONE})) \\ \quad \vee (\neg S(l, h, \text{NONE}) \wedge \neg P(l, h, \text{NONE}) \wedge \\ \qquad \neg T(l, h, \text{NONE}))]\end{array} \tag{12}$$

• **There exists a text span, the root, that spans over the entire text**

$$\begin{array}{l}\neg S(1, \text{N}, \text{NONE}) \wedge \neg P(1, \text{N}, \text{NONE}) \wedge \\ \neg T(1, \text{N}, \text{NONE})\end{array} \tag{13}$$

• **The status, type, and promotion set that are associated with a text span reflect the structural and nuclearity constraints that were discussed in section 2**

$$\begin{array}{l}[(1 \le h \le \text{N}) \wedge (1 \le l < h) \wedge \neg S(l, h, \text{NONE})] \rightarrow \\ \quad (\exists name, split\_point, s, n)[(l \le split\_point \le h) \\ \qquad \wedge (Nucleus\_first(name, split\_point, s, n) \vee \\ \qquad \quad Satellite\_first(name, split\_point, s, n))]\end{array} \tag{14}$$

Formula (14) specifies that whenever a test span $[l, h]$ denotes an internal node ($l < h$) in the final RS-tree, i.e., its status is not *none*, the span $[l, h]$ is built on the top of two text spans that meet at index *split_point* and either the formula denoted by *Nucleus_first* or *Satellite_first* holds.

$$\begin{array}{l}Nucleus\_first(name, split\_point, s, n) \equiv \\ \quad rhet\_rel(name, s, n) \wedge T(l, h, name) \wedge \\ \quad position(s, sp) \wedge position(n, np) \wedge \\ \quad l \le np \le split\_point \wedge split\_point < sp \le h \wedge \\ \quad P(l, split\_point, n) \wedge P(split\_point + 1, h, s) \wedge \\ \quad \{(name = \text{CONTRAST} \vee name = \text{JOINT}) \rightarrow \\ \qquad S(l, split\_point, \text{NUCLEUS}) \wedge \\ \qquad S(split\_point + 1, h, \text{NUCLEUS}) \wedge \\ \qquad (\forall p)[P(l, h, p) \rightarrow \\ \qquad \quad (P(l, split\_point, p) \vee \\ \qquad \quad P(split\_point + 1, h, p))]\} \wedge \\ \quad \{name = \text{SEQUENCE} \rightarrow \\ \qquad S(l, split\_point, \text{NUCLEUS}) \wedge \\ \qquad S(split\_point + 1, h, \text{NUCLEUS}) \wedge \\ \qquad (\forall p)(P(l, h, p) \rightarrow P(l, split\_point, p))\} \wedge \\ \quad \{(name \ne \text{SEQUENCE} \wedge name \ne \text{CONTRAST} \wedge \\ \quad name \ne \text{JOINT}) \rightarrow \\ \qquad S(l, split\_point, \text{NUCLEUS}) \wedge \\ \qquad S(split\_point + 1, h, \text{SATELLITE}) \wedge \\ \qquad (\forall p)(P(l, h, p) \rightarrow P(l, split\_point, p))\}\end{array} \tag{15}$$

Formula (15) specifies that there is a rhetorical relation with name *name*, from a unit s (in most cases a satellite) that belongs to span $[split\_point + 1, h]$ to a unit n, the nucleus, that belongs to span $[l, split\_point]$; that unit n is salient with respect to text span $[l, split\_point]$ and unit s is salient with respect to text span $[split\_point + 1, h]$; and that the type of span $[l, h]$ is given by the name of the rhetorical relation. If the relation is multinuclear, i.e., CONTRAST or JOINT, the status of the immediate sub-spans is NUCLEUS and the set of salient units for text span $[l, h]$ consists of all the units that make up the set of salient units that are associated with the two sub-spans. If the relation is a SEQUENCE relation, both sub-spans have NUCLEUS status, but the salient units for text span $[l, h]$ are given only by the salient units that are associated with the last member in the sequence, which in this case is realized first. If the relation is not multinuclear, the status of text span $[l, split\_point]$ is NUCLEUS, the status of text span $[split\_point + 1, h]$ is SATELLITE and the set of salient units for text span $[l, h]$ are given by the salient units that are associated with the subordinate nucleus span.

The difference between the formalization of the multinuclear relation of SEQUENCE and the other multinuclear relations stems from the fact that, unlike JOINT or CONTRAST, SEQUENCE is not symmetric. Formula *Satellite_first*$(name, split\_point, s, n)$ is a mirror image of (15) and it describes the case when the satellite that pertains to rhetorical relation *rhet_rel*$(name, s, n)$ belongs to text span $[l, split\_point]$, i.e., when the satellite goes before the nucleus. Due to space constraints, I do not reproduce it here.

An algorithmic view of RS-trees

Given the mathematical foundations of RS-trees, i.e., formulas (4)–(14), finding the RS-trees for a discourse described along the lines given in (2) and (3), for example, amounts to finding a model for the first-order theory that consists of formulas (2) to (14).

There are a number of ways in which one can proceed with an implementation: for example, a straightforward choice

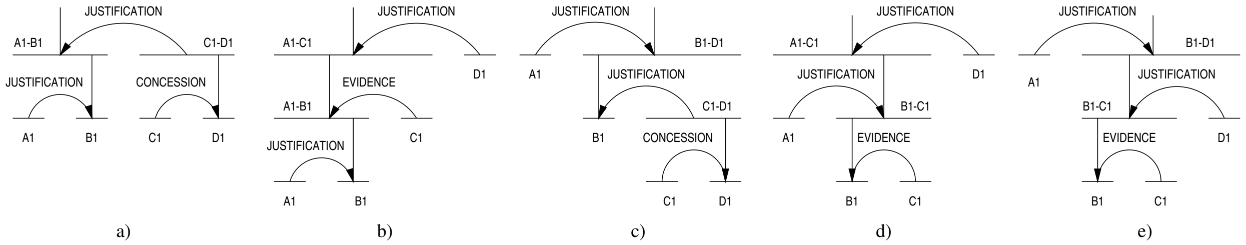

Figure 4: The set of all RS-trees that could be built for text (1).

is one that applies constraint-satisfaction techniques. Given a sequence U of N textual units, one can take advantage of the structure of the domain and associate with each of the $\mathrm{N}(\mathrm{N}+1)/2$ possible text spans a status, a type, and a salience or promotion variable whose domains consist in the set of objects over which the corresponding predicates S, T, and P range. This gives one a constraint-satisfaction problem with $3\mathrm{N}(\mathrm{N}+1)/2$ variables, whose domains are defined by formulas (6) to (10). The constraints associated with these variables are a one-to-one mapping of formulas (11) to (14). Finding the set of RS-trees that are associated with a given discourse reduces then to finding all the solutions for this constraint-satisfaction problem.

I have used Lisp and Screamer (Siskind & McAllester 1993), a macro package that provides constraint-satisfaction facilities, to fully implement a system that builds RS-trees. My program takes as input a linear sequence of textual units $U = u_1, u_2, \ldots, u_{\mathrm{N}}$ and the set of rhetorical relations that hold among these units. The algorithm builds automatically the corresponding constraint-satisfaction problem and then uses Screamer to find all the possible solutions for it. A simple procedure prints the RS-trees that pertain to each solution.

For example, for text (1), the program produces five RS-tree configurations (see figure 4). Among the set of trees in figure 4, trees 4.a and 4.b match the trees given in the introductory section in figure 1.a and 1.c. Trees 4.c–e represent trees that are not given in figure 1. Consequently, it follows that five RS-trees could be built on the top of text (1), and that tree 1.b is incorrect. It is easy to see that the reason that makes tree 1.b incorrect with respect to the formalization is that one of the constraints, i.e., the one that pertains to the rhetorical relation of *evidence* that is depicted between spans $[3,4]$ (C_1–D_1) and $[1,2]$ (A_1–B_1), does not hold. More precisely, the rhetorical relation of *concession* between C_1 and D_1 projects D_1 as the salient unit for text span $[3,4]$ (C_1–D_1). The initial set of rhetorical relations (2) depicts an evidence relation only between units C_1 and B_1 and not between D_1 and B_1. Since the nuclearity requirements make it impossible for C_1 to play both a satellite role in the span $[3,4]$ (C_1–D_1), and to be, at the same time, a salient unit for it, it follows that tree 1.b is incorrect.

The formalization and the algorithm that I presented here account for the construction of RS-trees in the cases in which the input specifies rhetorical relations between non-elementary spans as well. For example, if the input is enhanced such that besides the relations given in (2) it also contains the rhetorical relation *rhet_rel*(JUSTIFICATION, A_1, [B_1–D_1]), only the trees that are consistent with this extra constraint will be valid, i.e., trees 4.c and 4.e.

The formalization presented here distinguishes between correct and incorrect RS-trees only with respect to the original theory (Mann & Thompson 1988). Theme, focus, intention, or other pragmatic factors could rule out some of the trees that are produced by the algorithm; but a discussion of these issues is beyond the scope of this paper.

Conclusion

In this paper I provided a mathematical formulation of rhetorical structure trees that is based on the original Rhetorical Structure Theory (Mann & Thompson 1988) and the nuclearity features that pertain to natural language texts. On the basis of a first-order formulation of valid rhetorical structure trees, I implemented an algorithm that takes as input a sequence of textual units and a set of rhetorical relations that hold between those units, and that builds all the valid rhetorical structure trees that pertain to that sequence.

Acknowledgments. I am especially grateful to Graeme Hirst for long discussions and invaluable comments that helped me polish this work and to Jeff Siskind for bringing to my attention the similarity between charts and rhetorical structure trees, a similarity that catalyzed the emergence of the ideas presented in this paper. I am also grateful to Eduard Hovy, Ray Reiter, Manfred Stede, and Toby Donaldson for their comments on early drafts of the paper.

This reasearch was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Mann, W., and Thompson, S. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8 (3):243–281.

Matthiessen, C., and Thompson, S. 1988. The structure of discourse and 'subordination'. In Haiman, J., and Thompson, S., eds., *Clause combining in grammar and discourse*, volume 18 of *Typological Studies in Language*. John Benjamins Publishing Company. 275–329.

Siskind, J., and McAllester, D. 1993. Nondeterministic Lisp as a substrate for Constraint Logic Programming. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, AAAI–93, Seattle, 133–138.