LANGUAGE GENERATION 5
RST AND TEXT STRUCTURE PLANNING

Theme

RST, and the operationalization of RST into text plans. Its link with schemas. Intention, semantics, and text structure. The growth of text planning systems—from the RST Structurer to Moore and Paris to Cawsey, Dale, etc., to today. The need for additional planning—the tasks of microplanning.

Assignment: write a text planner using the sentence expansion realizer assigned for class.

1 Rhetorical Structure Theory (RST)

1.1 RST

RST is a (pre-)theory of text structure that turned out to have an important effect in computational work on text generation. After an analysis of hundreds of texts, Mann and Thompson (1988) produced a list of approx. 25 relations that, they claimed, were sufficient to describe the internal structure of all coherent English text. They defined coherence as follows:

> *A text is coherent when you can tell the function of each piece of a text with respect to all the others, using the functionals provided as RST relations.*

They represent discourse as a tree, using the rhetorical relations as labels for the branches. Incoherent texts are characterized by disjoint trees. The relations are divided into two classes—semantic and interpersonal. They include:

- **Semantic**: Cause, Elaboration (of various kinds), Sequence, etc.
- **Interpersonal**: Justify, Motivate, etc.

See the list of relations in the handout. Work through example 3.

Most of the relations are signaled by English words or phrases (Cause = "because" or "so", Sequence = "then", etc.); some are signaled otherwise (Elaboration uses a relative clause in English: "the man who was tall went away"). Most relations have two branches, a Nucleus (the principal branch) and a Satellite (the subsidiary one). Some, such as Sequence, are multi-nuclear, and have an arbitrary number of branches. Each branch is characterized by the constraints on the material that can fill it: a Sequence can only hold between an Event and another Event that follows it in time or space. The statements of the Nucleus and Satellite and joint N-S requirements are informal. A small exercise: Produce RST trees for the four texts in the accompanying handout.

You clearly see some shortcomings:

1. Inadequacy of relation definitions. From example 3, the fact that units 1 and 2 can be related by Elaborate-Attribute (when clause 1 is interpreted simply as a fact) or Motivate (when interpretation focuses on the meaning behind it, namely an invitation/request to attend the ballet). The fact that both readings are correct, and the distinction between Semantic and Interpersonal readings—see Section 5 below.

2. Incompleteness of relations: how many are there? More relations; see below.

3. Base level: why clauses? What about “the book that is blue is heavy” vs. “The blue book is heavy”? But if not clauses, what else?

1.2 Extended Sets of Discourse Relations

The frequent pairings of relations, and the ambiguity of cue words. So: more relations, of different kinds. Collection (Hovy and Maier, 1993, Hovy et al., 1992); taxonomy (Knott, 1994). Thee taxonomies show that you can develop new relations to an almost arbitrary degree of fineness, although 100 or so are probably more than enough for any computational system.

2 Formalizing RST Constraints using Cohen and Levesque’s Theory

In the previous lecture, we saw the value of schemas to help ensure coherent text structure. We also saw the value of Hobbs-like semantic connection graphs to represent the meaning of the discourse as it unfolds. Can one put the two together?

The answer is yes. One can decompose schemas into their component relations, for example using the RST relations, and then use the Nucleus and Satellite constraints, formalized, to capture (some of) the semantics Hobbs uses in his approach. Note that the RST constraints (N, S, and N+S) represent not only the facts of the world, but also the beliefs of the speaker and hearer.

The plan sounds appealing. Can one use this in a computer system? How can one operationalize RST relations and use them to produce coherent text?

Hovy (1987 and later) showed one way, using the BMB (mutual belief) notation of Cohen and Levesque to formulate relations’ constraints.

Cohen and Levesque (85) designed a theory of rational action and interaction expressed in a logic based on a possible-worlds semantics. They used the operators

(BEL x p) — P follows from X’s beliefs
(GOAL x p) — P follows from X’s goals
(BMB x y p) — P follows from X’s beliefs about what X and Y both believe
(AFTER a p) — P is true in all worlds after A occurs

They then derived some additional operators:

(KNOW x p) == p & (BEL x p)
(MUTUAL-KNOWLEDGE x y p) == p & (BMB x y p) & (BMB y x p)

and after some work derived some lemmas

(GOAL x p) & (BEL x (ALWAYS p→q)) → (GOAL x q)
(BMB y x (GOAL x p)) & (BMB y x (BEL x (ALWAYS p→q))) → (BMB y x (GOAL x q))

They then, after more work, defined the logical schema of imperatives (orders):

```
FORALL x y (MK x y (ATTEND y x)) -->
        (RESULT x (IMPER x y "y must do act")
                (BMB y x (GOAL x (BEL y (GOAL x (PRES-GOAL y (DONE y act)))))))
```

(when *x* tells *y* to do *act*, *y* believes that they both believe that *x* has the goal that *y* believes that *x* has the goal that *y* has the preservation goal (= continued goal) that *act* must be done by *y*.)

Finally, they produced 'summaries' of speech acts such as Request that capture the goals and beliefs of the participants. Here is a nonspecific request:

Initial conditions:

1. (BMB y x (SINCERE x (GOAL x (BEL y (GOAL x (PRES-GOAL y p)))))) &
2. (BMB y x (ALWAYS (COMPETENT y p))) &
3. (BMB y x (ALWAYS eventually-exists act (DONE y q act)
where q == (BEL y (RESULT y act p))))

Action:

(BMB y x (GOAL x (BEL y (GOAL x (PRES-GOAL y p)))))

Effect:

(BMB y x (GOAL x eventually p))

(when 1. *y* believes mutually that *x* is sincere in having the goal that *y* believe that *x* has the goal that *y* has the preservation goal to make *p* true, and 2. *y* also believes mutually that he/she is always competent to make *p* true after some act, and 3. *y* believes mutually that there always is some *act* that *y* believes will make *p* result and that in fact does so; then upon *y* believing mutually that *x* has the goal that *y* believe that *x* has the goal that y have the preservation goal to make *p* true; then *y* believes mutually that *x* has the goal that eventually *p* be true.)

These 'summaries' resemble the RST relations quite a lot, with the Initial Conditions being the N and S constraints and the Action being that which brings the N and S parts together—the joint N+S constraints. Once successfully done, the effect is the relation's effect on the mutual beliefs of speaker and hearer.

3 Multifunctionality of Discourse Structure Relations

An important debate in the early 1990s illustrated the differences between levels/types of interpretation of a discourse. Just as a sentence has syntactic, semantic, and informational structure, so a discourse (a set of sentences) can be analyzed at various levels.

Example:

"(1) Nothing can go faster than light, because (2) Einstein said so".

1. Surface analysis: (1) ←CAUSE— (2)
2. Semantics (physics): (1) —CAUSE→ (2)
3. Interpersonal: (BELIEF Speaker 1) ←CAUSE/JUSTIFY— (2)

After some initial confusion it is agreed by most researchers that it is useful to separate at least the following two levels:

- Informational/semantic: this deals with the information itself; its semantic structure
- Intentional/interpersonal: this deals with the speaker's goals, and desires to alter the belief states of the hearer

The papers by (Moore and Pollack, 1992) and following are important here.

4 The RST Structurer

4.1 Operationalizing Relations

With the RST Structurer (Hovy, 1988 and later), Hovy showed how RST relations, defined as above, could be linked to a knowledge base and operationalized as text plans and used in an AI NOAH-like expansion planner. Studying the paragraphs produced by Navy employees, he developed a set of plans that guaranteed the necessary structure.

RST relation/plan Sequence
Name: SEQUENCE
Goal/results:
((BMB SPEAKER HEARER (SEQUENCE-OF ?PART ?NEXT)))
Nucleus requirements/subgoals:
((BMB SPEAKER HEARER (TOPIC ?PART)))
Satellite requirements/subgoals:
((BMB SPEAKER HEARER (TOPIC ?NEXT)))
Nucleus+Satellite requirements/subgoals:
((NEXT-ACTION ?PART ?NEXT))
Nucleus growth points:
((BMB SPEAKER HEARER (CIRCUMSTANCE-OF ?PART ?CIR))
(BMB SPEAKER HEARER (ATTRIBUTE-OF ?PART ?VAL))
(BMB SPEAKER HEARER (PURPOSE-OF ?PART ?PURP)))
Satellite growth points:
((BMB SPEAKER HEARER (ATTRIBUTE-OF ?NEXT ?VAL))
(BMB SPEAKER HEARER (DETAILS-OF ?NEXT ?DETS))
(BMB SPEAKER HEARER (SEQUENCE-OF ?NEXT ?FOLL)))
Order: (NUCLEUS SATELLITE)
Relation-phrases: ("" "then" "next")
Activation-question:
"Could ?A be presented as start-point, mid-point, or end-point of some succession of items along some dimension? — that is, should the hearer know that ?A is part of a sequence?"

Note the important parts:
- the Nucleus contains some data item ?PART
- the Satellite contains some other data item ?NEXT
- the knowledge base must contain the relation (NEXT-ACTION ?PART ?NEXT) — that is, it must be true in the domain that ?NEXT follows ?PART
- the Order specifies which comes first in the text: Nucleus or Satellite
- the Relation-phrases provide some useful words that might be included in the text

If this is all true, the text planner can conjoin ?PART and ?NEXT in this order, and specify to the realizer that the phrase "then" or "next" be used.

But this is not all. It is possible to say more about ?PART and about ?NEXT, individually, before conjoining them. But one cannot simply say anything, in any order. So after ?PART, still in the Nucleus, one can search for: Circumstances of ?PART, Attributes of ?PART, or Purposes of ?PART, in that order. Similarly, one can search for additional aspects of ?NEXT, in the Satellite.

The **Growth Points** are what makes the relation act like a McKeown-style schema. That is, once you have decided on a

(generate-clause ?PART) (generate-word "then") (generate-clause ?NEXT)

you have on the goal stack the following goals:

(BMB SPEAKER HEARER (Circumstance ?PART ?CIR))
(BMB SPEAKER HEARER (Attribute-of ?PART ?VAL))
(BMB SPEAKER HEARER (Purpose-of ?PART ?PURP))
...

for each of which you have another RST relation plan.

These Growth Points are not just arbitrary—they are what ensures coherence, and must therefore be carefully assembled. Different domains, genres, and uses will require different relation/plans.

The initial goal starts the whole process, and will become the root node of the RST structure governing the paragraph. It could, for example, be

(BMB SPEAKER HEARER (Sequence-of Event1 Event2))

It is placed on the goal stack. When this goal matches the Effect field of the Sequence plan operator, the first node of the tree is established.

The **Order** and **Relation-Phrases** specify how the planned results (the leaves of the RST tree) go to the sentence realizer —what order, and what word(s) to use between them— so as to ensure that the eventual paragraph will be coherent.

Thinking back to the sentence realizer, you may wonder what all the fuss is about. Why could one not write the RST relation/plans in a simpler form? Indeed one can; much of the notation was inherited from the desire to conform to theory. Here's the relation/plan Sequence again:

Sequence[head] → [head] Circumstance[head] Attribute[head] Purpose[head] "then"
[head:next-action] Attribute[head:next-action]
Details[head:next-action] Sequence[head:next-action]

Where the [head] and [head:next-action] alone stand for the Nucleus and Satellite sentences. There's one principal difference though: generating a sentence with a grammar rule, you have to be sure to match all the constituents against the input frame; here the growth points are all optional.

4.2 The RST Text Structure Planner

These relations are then used in the RST Text Structure Planner, which performs cascaded expansion (using a simple version of the NOAH expansion algorithm from (Sacerdoti, 1977)) just like the sentence realizer algorithm.

Instead of grammar rules, you use text plans (with growth points)

Instead of a case frame for input, you use a communicative goal pointing into a knowledge base of facts

Instead of lexical items, you bottom out in single-clause calls to the realizer

Algorithm:

Input: communicative goal and a Knowledge Base (KB) of facts (your sentence representations, all linked).

Each goal is a pair [RST-effect, KB-item-frame]

place goal on stack
until stack empty, loop:
pop next goal from stack
if it is a Realizer call, add discourse phrase and place on Realizer Queue
else find appropriate relation/plan(s) by matching the goal's Effect to Effect fields
match N and S requirements against the KB
select a relation/plan and instantiate it
extend the discourse tree (take care of N and S order)
load instantiated N and S growth points onto the stack, in order
(optional) go through Realizer Queue and do cleanup (= early microplanning)
send each item on the Realizer Queue to the Realization Engine

This algorithm uses a Goal Stack and a Realizer Queue, to hold the outputs of the text planner (its leaves), which are then the input to the Realizer.

See the handout for a detailed example.

5 Text Planners Galore

Almost immediately afterward several extensions of this approach appeared: (Moore and Paris, 1989; also at ISI; Maybury, 1990; Cawsey, 1991; Suthers, 1993). They focused on some core problems: the inadequacy of RST relation/plans for selecting what to include, and the strangeness of forming communicative goals from many of them (you may have a goal to Justify, but surely not a goal to "Make a Sequence".

(Moore and Paris, 1991): Communicative goals, text plans (with examples), and use in EES.

(Maybury, 1990): illocution and perlocution.

(Dale, 1990): referring expressions.

See the handout for examples.

6 Problems

- Inadequacy of formulation of communicative goals. What about interpersonal goals, style, etc.?
- Inability to plan longer-than-paragraph texts in practice. Hybrids of schemas and text plans.
- Generation gap between text plan and sentence generator input: need for Sentence Planning.
- One-way text planners are not yet conversational systems!

7 Optional Background Readings

RST formalization and use:

Cohen, P.R. and Levesque, H.J. 1985. Speech Acts and Rationality. *Proceedings of the 23rd ACL Conference*, Chicago (49–59).

Moore, J.D. and Pollack, M.E. 1993. A Problem for RST: The Need for Multi-Level Discourse Analysis. Squib in *Computational Linguistics* 18(4).

Other text planners:

Dale, R. 1988. Generating Referring Expressions in a Domain of Objects and Processes. Ph.D. dissertation, University of Edinburgh.

Hovy, E.H. 1988. Planning Coherent Multisentential Text. *Proceedings of the 26th ACL Conference*, Buffalo (163–169).

Maybury, M.T. 1990. Planning Multisentential English Text Using Communicative Acts. Ph.D. dissertation, Cambridge University. Also available as RADC Technical Report 90-411.

Moore, J.D. 1989. A Reactive Approach to Explanation in Expert and Advice-Giving Systems. Ph.D. dissertation, University of California in Los Angeles.

Planner:

Sacerdoti, E. 1977. *A Structure for Plans and Behavior*. Amsterdam: North-Holland.

8 Assignment

Handed out in class.

1

Rhetorical Structure Theory: A Theory of Text Organization

William C. Mann
USC Information Sciences Institute

Sandra A. Thompson
Linguistics Department
UC Santa Barbara

Abstract

Rhetorical Structure Theory is a descriptive theory of a major aspect of the organization of natural text. It is a linguistically useful method for describing natural texts, characterizing their structure primarily in terms of relations that hold between parts of the text. This paper establishes a new definitional foundation for RST. Definitions are made more systematic and explicit, they introduce a new functional element, and incidentally reflect more experience in text analysis. Along with the definitions, the paper examines three claims and findings of RST: the predominance of nucleus/satellite structural patterns, the functional basis of hierarchy, and the communicative role of text structure.

Sample text from ISI email:

1. It's new brochure time,
2. and that means a chance for new project write-ups.

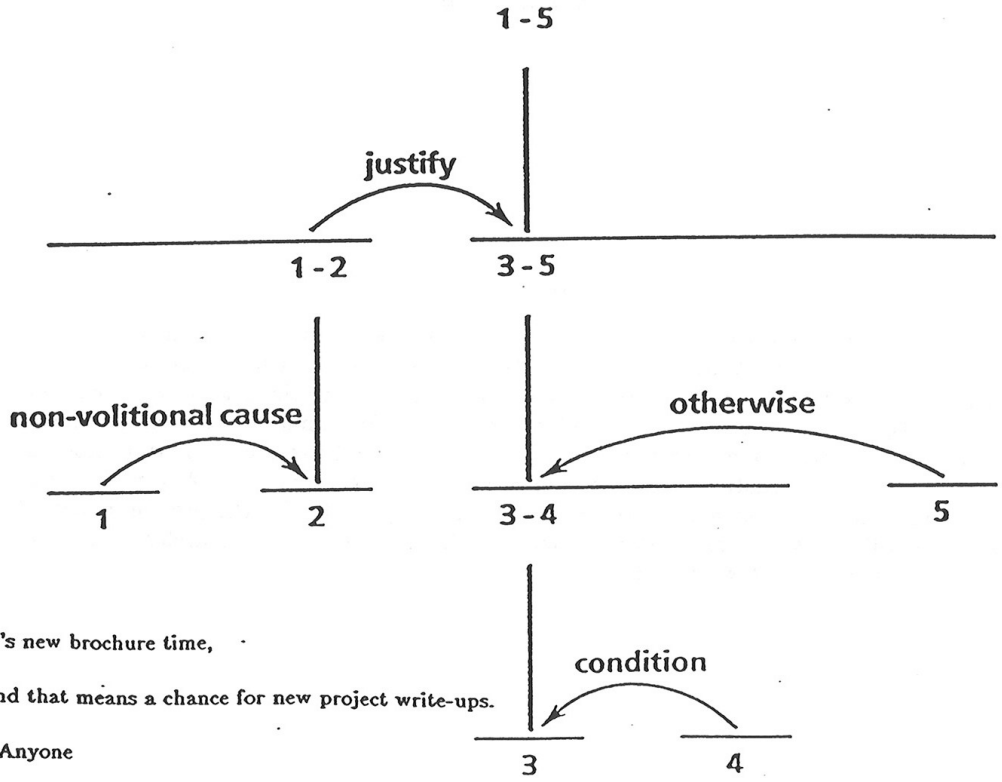
3a Anyone

4. desiring to update their entry in this brochure

3b. should have their copy in by Dec. 1.

5. Otherwise the existing entry will be used.

RHETORICAL STRUCTURE THEORY



1. It's new brochure time,
2. and that means a chance for new project write-ups.
3A. Anyone
4. desiring to update their entry in this brochure
3B. should have their copy in by Dec. 1.
5. Otherwise the existing entry will be used.

(Mann and Thompson 88)

Table 1: Organization of the Relation Definitions

Circumstance	Antithesis and Concession
Solutionhood	Antithesis
Elaboration	Concession
Background	Condition and Otherwise
Enablement and Motivation	Condition
Enablement	Otherwise
Motivation	Interpretation and Evaluation
Evidence and Justify	Interpretation
Evidence	Evaluation
Justify	Restatement and Summary
Relations of Cause	Restatement
Volitional Cause	Summary
Non-Volitional Cause	Other Relations
Volitional Result	Sequence
Non-Volitional Result	Contrast
Purpose	

relation name: EVIDENCE
constraints on N: R might not believe N to a degree satisfactory to W^6
constraints on S: The reader believes S or will find it credible.
constraints on the N + S combination:
R's comprehending S increases R's belief of N
the effect: R's belief of N is increased
locus of the effect: N

relation name: JUSTIFY
constraints on N: none
constraints on S: none
constraints on the N + S combination:
R's comprehending S increases R's readiness to accept W'S right to present N
the effect: R's readiness to accept W's right to present N is increased
locus of the effect: N

relation name: BACKGROUND
constraints on N: R won't comprehend N sufficiently before reading text of S
constraints on S: none
constraints on the N + S combination:
S increases the ability of R to comprehend an element in N
the effect: R's ability to comprehend N increases
locus of the effect: N

---

relation name: ANTITHESIS
constraints on N: W has positive regard for the situation presented in N
constraints on S: none
constraints on the N + S combination:
the situations presented in N and S are in contrast (cf. CONTRAST, i.e., are (a) comprehended as the same in many respects (b) comprehended as differing in a few respects and (c) are compared with respect to one or more of these differences); because of an incompatibility that arises from the contrast, one cannot have positive regard for both the situations presented in N and S; comprehending S and the incompatibility between the situations presented in N and S increases R's positive regard for the situation presented in N
the effect: R's positive regard for N is increased
locus of the effect: N

---

relation name: MOTIVATION
constraints on N: presents an action in which R is the actor (including accepting an offer), unrealized with respect to the context of N
constraints on S: none
constraints on the N + S combination:
Comprehending S increases R's desire to perform action presented in N
the effect: R's desire to perform action presented in N is increased
locus of the effect: N

---

relation name: ENABLEMENT
constraints on N: presents R action (including accepting an offer), unrealized with respect to the context of N
constraints on S: none
constraints on the N + S combination:
R comprehending S increases R's potential ability to perform the action presented in N
the effect: R's potential ability to perform the action presented in N increases
locus of the effect: N

---

relation name: CIRCUMSTANCE
constraints on N: none
constraints on S: S presents a situation (not unrealized)
constraints on the N + S combination:
S sets a framework in the subject matter within which R is intended to interpret the situation presented in N
the effect: R recognizes that the situation presented in S provides the framework for interpreting N
locus of the effect: N and S

---

relation name: SEQUENCE
constraints on N: multi-nuclear
constraints on the combination of nuclei:
A succession relationship between the situations is presented in the nuclei[23]
the effect: R recognizes the succession relationships among the nuclei.
locus of the effect: multiple nuclei

---

relation name: ELABORATION
constraints on N: none
constraints on S: none
constraints on the N + S combination:
S presents additional detail about the situation or some element of subject matter which is presented in N or inferentially accessible in N in one or more of the ways listed below. In the list, if N presents the first member of any pair, then S includes the second:[18]

1. set : member
2. abstract : instance
3. whole : part
4. process : step
5. object : attribute
6. generalization : specific

the effect: R recognizes the situation presented in S as providing additional detail for N. R identifies the element of subject matter for which detail is provided.
locus of the effect: N and S

---

relation name: CONDITION
constraints on N: none
constraints on S: S presents a hypothetical, future, or otherwise unrealized situation (relative to the situational context of S)
constraints on the N + S combination:
Realization of the situation presented in N depends on realization of that presented in S
the effect: R recognizes how the realization of the situation presented in N depends on the realization of the situation presented in S
locus of the effect: N and S

relation name: OTHERWISE
constraints on N: presents an unrealized situation
constraints on S: presents an unrealized situation
constraints on the N + S combination:
realization of the situation presented in N prevents realization of the situation presented in S
the effect: R recognizes the dependency relation of prevention between the realization of the situation presented in N and the realization of the situation presented in S
locus of the effect: N and S

relation name: NON-VOLITIONAL CAUSE
constraints on N: presents a situation that is not a volitional action
constraints on S: none
constraints on the N + S combination:
S presents a situation that, by means other than motivating a volitional action caused the situation presented in N; without the presentation of S, R might not know the particular cause of the situation; a presentation of N is more central than S to W's purposes in putting forth the N-S combination.
the effect: R recognizes the situation presented in S as a cause of the situation presented in N
locus of the effect: N and S

relation name: JUSTIFY
constraints on N: none
constraints on S: none
constraints on the N + S combination:
R's comprehending S increases R's readiness to accept W'S right to present N
the effect: R's readiness to accept W's right to present N is increased
locus of the effect: N

From Hovy et al 92: Employing knowledge resources

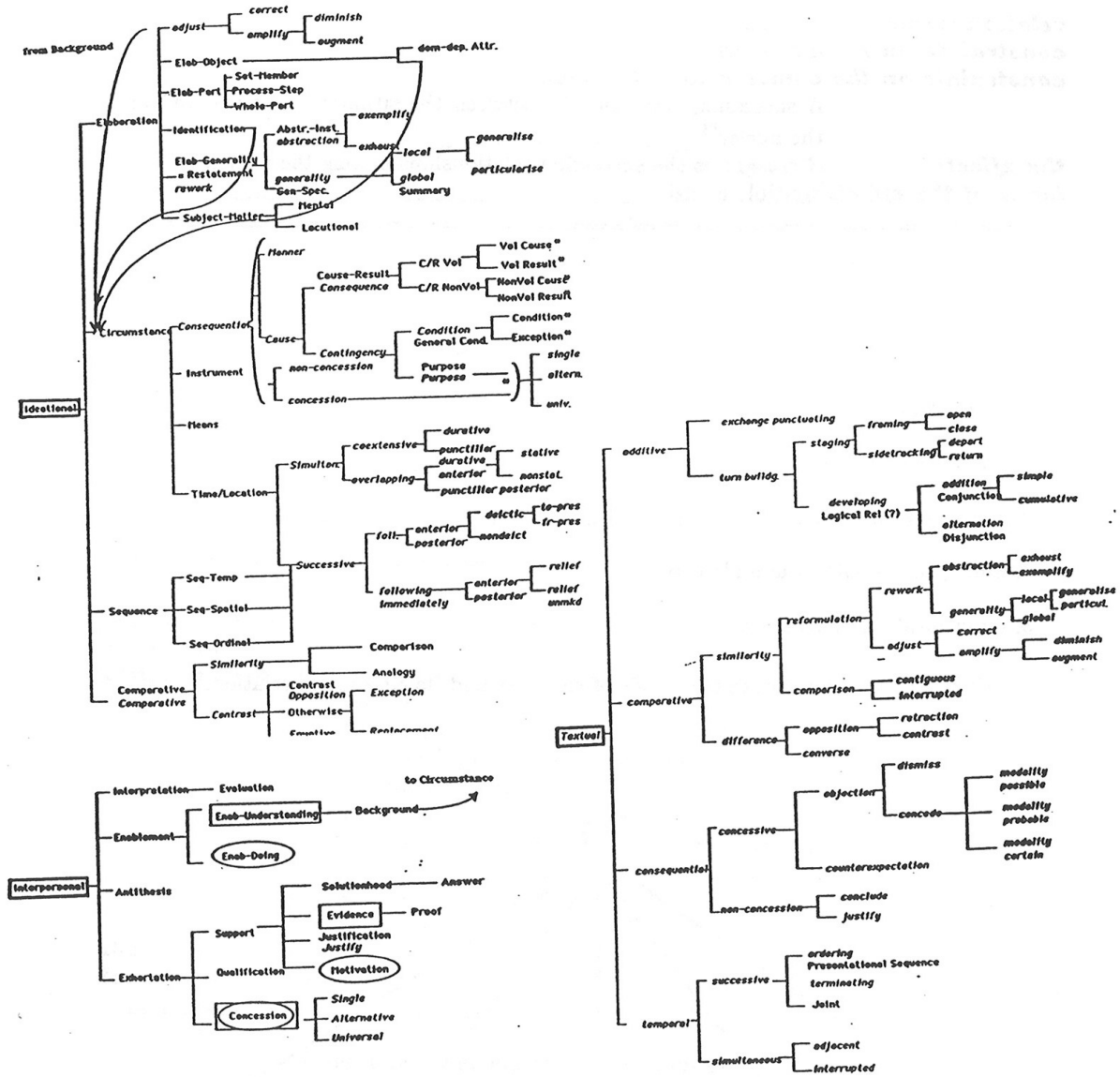


Figure 3: Discourse structure relation networks.

(1.)

1. The Los Angeles Chamber Ballet (the ballet company I'm dancing with) is giving 4 concerts next week ...

2. Tickets are $7.50 except for the opening night ...

3. The show is made up of new choreography and should be very entertaining.

4. I'm in 3 pieces.

(2.)

1. P. M. has been with KUSC longer than any other staff member.

2. While attending Occidental College,

3. where he majored in philosophy,

4. he volunteered to work at the station as a classical music announcer.

5. That was in 1970.

(3.)

1. It's new brochure time,

2. and that means a chance for new project write-ups.

3A. Anyone

4. desiring to update their entry in this brochure

3B. should have their copy in by Dec. 1.

5. Otherwise the existing entry will be used.

(4.)

1. Employees are urged to complete new beneficiary designation forms for retirement or life insurance benefits

2. whenever there is a change in marital or family status.

....

5A. Employees

6. who are not sure of who is listed as their beneficiary

5B. should complete new forms

7. since the retirement system and the insurance carrier use the most current form

8. to disburse benefits

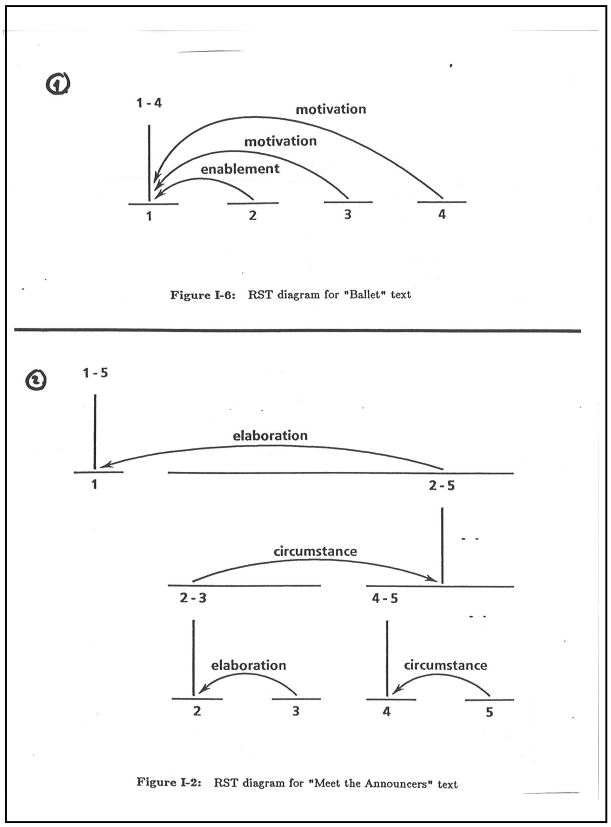


Figure I-6: RST diagram for "Ballet" text

Figure I-2: RST diagram for "Meet the Announcers" text