Rhetorical Relations in Instructional Text Generation

Keith Vander Linden
Department of Computer Science
University of Colorado
Boulder, CO 80309-0430
email: linden@cs.colorado.edu

May 31, 1993

The IMAGENE project has studied the expression of actions in the context of instructional text generation. The approach employs a rather traditional interpretation of Rhetorical Structure Theory (RST) (Mann and Thompson, 1989), using it both as a descriptive tool and as a constructive tool (Mann and Thompson, 1987). No explicit representation of and reasoning about intentions was employed. In this light, the project can serve as a data point in the broader discussion of the use of rhetorical relations and of their interaction with intentions.

This contribution will begin with a brief overview of the IMAGENE project[1], focusing on the nature of the use of RST in the project and, in particular, on the procedural basis of the inventory of rhetorical relations employed. It will conclude with a discussion of the precise characteristics of the project that appear to have warranted the lack of specific concern with intentions.

The Use of Relations

In the IMAGENE project, *instructional text* is taken to refer exclusively to written, procedural directions prescribing the performance some sequence of actions to the reader. This type of text can be seen as the expression of a set of actions bearing procedural relationships with one another. In this light, two tasks that an instructional text generator must perform are, first, to choose, for each action expression, the rhetorical relation it will hold with the other actions that best conveys their procedural relationships, and, secondly, to choose the precise grammatical form that will signal this rhetorical relation.

It is tempting to address these two tasks at an intuitive level, identifying both the rhetorical status and the grammatical form that appear to most effectively express various types of actions and their relations. The problem with this approach is that it is unclear how accurate our intuitions in this matter are. As an alternative, the IMAGENE project was based on a detailed function to form study of a corpus of instructional texts, currently made up of approximately 1000 clauses of instructional text (6000 words) taken from manuals and instructions concerning a variety of devices and processes, taken from 17 different sources[2]. This corpus is represented in a relational database representing the rhetorical and grammatical aspects of the text.

The corpus was analyzed and RST structures were built for all of the text. This analysis of rhetorical status made use of three nucleus-satellite relations: Purpose, Precondition, Result, and

---

[1]More detail can be found elsewhere (Vander Linden, 1993a; Vander Linden, 1993b).
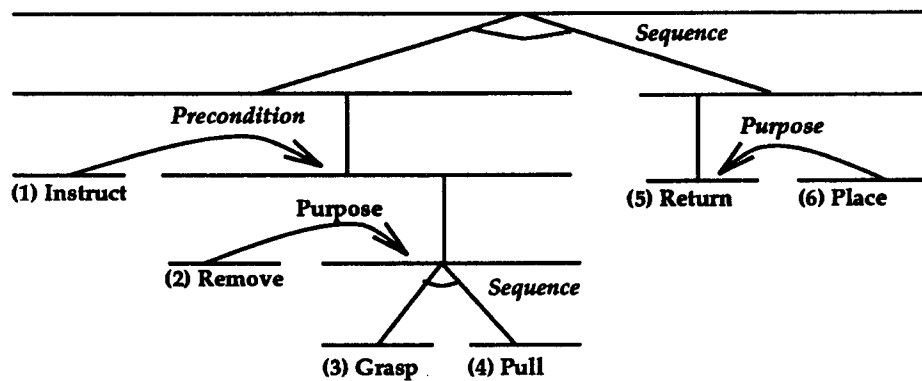[2]This is in interesting contrast to the form to function study performed by Knott and Dale (1992).
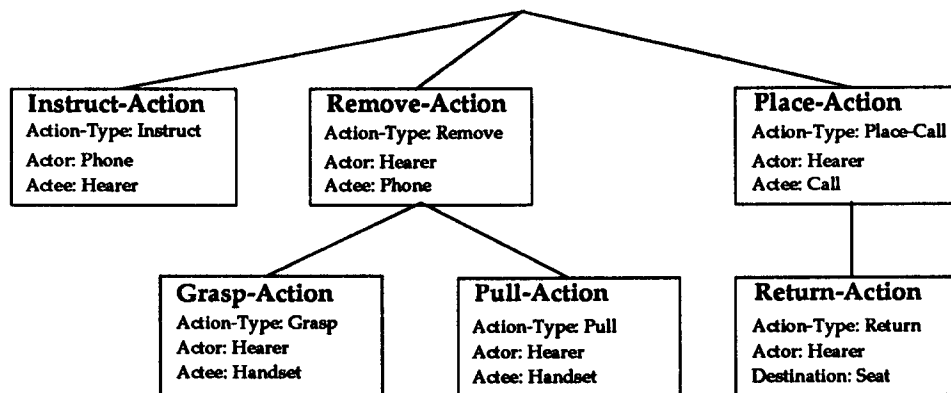
Figure 1: The RST Analysis for the Example.

Figure 2: The Procedural Relations for the Example.

two joint schemas: Sequence and Concurrent. The Purpose relation and the Sequence schema are taken directly from the RST specification. The Precondition and Result relations are simple amalgams of Circumstance and Condition, and Volitional and Non-Volitional Result respectively. The Concurrent schema is a simple extension of the Sequence schema.

As an example, consider the following passage from a telephone instruction manual:

[1] When instructed (approx. 10 sec.) [2] remove phone [3] by firmly grasping top of handset [4] and pulling out. [5] Return to seat [6] to place calls.

The RST analysis for this text is shown in Figure 1. The procedural relationships that lie behind the expressions in this passage are shown in Figure 2.

This set of relations and schemas, which has proven effective in analyzing instructional text, is based on the notions of hierarchical and non-linear plans and the use of preconditions and postconditions in automated planners[3]. During descriptive analysis, Purpose is identified with the expression of what is called the *generation* relationship (Goldman, 1970). Precondition and Result

---

[3]See Mellish and Evans (1988) for a discussion of the these issues are they relate to text generation.

141

are identified with the expression of actions as pre or post-conditions for other actions. Sequence schemas are identified with the expression of sequential actions and, similarly, Concurrent schemas with the expression of concurrency.

Given this coding of the rhetorical status of action expressions, coupled with the coding of the grammatical form of the expressions, a functional analysis was performed which identified systematic co-variation between functions and forms in the corpus. It turned out that a set of approximately 70 features of the communicative environment — in terms of Systemic-Functional Linguistics, elements of the ideational, textual, and interpersonal metafunctions (Halliday, 1985) — were sufficient to produce a broad analytical coverage of the rhetorical status and grammatical forms used in the corpus. These features were then coded in a single system network which formed the basis of the constructive use of RST within the IMAGENE instructional text generation system.

Where Intentions Fit In

Given that Moore and Pollack's (1992) examples of the need for simultaneous representation of relations and intentions are so compelling, there must some explanation of why the IMAGENE project was successful within the orthodox RST tradition. This concluding section will suggest two characteristics of the IMAGENE study that appear to have been instrumental in this regard and discuss their implications for the appropriate roles of relations and intentions.

The first characteristic of the IMAGENE project was its focus on local rhetorical relations in written instructional text in English. There are a number of sub-issues related to this focus of concern, all of which tend to lend themselves to a traditional RST approach:

Written rather than interactive discourse — A number of studies in the context of interactive discourse have emphasized the need for separate representation of intentions (Fox, 1988; Grosz and Sidner, 1986; Moore and Paris, 1989). This mechanism allows the system to deal with, for example, conversational repair, an issue which is not prevalent in written text.

Instructional text rather than other genres — Instructional text does not tend to make use of the deep and multi-faceted intentions that are common in argumentative and persuasive text (such as was the case in the "Come home by 5:00" example cited by Moore and Pollack). Instructional texts tend to be more straight forward expressions of actions and the procedural relations among them. In fact, the definition of the instructional genre itself makes reference to the single fundamental intention of expressing a procedure in an effective way (termed the "deep" intention by Delin et al. (1993)), an intention which has manifested itself in a number of standardized, domain-specific forms of expression commonly used by technical writers (termed *Domain Communication Knowledge* by Kittredge et al. (1991)).

Local rather than global relations — Not only did the IMAGENE project specifically address instructional text, but it has exclusively addressed the use of local rhetorical relations. Quite often the difficulty observed with RST analyses has been at higher levels.

The second, and perhaps most significant characteristic of the IMAGENE project is its focus on the problems of rhetorical status selection and grammatical form selection. No attempt was made to address the issue of content selection, indeed the corpus-based methodology employed would not provide a completely satisfying basis on which to address this issue. IMAGENE takes as input a process structure, such as the one shown in Figure 2, and does very little reasoning concerning

what to say (aside from pruning the process tree structure in some cases). This issue of content selection appears to be a crucial contribution of intentions.

References

Delin, J., Scott, D., and Hartley, T. (1993). Knowledge, intention, rhetoric: Levels of variation in multilingual instructions. In this volume.

Fox, B. A. (1988). Robust learning environments: The issue of canned text. Technical Report 88-5, Institute of Cognitive Science, The University of Colorado.

Goldman, A. I. (1970). *A Theory of Human Action*. Prentice Hall, Englewood Cliffs, NJ.

Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3).

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar*. Edward Arnold, London.

Kittredge, R., Korelsky, T., and Rambow, O. (1991). On the need for domain communication knowledge. *Computational Intelligence*, 7(4):305–314.

Knott, A. and Dale, R. (1992). Using linguistic phenomena to motivate a set of rhetorical relations. Technical Report HCRC/RP-39, Human Communication Research Centre, University of Edinburgh.

Mann, W. C. and Thompson, S. A. (1987). Rhetorical structure theory: Description and construction of text structures. In Kempen, G., editor, *Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics*. NATO Scientific Affairs Division, Martinus Nijhoff.

Mann, W. C. and Thompson, S. A. (1989). Rhetorical structure theory: A theory of text organization. In Polanyi, L., editor, *The Structure of Discourse*. Ablex, Norwood, NJ. To appear, currently available as ISI tech. report ISI/RS-87-190.

Mellish, C. (1988). Natural language generation from plans. In Zock, M. and Sabah, G., editors, *Advances in Natural Language Generation An Interdisciplinary Prespective*, volume 1, chapter 7. Ablex. Selected readings from the 1st European NLG Workshop, the Abbaye de Rouaumont, 1987.

Moore, J. D. and Paris, C. L. (1989). Planning text for advisory dialogues. In *Proceedings of the 27th Annual Meeting of Association for Computational Linguistics*, 26–29 June, Vancouver, B.C. Also available as ISI tech. report ISI/RR-89-236.

Moore, J. D. and Pollack, M. E. (1992). A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics*, 18(4):537–544. Squibs and Discussions.

Vander Linden, K. (1993a). Generating effective instructions. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*, June 18–21, Boulder, CO. To appear.

Vander Linden, K. (1993b). *Speaking of Actions: Choosing Rhetorical Status and Grammatical Form in Instructional Text Generation*. PhD thesis, University of Colorado.