

On the necessity of intentions and (at least) the usefulness of rhetorical relations: A position paper

Vibhu O. Mittal and Cécile L. Paris

Information Sciences Institute and Department of Computer Science
University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292
{mittal, paris}@isi.edu

In this position paper, we argue for the need for a generation system to represent communicative goals (i.e., goals that express the speaker's intent to affect the hearer's mental attitudes), at least if certain tasks are to be achieved. We further argue that, while rhetorical relations might be recoverable from other factors, they appear to be a useful level of abstraction to maintain about the discourse structure to avoid costly reasoning. However, we believe a source of confusion comes from the fact that the term 'rhetorical relations' encompasses various different constraints about a discourse structure, and that there is a need to be clear about what a 'rhetorical relation' is.

1 Need for communicative goals

Typically, a speaker has some goal in mind when producing an utterance. This can be a very concrete goal such as "describe an object", "motivate the user to perform an action", or a much vaguer one, such as discussing an event or simply providing information about a subject. We believe that communicative goals should be represented explicitly in a generation system for several reasons:

- To select the facts to be expressed, as typically, one does not utter everything one knows, and one's communicative intentions influences strongly what is to be said.
- To determine the ordering of facts to be expressed, in the event that these facts have been determined ahead of time. Indeed, typically, several orderings will be possible, but only a subset of these will be appropriate for the speaker's intention(s).
- To handle dialogue (feedback): it has already been shown that a record of the rhetorical relations holding in an utterance is not enough to be able to participate in a dialogue (specifically to recover from communication failures) [6]. This is because, as there is not always a one to one relationship between rhetorical relations and communicative goals, communicative goals cannot always be recovered from a record of the rhetorical relationships and the communicative goal is necessary to handle follow-up questions.
- To take into account any non-textual phenomena: for instance, there may be many ways to elaborate on a specific aspect in a presentation that includes both text and graphics: the specific manner of elaboration will depend upon the communicative goal to be achieved (e.g., providing information about the position of a knob in the control panel may result in a picture of the whole panel, while providing information in order to identify the knob may result in a descriptive phrase such as "the hexagonal knob").

We thus argue that communicative goals are important to represent in a generation system and are an essential part of discourse structure (which is not to say they are sufficient, however).

Some researchers have argued that communicative goals are not necessary for coherent text generation, and have cited architectures such as schemas and systems based on "domain knowledge" to illustrate their point. In our view, this simply illustrates how one can compile out the intentional knowledge necessary

to achieve coherent discourse within a specific domain and context, not that intentions are not an essential part of the discourse structure. It is important to realize that, while for some generation tasks, schemas that compile out intentional information are enough, systems based on such generation schemes are then unable to handle other discourse phenomena such as dialogue.

Finally, some researchers have conflated the notion of communicative goals with that of rhetorical relations, arguing that it is sufficient to reason about rhetorical relations that hold between certain facts to construct a coherent text, and further arguing that the choice of a rhetorical relation indicates a specific communicative goal. We do not agree with these arguments, however, at least not with respect to the current theories of rhetorical relations. Indeed, in general, there is not a one-to-one mapping between intentions and rhetorical relations [6]. Consider for example the following dialogue, taken from [6]:

S: (1) Remove the cover. (2) You'll need the Phillips screwdriver. (3) It's in the top drawer of the toolbox. Do you have it?

H: No.

Using an RST-type analysis [4], (3) is a CIRCUMSTANCE to (2). In terms of intentions, (2) is uttered in an attempt to make the hearer identify and find the screwdriver. Suppose the hearer is not able to find the screwdriver, because he or she cannot identify it. At this point, the speaker needs to recognize that the intention to make the hearer identify the screwdriver has not been achieved. However, the speaker cannot recover this intention simply from knowing that there is a CIRCUMSTANCE relation between (2) and (3). CIRCUMSTANCE could have been used to achieve a variety of other intentions. Similarly, the communicative goal could have been achieved using other rhetorical relations such as, for example, ELABORATE-OBJECT-ATTRIBUTE. (For more details, see [6]).

2 Rhetorical relations

Doing text analysis, researchers have found that coherent texts tend to exhibit patterns in their organization, with a limited set of relationships between their components. As a result, people have attempted to characterize these relationships in terms of "rhetorical relations" (or schemas of rhetorical predicates). Typically, the relationships identified above arise because of various factors, including:

- There is a *semantic* relation that holds between the various semantic units (or sets of semantic units taken as a whole) – for example, given the facts: <John has a black car> <John's car is a Honda>, since 'being a Honda' is an attribute of John's car, an item introduced in the first fact, we could say: "John has a black car. It is a Honda" and draw an "elaborate-object-attribute" between the two clauses.
- *Intentions* also sometimes have a *canonical clustering*: for example, the goal to provide evidence for a claim often appears when a claim is given. In the resulting text, the relation "evidence" can be drawn between the two spans of text that result from the two intentional goals.
- A recognized pattern of *thematic progression* is present in the text, giving rise to what has been called to a "textual relation" between units of text.

Essentially, these standard patterns (or standard sets of constraints) have been given names (the names of the "rhetorical relations", e.g., "elaboration", "evidence", or "additive"). These labels seem useful to perform certain types of reasoning on a text, such as the choice of syntactic realization (including deciding where a sentence break should occur, or the choice of a cue phrase), e.g., [6, 10], without reasoning explicitly about the actual constraints that hold, and which the speaker wanted to highlight. (Note that this is not to say that there is a clear unambiguous mapping between the rhetorical relation and the resulting syntactic realization).

For example, if a system knows that a unit of text U1 is in PURPOSE relation to another unit U2, then the realization component can select an appropriate syntax to verbalize U1 and U2 (e.g., “in order to...”), without re-reasoning about the constraints that hold between the information presented in U1 and U2.

Similarly, when several constraints hold and the speaker chooses to highlight one, this can be indicated in the discourse structure by the appropriate preferred label, which then constrains the realization component to choose the appropriate syntactic structure allowing the hearer to recognize the preferred interpretation. For example, consider the propositions “turn on the light” and “flip the switch” (taken from [9]). These propositions stand in several semantic relations with respect to each other: e.g., at least, CAUSE-EFFECT and PURPOSE. Given the speaker’s intention to make the hearer know how to turn the light, depending, for example, on the focus the speaker wishes to have, several texts can be generated, each would highlight a different relation, and each would be analyzed with a different RST-relation between the two clauses (e.g., “to turn on the light, flip the switch” or “flipping the switch will cause the light to be turned on”). If the discourse structure is annotated with the interpretation the speaker wishes to prefer, the appropriate sentence can be generated.

If, then, when generating text, one keeps track of these rhetorical relations to indicate the reasoning that has already taken place, the constraints that are already known to hold between various units, and which the speaker wishes to highlight, one avoids having to perform that reasoning again, for example, to choose a cue phrase. While this might not be important when the relation is a simple semantic one (e.g., <x> is the color of <y>) as this check can be done easily, it is useful (computationally speaking) when the reasoning that took place is complex (even on a semantic level, for example, to figure out that <x> is a cause of <y>, when no causal link exists in the knowledge base). It is also important when several relations exist, different syntactic realization highlight different relations, and one was preferred by the speaker. In this sense, current relations can be considered as ‘macros’ that represent a set of constraints that hold between various units (semantic units, textual units or intentions). As such, they might turn out to be an essential level of abstraction to represent as part of a discourse structure in a generation system to control realization.

In general, however, there are serious problems with current theories of rhetorical relations such as RST. Several problems stand out:

- It is not always clear what the label actually stands for and what exactly it is used for. The constraints they represent are not well defined, may even differ depending on the surrounding context, and may encompass several different aspects of the discourse.
- As already pointed out, these relations represent different types of constraints. This has also been discussed by other researchers [1–3, 5–8]. Conflating all these factors by using one term “rhetorical relations” leads to confusion.
- Few theories allow for several relations to hold at the same time.¹ Yet, semantic constraints, relations between communicative goals and textual relations usually co-occur [2, 7]. In fact, not only do they co-occur between two units of text, but they may also give rise to different non-isomorphic structures. However, all these relations (or constraints) contribute to generating an appropriate text and all should be represented and taken into account. Furthermore, several constraints of the same type might be present.
- To do an analysis in terms of rhetorical relations such as RST, one is often implicitly also doing an intentional analysis. For example, consider the sentence:

(1) if you cook tonight, (2) I’ll take you out to the movies.

¹Notable exceptions are theories based on Systemics Linguistics, e.g., [5].

Based purely on semantic ground, the sentences would be analyzed as having a condition between their two clauses. The other interpretation of MOTIVATION requires the analyzer to also do the intentional analysis. In truth, of course, both relations co-exist. (This is also discussed in [7].)

3 Conclusions

Communicative goals are necessary to build a generation system capable of both generating coherent discourse and being able to respond to feedback and follow-up questions. Furthermore, rhetorical relations are a useful computational tool to represent constraints we currently don't totally understand, avoid duplicating reasoning from first principles, and provide an appropriate level of interface with the realization component of a system. A theory of rhetorical relations, however, is far from complete and would gain from identifying more clearly what the relations stand for and how they are used.

References

- [1] HOVY, E. H., LAVID, J. L., MAIER, E., MITTAL, V. O., AND PARIS, C. L. Employing Knowledge Resources in a New Text Planner Architecture. In *Aspects of Automated Natural Language Generation*, R. Dale, E. Hovy, D. Rösner, and O. Stock, Eds. Springer-Verlag, Berlin, 1992, pp. 57–73.
- [2] MAIER, E. A. The Extension of a Text Planner for the Treatment of Multiple Links between Text Units. In *Proceedings of the 4th European Workshop on Text Generation* (Pisa, Italy, 1993), pp. 103–114.
- [3] MAIER, E. A., AND HOVY, E. H. Organizing Discourse Structure Relations using Meta-Functions. In *New Concepts in Natural Language Processing*, H. Horacek and M. Zock, Eds. Pinter Publisher, London, 1993, pp. 69–86.
- [4] MANN, W. C., AND THOMPSON, S. A. Rhetorical Structure Theory: towards a functional theory of text organization. *Text* 8 (1988), 243–281.
- [5] MARTIN, J. R. *English Text: System and Structure*. Benjamins, Amsterdam, 1990. (In Press).
- [6] MOORE, J. D., AND PARIS, C. L. Planning Text for Advisory Dialogues: Capturing Intentional, Rhetorical and Attentional Information, 1992. Technical Report from the University of Pittsburgh, Department of Computer Science (Number 92–22) and from USC/ISI, # RS 93-330.
- [7] MOORE, J. D., AND POLLACK, M. E. A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics* 18, 4 (1992).
- [8] REDEKER, G. Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics* 14 (1990), 367–381.
- [9] SCOTT, D. Introduction to the Session on Intersegment Relatedness, 1993. Presented at the Burning Issues in Discourse Workshop, Maratea, Italy, 1993.
- [10] SCOTT, D., AND DE SOUZA, C. Getting the Message across in RST-Based Text Generation. In *Current Research in Natural Language Generation*, R. Dale, C. Mellish, and M. Zock, Eds. Academic Press, Boston, 1990, pp. 47–74.