3. On the text surface the transition from one communicative goal to another can, for instance, be observed from the text layout. Where a new goal sets in, the paragraph structure is often interrupted and a new paragraph begins.

4. Also, the theme of the first sentence of such a new paragraph is in most cases not related to any element in the previous sentence, which is the case when subsequent sentences belong to the same communicative goal. Instead, a new lexeme expressing the new global focus is usually preferred.

The following text fragment exemplifies some of the phenomena identified above.

Example 3: Topic Shift, Paragraph Structuring

> (..) As earlier, more than half of *imported* soft drinks came from Austria, followed by West Germany and Belgium/Luxembourg.
> Also *exports* of mineral water continued to expand rapidly. (...)

The two sentences given in Example 3 include a transition between two text segments which follow two different communicative goals. As an effect, the paragraph topic changes from *import* to *export*, the new paragraph introduces the new topic by placing it at thematic position and a new (surface) paragraph is created.

The Interaction Between Communicative Goals and Rhetorical Relations

From what has been said above it can be concluded that linguistic surface signals as discussed above are ways to realize virtual constructs like rhetorical relations and communicative goals. This means that the interaction between communicative goals and rhetorical relations is one of *realization*. Rhetorical relations are employed to achieve communicative goals. There are possibly other ways of interaction between goals and relations, but they are difficult to observe and of rather speculative nature. Therefore, we restricted the description of the interaction between goals and relations to what is observable.

As pointed out in [Maier and Hovy, '91] three types of relations can be distinguished: ideational, interpersonal and textual relations. Descriptive texts can be characterized by the preferred use of ideational relations while interpersonal relations occur in genres with a high degree of reader involvement (advertisements, personal letters, etc.). Textual relations are unspecific with respect to text types although subsets of the textual relations might be preferably used for some genres ([Maier '93]). Various types of communicative goals are responsible for the use of either ideational or interpersonal relations - in [Maier and Hovy, '91] called "ideational" and "interpersonal" goals, respectively. This labeling does not refer to the nature of the communicative goals; it rather refers to the type of text to be generated and the type of relations to be used. Instead, communicative goals have to be considered an interpersonal device, since they deal with the intentions to be achieved by means of the discourse. In Systemic Functional Linguistics this is exactly what the interpersonal metafunction is about.

The Representation of Rhetorical Relations and Communicative Goals

Both communicative goals and rhetorical relations have been taxonomized and represented in declarative knowledge resources ([Paris and Maier, '91]), which are part of the text planning system described in [Hovy et al., '92]. Both resources are implemented in a way that the selection

of an item (a relation, a goal) results in the execution of associated *realization statements*, which achieve the effects discussed above (e.g. topic shift, preselection of a subset of rhetorical relations,..). In the following we discuss both knowledge resources in turn.

The rhetorical relations are represented in a network, which is traversed during text planning in order to find the best relation to connect the new proposition to the previous text. The realization statements specified for this relation are then executed. Below we give an example for the representation of a relation.

relation: id-sequence
inquiry: id-sequence-query
realization: (SELECT-KNOWLEDGE sequence)
               (PREFER-THEMATIC-PROGRESSION theme-theme)
               (GROW-TREE id-sequence)

The selection of the relation ID-SEQUENCE, which is typically employed to link chronological events, triggers three follow-up actions:

- an event is selected from the knowledge base which stands in a succession relationship to the event which has just been mentioned. Also, relevant information linked to that "new" event (actors, temporal features) has to be retrieved (SELECT-KNOWLEDGE);
- a certain pattern of thematic progression, which is favored by the relation at hand, is determined. In example 2 above, the theme of the chronologically linked sentences is the same ("water") throughout the whole text (function PREFER-THEMATIC-PROGRESSION);
- the text plan is incremented by the new information and linked to the preceding context by means of the relation ID-SEQUENCE (GROW-TREE).

In a similar way, the effect of choosing a communicative goal imposes constraints on the document planning environment. Depending on the type of goal, various realization statements are executed. We distinguish (1) goals responsible for the generation of text segments and (2) goals contributing to the choice of medium and presentation form of an utterance. We will give an example for each:

communicative goal: describe-group-topics
type: describe-group
realization: (PUSH-ON-GOAL-STACK -none-)
                      (HIGHLIGHT-RELATIONS (elaborate-group elaborate-person))
                      (CHANGE-TOPIC to-group)
                      (PREFER-FOCUS (group))

communicative-goal: describe-by-showing
type: describe
realization: (PREFER-PRESENTATION-TYPE picture)

The goals related to the production of text, of which our first example is an instance, can induce the following realization statements:

- if the text unit can be composed of further subtexts and if there are subgoals available to represent these text units they have to be pushed on a goal stack in the order they are supposed to appear in the text (PUSH-ON-GOAL-STACK).

- the relations which are typically used in the text units represented by the goal have to be marked as preferable (HIGHLIGHT-RELATIONS).

- with every change of the communicative goal, the global topic to be dealt with changes accordingly. This change in topic is brought about by the function CHANGE-TOPIC. In terms of the text planning process this function determines the hub from which the generation of this new segment has to be started. The hub represents the instance, where the knowledge selection and the navigation in the knowledge base with respect to the text unit starts.

- where possible thematic progression is determined by means of rhetorical relations; if the context is empty and there is no relation available - for example when a new text unit is generated - or if thematic progression cannot be constrained by the relation chosen the default focus of the paragraph as specified by PREFER-FOCUS is taken.

Goals concerned with the choice of the best way to present information activate only one type of realization statements, which restrict the presentation types to be chosen (PREFER-PRESENTATION-TYPE).

Based on these ideas a new component for the treatment of communicative goals in the framework of Multimedia Document Generation has been developed. This component integrates goals necessary for text planning with intentions employed by the so-called 'Pragmatic Model', which fulfills the task of a presentation planner. This builds on experience developed with the AlFresco project ([Stock et al., forthcoming]).

References

[Hovy et al., '92] E. Hovy, J. Lavid, E. Maier, V. Mittal and C. Paris. Employing Knowledge Resources in a New Text Planner. In: R. Dale, E. Hovy, D. Rösner and O. Stock (eds.). *Proceedings of the 6th International Workshop on Natural Language Generation*, p.57-72, Springer, 1992.

[Maier and Hovy, '91] E. Maier and E.H. Hovy. A Metafunctionally Motivated Taxonomy for Discourse Structure Relations. In: H. Horacek and M. Zock (eds.). *Proceedings of the Third European Workshop on Natural Language Generation*, Judenstein, Austria, 1991.

[Maier, '93] E.A. Maier. The Extension of a Text Planner for the Treatment of Multiple Links Between Text Units, in: *Proceedings of the 4th European Workshop on Natural Language Generation*, April 28-30, 1993, Pisa, Italy, p.103-114. Also available as IRST Technical Report No.9301-15, IRST, Trento, Italy, January 1993.

[Matthiessen and Bateman, '91] C.M.I.M. Matthiessen and J.A. Bateman. *Text Generation and Systemic-Functional Linguistics - Experiences from English and Japanese*. Pinter Publishers, 1991.

[Paris and Maier, '91] C.L. Paris and E.A. Maier. Knowledge Resources or Decisions ? In: *Proceedings of the IJCAI-91 Workshop on Decision Making Throughout the Generation Process*, Syndney, Australia, 1991.

[Stock et al., forthcoming] O. Stock and the AlFresco Project Team. AlFresco - Enjoying the Combination of Natural Language Processing and Hypermedia for Information Exploration. In: M.T. Maybury (ed.). *Intelligent Multimedia Interfaces*. AAAI press, forthcoming.

On Structure and Intention

Mark T. Maybury*

Abstract

This position paper contrasts rhetorical structuring of propositions with intentional decomposition using communicative acts. We discuss the kinds of information current explanation planners capture in their plan operators and propose extensions to these. In Maybury (1992b) we detail how these plans can and have been extended to capture a more general notion of communication as action, describing other types of communicative acts such as graphical acts and discourse acts. Our current efforts (Maybury, 1992b, forthcoming) are focused on developing a taxonomy of multimedia communication acts which attempt to distinguish semantic relations, rhetorical relations and intentions.

Rhetorical Structuring versus Intention Decomposition

A number of researchers have investigated using structural analyses of text, including Rhetorical Structure Theory (RST) (Mann and Thompson. 1987), as the basis for explanation planning architectures. For example, using rhetorical relations such as background and elaboration, Hovy's (1988) system constructs a rhetorical structure over a given set of propositions (See Figure 1a). Moore's (1989) system also constructs a rhetorical structure, however, the leafnodes of the resulting tree are illocutionary acts (e.g., inform) with associated propositions. While we agree that text contains relations between parts, we also concur with the position held by Suthers (1991) and others that rhetorical relations, in their current form, conflate a number of issues including intention, structure, linear precedence, and epistemological distinctions. Hovy (1990) details problems with RST approaches to paragraph planning, including algorithmic problems and, more seriously, problems with the theory and representation of coherence relations.

In contrast to RST-based planners but similar to rhetorical schema based generators, our explanation planning architecture uses "rhetorical predicates" (e.g., attribution, evidence, enablement) to abstractly characterize epistemological content and relations in the underlying knowledge base. As in McKeown (1982), some of these predicates indicate local relations (e.g., illustration) and have associated cue words (e.g., "for example") or associated semantic actions (e.g., "contains," "enables"). However, other predicates, such as attribution or definition, have no marked relation to their surrounding text (only the weak notion of elaboration). In our attempts to generate the range of text types ranging from narration to argument, we have found the need to develop a correspondingly broad range of rhetorical predicates, including logical-definition, synonymic-definition, constituency, classification, evidence, motivation, etc.

---

*Mail Stop K329, Artificial Intelligence Center, The MITRE Corporation, Burlington Road, Bedford, MA 01730. (617) 271-7230. maybury@linus.mitre.org.

74

We use these same rhetorical predicates to abstractly mark the epistemological content of speech acts (e.g., request or inform). An example action in our system might be `INFORM(#<system>, #<user-023>, logical-definition(#<Ferrari-Testarossa>))` which says "have the system inform user-023 of the logical definition of the object, `#<Ferrari-Testarossa>`," which might eventually result in the utterance "A Ferrari Testarossa is a fast, sleek Italian sports car". In order to retrieve the content for a "logical definition" predicate, we must not only look up the genus of the entity, but also calculate its differentia, or distinguishing characteristics (Maybury, 1990). Thus, the relation between rhetorical predicates and semantic relations in the underlying knowledge base is not a simple one-to-one mapping; in some cases the content must be calculated. Moreover, content may be modulated by context or by a user model (e.g., choosing the perspective from which to view an object, if it has multiple superordinates (McCoy, 1985)).

Our architecture actually distinguishes between illocutionary acts (e.g., inform, request) and surface speech/locutionary acts (e.g., assert, command, suggest) which have associated surface forms (e.g., declarative imperative, interrogative mood). In our architecture, the organization and structure of illocutionary speech acts such as the above inform action is accomplished by more abstract rhetorical acts (e.g., describe, compare, argue). Rhetorical acts characterize the communicative action performed by one or more utterances, and correspond to the text types such as description, narration, and exposition. Because our focus has been on formalizing the communicative actions that underlie texts, we have worked toward a unified view of rhetorical and speech acts. Therefore, our approach can be seen as an extension of theoretical work which views language as purposeful behavior (Austin, 1962; Searle, 1969) and of computational implementations of speech acts (Cohen, 1978; Allen, 1979; Appelt, 1982). As we dicuss below, we have also investigated using the notion of rhetorical acts to characterize both linguistic and non-linguistic acts, resulting, for example, in mixed text and graphics.

We formalize communicative acts (speech acts and rhetorical acts) as plan operators. A hierarchical planner reasons about these operators in order to produce a text plan (an executable action decomposition) that achieves some given discourse goal (see Figure 1b). The planner actually produces two structures: the action decomposition shown in Figure 1b as well as a corresponding effect decomposition in which each level represents to the effects achieved by each act in the action decomposition. In the architecture implemented in our system TEXPLAN, the decomposition of plan operators captures the hierarchical structure and order of intentions underlying text. Thus our architecture differs from work in planned rhetorical relations (Hovy, 1988; Moore, 1989) in that it recognizes and formalizes the distinction between the rhetorical relations in a text (e.g., evidence, enablement, purpose) and the rhetorical acts establishing these. And as we will discuss in a detailed position paper, there are also differences in the representation of preconditions and effects.

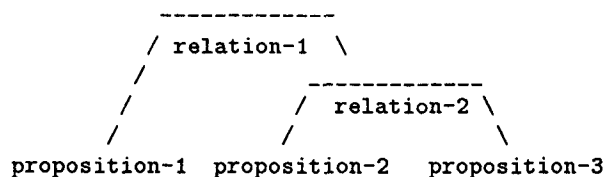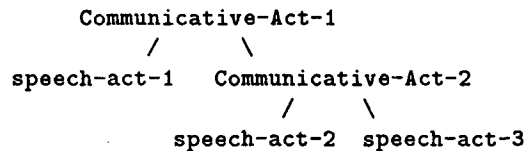Figure 1a. The Content and Structure of Resulting Explanations, Relation-Based:

```
        _____
       / relation-1  \
      /         _____
     /         /  relation-2 \
    /         /               \
proposition-1 proposition-2  proposition-3
```

Figure 1b. Communicatiove-act-based.

```
        Communicative-Act-1
           /        \
  speech-act-1   Communicative-Act-2
                    /      \
             speech-act-2  speech-act-3
```

Conclusion

In our research we have found that there are at least four generic types of text: description, narration, exposition, and argument. These text types form the basis of explanations which convey different propositional content (e.g., entities and relations versus events and states), have particular intended effects on the addressee's knowledge, beliefs, and desires, and are compositional (e.g., narration can invoke description). In the extended position paper we contrast two architectures for explanation planning: rhetorical structuring of propositions versus communicative act-based explanation planning. In our work we consider the structure of plan operators, including issues of constraints, preconditions, effects, and decomposition, and have discussed (Maybury, 1992b) how current representations might be extended, and also consider the applicability to plan multimedia explanations and discourse. After considering issues concerning plans and focus models, we conclude by indicating that current plan-based architectures suffer from a number of fundamental architectural deficiencies that stem from the current state of the art in planning techniques. This situation is exacerbated by the current lack of understanding of the nature of and relationshiip among attention, intensions and rhetorical relations.

References

Allen, J. F. 1979. A Plan-based Approach to Speech Act Recognition. Ph.D. dissertation, Department of Computer Science, University of Toronto, Toronto, Canada.

Appelt, D. E. March, 1982. Planning Natural Language Utterances to Satisfy Multiple Goals. SRI Technical Note 259.

Austin, J. 1962. How to do Things with Words. editor J. O. Urmson. England: Oxford University Press.

Cawsey, A. 1989. "Explanatory Dialogues." Interacting with Computers 1(1):69-92.

Cohen, P. R. 1978. On Knowing What to Say: Planning Speech Acts. University of Toronto TR-118.

Grosz, B. J. and C. Sidner, 1989. Plans for Discourse. Intentions and Communications, editors P. Cohen, J. Morgan and M. Pollack. MIT Press. [Harvard University TR-11-87].

Hovy, E. 1988. Planning Coherent Multisentential Text. Proceedings of the 26th Meeting of the ACL, Buffalo, NY, June 7-10, 1988. 163-169.

Hovy, E. 1990. Unresolved issues in paragraph planning. In Dale, R., Mellish, C., Zock, M. Current Research in Natural Language Generation, London: Academic Press.

Mann, W. C. and S. A. Thompson. 1987. "Rhetorical Structure Theory: Description and Construction of Text Structures." Natural Language Generation, editor G. Kempen. 85-95. Dordrecht: Martinus Nijhoff.

Maybury, M. T. 1990. "Generating Natural Language Definitions from Classification Hierarchies" in Susanne Humphrey, ed., 1991. ASIS Monographs Series, Advances in Classification Research and Application: Proceedings of the 1st ASIS SIG/CR Classification Research Workshop, Toronto, Canada, November 4, 1990, Learned Information: Medford, NJ, ISBN 0-938734-53-9.

Maybury, M. T. 1991a. "Planning Multimedia Explanations using Communicative Acts", Proceedings of the Ninth National Conference on Artificial Intelligence, AAAI-91, July 14-19, 1991, Anaheim, CA.

Maybury, M. T. 1991b. Planning Multisentential English Text using Communicative Acts. Cambridge University Ph.D. dissertation.

Maybury, M. T. 1991c. "Topical, Temporal, and Spatial Constraints on Linguistic Realization" Computational Intelligence: Special Issue on Natural Language Generation. Volume 7(4), December, 1991.

Maybury, M. T. April, 1992a. "A Critique of Text Planning Architectures" Journal of the International Forum on Information and Documentation (IFID). 17(2):7-12. Special issue on the Bijormi Text Generation Symposium, Bijormi, Georgia, USSR, 23-37 September, 1991.

Maybury, M. T. August, 1992b. "Communicative Acts for Explanation Generation" International Journal of Man-Machine Studies. 37(2):135-172.

Maybury, M. T. forthcoming Intelligent Multimedia Interfaces. AAAI/MIT Press.

McCoy, K. F. December, 1985. Correcting Object-Related Misconceptions. Ph.D. dissertation, University of Pennsylvania TR MS-CIS-85-57, Philadelphia, PA.

McKeown, K. R. 1982. Generating Natural Language Text in Response to Questions About Data Base Structure. Ph.D. dissertation, University of Pennsylvania TR MS-CIS-82-5.

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

Paris, C. L. 1987. The Use of Explicit User Models in Text Generation: Tailoring to a User's Level of Expertise. Ph.D. dissertation, Columbia University, NY.

Pollack, M. 1986. Inferring Domain Plans in Question-answering. University of Pennsylvania Ph.D. dissertation, Philadelphia, PA.

Searle, J. R. 1969. Speech Acts. Cambridge University Press.

Sidner, C. L. 1979. Toward a Computational Theory of Definite Anaphora Comprehension in English Discourse. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Suthers, D. 1991. Task-Appropriate Hybrid Architectures for Explanation. AAAI-91 Workshop on Evaluation of Explanation Planning Architectures, Anaheim, CA, 14 June, 1991.