Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Research
Report
RR-92-45

The Design of Illustrated Documents as a Planning Task

Elisabeth André, Thomas Rist

September 1992

Deutsches Forschungszentrum für Künstliche Intelligenz
GmbH

Postfach 20 80
D-6750 Kaiserslautern, FRG
Tel.: (+49 631) 205-3211/13
Fax: (+49 631) 205-3210

Stuhlsatzenhausweg 3
D-6600 Saarbrücken 11, FRG
Tel.: (+49 681) 302-5252
Fax: (+49 681) 302-5341

Deutsches Forschungszentrum für Künstliche Intelligenz

The German Research Center for Artificial Intelligence (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI) with sites in Kaiserslautern and Saarbrücken is a non-profit organization which was founded in 1988. The shareholder companies are Atlas Elektronik, Daimler Benz, Fraunhofer Gesellschaft, GMD, IBM, Insiders, Mannesmann-Kienzle, Philips, SEMA Group Systems, Siemens and Siemens-Nixdorf. Research projects conducted at the DFKI are funded by the German Ministry for Research and Technology, by the shareholder companies, or by other industrial contracts.

The DFKI conducts application-oriented basic research in the field of artificial intelligence and other related subfields of computer science. The overall goal is to construct *systems with technical knowledge and common sense* which - by using AI methods - implement a problem solution for a selected application area. Currently, there are the following research areas at the DFKI:

- Intelligent Engineering Systems
- Intelligent User Interfaces
- Intelligent Communication Networks
- Intelligent Cooperative Systems.

The DFKI strives at making its research results available to the scientific community. There exist many contacts to domestic and foreign research institutions, both in academy and industry. The DFKI hosts technology transfer workshops for shareholders and other interested groups in order to inform about the current state of research.

From its beginning, the DFKI has provided an attractive working environment for AI researchers from Germany and from all over the world. The goal is to have a staff of about 100 researchers at the end of the building-up phase.

Prof. Dr. Gerhard Barth
Director

The Design of Illustrated Documents as a Planning Task

Elisabeth André, Thomas Rist

DFKI-RR-92-45

A version of this paper will appear in:
Intelligent Multimedia Interfaces, M. Maybury (ed.), AAAI Press

This work has been supported by a grant from The Federal Ministry for Research and Technology (FKZ ITW-8901 8).

Table of Contents

The Design of Illustrated Documents as a Planning Task

Elisabeth André, Thomas Rist
German Research Center for Artificial Intelligence (DFKI)
W-6600 Saarbrücken 11, Stuhlsatzenhausweg 3, Germany
Phone: (+49 681) 302-5252, Fax: (+49 681) 302-5341
E-mail: {andre, rist}@dfki.uni-sb.de

Abstract:

Not only the generation of text, but also the generation of multimodal documents can be considered as a sequence of communicative acts which aim to achieve certain goals. For the realization of a system able to automatically generate illustrated documents, a plan-based approach seems adequate. To represent knowledge about how to present information, we have designed presentation strategies which relate to both text and picture production. These strategies are considered as operators of a planning system. However, a conventional hierarchical planner for determining the contents and the rhetorical structure of a document has proven inappropriate to handle the various dependencies between content determination, mode selection and content realization. To overcome these problems, a new planning scheme has been developed that supports data transfer between the content planner and the mode-specific generation components and allows for revising an initial document structure.

1 Introduction

Recently, there has been increasing interest in the design of systems generating multimodal output. Research in this area addresses the analysis and representation of presentation knowledge (cf. [Arens et al., this volume]) as well as computational methods for the automatic synthesis of multimodal presentations (cf. [Badler et al. 91], [Feiner/McKeown 91], [Marks/Reiter 90], [Maybury, this volume], [Roth et al. 91] and [Wahlster et al. 91]). There is general agreement that a multimodal presentation system cannot simply merge the results of the mode-specific generators, but has to carefully tailor them to each other. Such tailoring requires knowledge concerning the functions of textual and pictorial document parts and the relations between them. Furthermore, a presentation system must be able to handle the various dependencies between content planning, mode selection and content realization.

In the following, we will show that many concepts applied in natural language generation, such as communicative acts and coherence relations, can be adapted to the generation of text-picture combinations. We will present an approach that integrates content planning and mode selection and allows for interaction with mode-specific generators. This approach has been integrated into the multimodal presentation system WIP (cf. [André et al., this volume]) which generates illustrated instructions for technical devices.

2 The Structure of Illustrated Documents

Our approach is based on the assumption that not only the generation of text, but also the generation of multimodal documents can be considered as an act sequence that aims to achieve certain goals (cf. [André/Rist 90a]). We presume that there is at least one act that is central to the goal of the whole document. This act is referred to as the *main act*. Acts supporting the main act are called *subsidiary acts*.[1] Main and subsidiary acts can, in turn, be composed of main and subsidiary acts. The root of the resulting hierarchical structure generally corresponds to a complex communicative act such as describing a process, and its leaves are elementary acts, i.e., speech acts (cf. [Searle 69]) or pictorial acts (cf. [Kjorup 78]).

The structure of a document is, however, not only determined by its hierarchical act structure, but also by the role acts play in relation to other acts. In textlinguistic studies, a variety of coherence relations between text segments has been proposed (e.g., see [Grimes 75] and [Hobbs 78]). Perhaps the most elaborated set is presented in RST-theory (cf. [Mann/Thompson 87]). Examples of RST-relations are *Motivation*, *Elaboration*, *Enablement*, *Interpretation* and *Summary*. Text-picture researchers have investigated the role a particular picture plays in relation to accompanying text passages. E.g., Levin has found five primary functions (cf. [Levin et al. 87]): *Decoration*, *Representation*, *Organization*, *Interpretation* and *Transformation*. Hunter and colleagues distinguish between: *Embellish*, *Reinforce*, *Elaborate*, *Summarize* and *Compare* (cf. [Hunter et al. 87]). An attempt at a transfer of the relations proposed by Hobbs to pictures and text-picture combinations has been made in [Bandyopadhyay 90]. Unfortunately, text-picture researchers only consider the communicative functions of whole

[1] This distinction between main and subsidiary acts essentially corresponds to the distinction between *global* and *subsidiary speech acts* in [Searle 69], *main speech acts* and *subordinate speech acts* in [Van Dijk 80], *dominierenden Handlungen* and *subsidiären Handlungen* in [Brandt et el. 83] and between *nucleus* and *satellites* in the RST-Theory proposed in [Mann/Thompson 87].

pictures, i.e., they do not address the question of how a picture is organized. To get an informative description of the whole document structure, one has to consider relations between picture parts or between picture parts and text passages too. E.g., a portion of a picture can serve as background for the rest of the picture or a text passage can elaborate on a particular section of a picture.

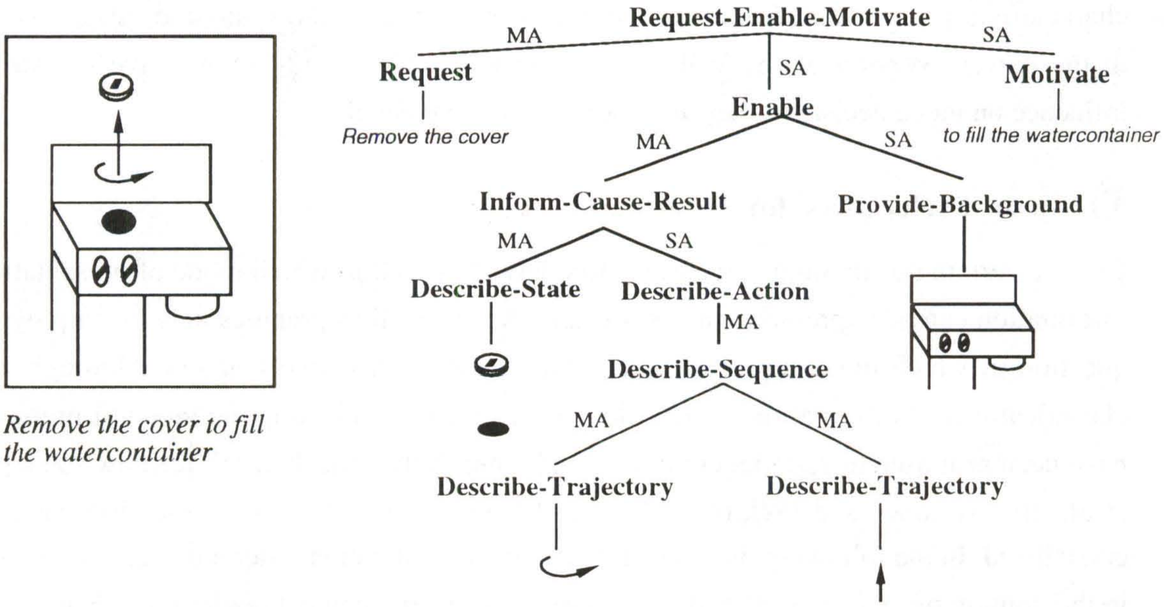

Fig. 1: A Document Fragment[2] and its Structure

In Fig. 1, an example document fragment and its discourse structure are shown. The goal of this document fragment is to instruct the user in removing the cover of the water container of an espresso machine. The instruction can be considered as a composite goal comprising a request, a motivation and an enablement part. The request is conveyed through text (main act (MA)). To motivate that request, the author has referred to a superordinate goal, namely filling the water container (subsidiary act (SA)). The picture provides additional information which enables the addressee to carry out the request (subsidiary act). The generation of the picture is also subdivided into a main act, which describes the result and the actions to be performed, and a subsidiary act, which provides the background to facilitate orientation.

[2] The example is a slightly modified and translated version of instructions for the Philips espresso-machine HD 5649.

3 Design Criteria for Text-Picture Combinations

When designing an illustrated document, an author has to decide which mode combination is the most suitable for meeting his goals. The decision-making process for mode selection is influenced by different factors including: the kind of information content, the communicative functions that textual and pictorial document parts ought to fill in a presentation, resource limitations (e.g., limitations due to the output medium, or space and time restrictions), user characteristics (e.g., trained or untrained in reading visual displays) and the user's task. Since in the current version of the WIP system the first and second factors have the strongest influence on mode decisions, they are examined in more detail.

3.1 Mode Preferences for Information Types

Given a certain information content, we first have to check in which mode of presentation the information can be expressed. In cases where text as well as graphics may be employed, the question of which mode conveys the information most effectively arises. Although several classifications of information content that are relevant for selecting the mode of presentation have been proposed (e.g., [Bieger/Glock 84], [Roth/Mattis 90], [Feiner/McKeown 91], [Arens et al., this volume] and [Whittaker/Walker 91]), an exhaustive classification has not yet crystallized. In the following, we will present some classification criteria that are of importance in the domain of maintenance and repair instructions for technical devices. Of course, further criteria are necessary, in particular when shifting to another domain.

Concrete information: Information concerning visual properties of concepts (such as shape, color and texture) is classified as concrete. We regard events and actions as concrete if they involve physical objects and if their occurrence causes visually perceptible changes. Since pictures seem to be superior in teaching perceptual concepts (e.g., see [Molitor et al. 89]), graphics will be used in preference to text when presenting concrete information.

Spatial information: Since space is conceptualized mainly through objects, the category of spatial information primarily includes information concerning the location, orientation and composition of objects. Furthermore, physical events and actions mostly have a spatial component. Since a movement of a physical object can be characterized by means of spatial concepts (such as the direction of movement or the starting and end position), actions and events also get the attribute *spatial* if they involve movements of physical objects. In deciding how to present spatial information, we can partly fall back on empirical psychological studies. E.g., Bieger and Glock (cf. [Bieger/Glock 86]) found that in assembly instructions spatial information is perceived faster if pictures are used; on the other hand, subjects confronted with textual presentations make fewer mistakes when carrying out instructions. Thus, if the emphasis is on speed, pure pictorial presentations of spatial information should be preferred.

Temporal information: In the domain of operating instructions, the temporal relations between states, events or actions play an important part. The sequential order of events can be effectively communicated by arranging pictures from top to bottom or from left to right. In some cases, subsequent events can even be depicted in a single picture (cf. Fig. 1). While precedence relations can be easily communicated through pictures, the fact that two events overlap in time is hard to express pictorially. Furthermore, for a number of time specifications, such as *mostly, periodically* or *in the future*, textual presentations should be preferred in order to avoid misconceptions.

Covariant information: Covariant information expresses a semantic relationship between at least two pieces of information that vary together. Such relationships are: *cause/effect*, *action/result*, *problem/solution*, *condition*, and *concession*.[3] Cause/effect and action/result relationships are often expressed through a single picture, a sequence of pictures or through a text-picture combination. The presenter has, however, to consider that cause/effect and action/goal relationships between (parts of) pictures are often interpreted as pure temporal relationships. If it is not certain whether the addressee recognizes the intended relationship, text should be used in preference to graphics. To ensure that a problem/solution relationship is correctly interpreted, the problem should be presented in text unless a kind of picture language is used (e.g., in [Strothotte/Schmid 90], a question mark indicates that a picture presents a problem.). The relationships *condition* and *concession* can hardly be expressed by graphics without verbal comments.

Quantification: In general, it is very difficult to graphically depict quantifiers. Even if quantification is to be done over finite sets of physical objects and it seems to be straightforward to communicate quantifying information by graphically enumerating instances, a viewer will be confused if he does not recognize whether the picture is meant to show a complete set or most/some/any/exactly-n/etc. instances. Apart from this, such pictorial enumerations tend to be long-winded and waste space in a document.

Negation: Although there is no "natural" way to graphically express negation, some kinds of negation are frequently expressed using conventionalized graphical symbols. Perhaps the most widespread convention is the use of overlaid crossing bars. E.g., in graphical warnings where a technical device is shown in a particular state, crossing bars indicate that this state must not be

---

3 These relations also appear in RST-theory to describe a semantic relation between real world entities.

achieved. However, a viewer may have difficulties in figuring out the scope of a negation symbol. Furthermore, it is questionable whether already conventionalized negation symbols can be employed for other kinds of negation, e.g., to express the absence of objects or object attributes.

3.2 Achievement of Communicative Goals

As mentioned before, mode decisions depend not only on the kind of information to be communicated, but also on the communicative function of an utterance. There is no doubt that many communicative acts (e.g., *describe*, *inform* or *warn*) can be accomplished with pictures (cf. [Novitz 77]). In this section, we will concentrate on communicative functions that pictures fulfill in relation to text or other pictures. Some of these functions have also been identified by text-picture researchers, and most of them correspond to pragmatic relations in RST-theory.

Attract-Attention: The text directs the addressee's attention to special aspects of the picture/text. E.g. directives, such as "Look at ..." can be used to tell the addressee what is important in a picture. Furthermore, a part of a picture can emphasize other document parts, e.g., arrows pointing to important objects.[4]

Compare: Two document parts provide a comparison between several concepts. To emphasize the differences or parallels between the concepts, the same presentation modes should be used for describing the concepts.

Elaborate: One part of a document provides further details about another part. Text can elaborate on a picture, e.g., by specifying attributes of an object shown in the picture. On the other hand, a picture can elaborate on text, e.g., by showing an object belonging to a verbally described class. Pictures can also elaborate on other pictures, e.g., consider an inset that shows further details of a depicted object.

Enable: The picture/text provides additional information in order to enable the addressee to perform the requested action. E.g., a request may be accompanied by a picture showing how an action should be carried out. The request is typically conveyed by text.

Elucidate: One document part provides an explanation or interpretation of another part. E.g., text can be used to express the meaning of a picture or to clarify graphical techniques. While text can explain pictures or text passages, pictures can explain text, but normally not other pictures (cf. [Muckenhaupt 86]).

---

[4] This situation must not be confused with situations where a document part indirectly attracts the addressee's attention because of its visual appearence (e.g., because of its size, position or color).

Label: A piece of text serves as a label for a portion of the document. Typical examples of the label-relationship are: *headline/paragraph*, *caption/figure* and *name/picture part.*

Motivate: The addressee is to be motivated to comply with a request. This goal can be met by means of pictures or by means of text. Consider an advertisement showing a cup of steaming coffee to motivate people to buy this coffee. Typically, the request implicitly follows from the context or is explicitly conveyed through text.

Evidence: The picture/text produces evidence for a verbal claim. Since pictures increase authenticity (cf. [Smith/Smith 66]), they are well suited to support a claim. Typical examples are TV news.

Background: One document part establishes the context for the other. E.g., text may provide the necessary background information for a picture that shows a device from an extraordinary perspective. Background can also be provided by parts of a picture, e.g., a picture of an object may include further objects in order to reduce ambiguities by showing the object's spatial context.

Summarize: The picture/text provides an organized, reduced form of the text structure. E.g., a picture may be presented in advance to show the most important parts of a machine which are described in detail by text. On the other hand, text may be used to summarize the contents of a picture.

4 Representation of Presentation Knowledge

To generate multimodal presentations, we have defined a set of presentation strategies that can be selected and combined according to a particular presentation task. Such presentation strategies reflect general presentation knowledge as indicated in the preceding section, or they embody more specific knowledge of how to present a certain subject.

To represent presentation strategies, we follow the approach proposed by Moore and colleagues (cf. [Moore/Paris 89] and [Moore/Swartout 89]) to operationalize RST-theory for text planning. The strategies are represented by a name, a header, an effect, a set of applicability conditions and a specification of main and subsidiary acts. Whereas the header of a strategy is a complex communicative act (e.g., to enable an action), its effect refers to an intentional goal (e.g., the user knows a particular object).[5] To represent intentional goals, we use the same notation as in

[5] In [Moore/Paris 89], this distinction between header and effect is not made because the effect of their strategies may be an intentional goal as well as a rhetorical relation.

Hovy's RST planner (cf. [Hovy 88]). The expression (Goal P x) stands for: The presenter P has x as a goal. (Bel P x) should be read as: P believes that x is satisfied. (BMB P A x) is an abbreviation for the infinite conjunction: (Bel P x) & (Bel P (Bel A x)) & (Bel P (Bel A (Bel P x))), etc. The applicability conditions specify when a strategy may be used and constrain the variables to be instantiated. The main and subsidiary acts form the kernel of the strategies. Examples of presentation strategies are shown below. The first strategy can be used to request the user to perform an action. Whereas text is used to perform the main acts, the mode for the subsidiary acts is open. In this strategy, three kinds of acts occur: the elementary act S(urface)-Request, three referential acts for specifying the action and the semantic case roles associated with the action (Activate), and two complex communicative acts (Motivate and Enable).

```
[S1]  Name:
          Request-Enable-Motivate
      Header:
          (Request P A ?action T)6
      Effect:
          (BMB P A (Goal P (Done A ?action)))
      Applicability Conditions:
          (And  (Goal P (Done A ?action)) (Bel P (Complex-Operating-Action ?action))
                (Bel P (Agent ?agent ?action)) (Bel P (Object ?object ?action)))
      Main Acts:
          (S-Request P A (?action-spec (Agent ?agent-spec) (Object ?object-spec)))
          (Activate P A (Action ?action) ?action-spec T)
          (Activate P A (Agent ?agent) ?agent-spec T)
          (Activate P A (Object ?object) ?object-spec T)
      Subsidiary Acts:
          (Motivate P A ?action ?mode-1)
          (Enable P A ?action ?mode-2)
```

The second and third strategies may be employed to show the orientation of an object and to enable its identification in a picture (see also [André/Rist 90b]).

```
[S2]  Name:
          Describe-Orientation
      Header:
          (Describe P A (Orientation ?orientation) G)
      Effect:
          (BMB P A (Has-Orientation ?orientation ?x))
      Applicability-Conditions
          (Bel P (Has-Orientation ?orientation ?x))
      Main Acts:
          (S-Depict P A (Orientation ?orientation) ?p-orientation ?pic)
      Subsidiary Acts:
          (Achieve P (BMB P A (Identifiable A ?x ?px ?pic)) ?mode)
```

---

6 T stands for text, G for graphics.

```
[S3] Name:
        Provide-Background
     Header:
        (Background P A ?x ?px ?pic G)
     Effect:
        (BMB P A (Identifiable A ?x ?px ?pic))
     Applicability Conditions:
        (AND (Bel P (Image-of ?px ?x ?pic))
             (Bel P (Perceptually-Accessible A ?x))
             (Bel P (Part-of ?x ?z)))
     Main Acts:
        (S-Depict P A (Object ?z) ?pz ?pic)
     Subsidiary Acts:
        (Achieve P (BMB P A (Identifiable A ?z ?pz ?pic)) ?mode)
```

When defining presentation strategies, one has to decide whether to define relatively specific strategies by anticipating important design decisions, e.g., about mode selection, or whether to define more general presentation strategies, e.g., by leaving mode decisions open. By constraining design decisions, we can avoid situations in which decisions have to be retracted because they are not realizable. However, we have to take care that we do not unnecessarily restrict the set of possible designs. Strategy [S1] can be considered as a compromise between these two approaches. Whereas the mode for the subsidiary acts is left open, the strategy prescribes text for the main acts.

Since there may be several strategies for achieving a certain goal, we need criteria for ranking the effectiveness, the side-effects and costs of executing presentation strategies. To formulate selection criteria, we use meta rules.

```
[M1] IF (IS-A ?current-attribute-value Spatial-Concept)
        THEN (Dobefore *graphics-strategies* *text-strategies*)
```

E.g., the metarule [M1] suggests a preference for graphics over text when presenting spatial information. The studies listed in section 3 form the theoretical basis of such meta rules.

5 The Presentation Planning Process

To automatically generate documents, one not only has to identify and represent relevant presentation knowledge, but also has to operationalize the synthesis process.

5.1 The Basic Planning Scheme

Presentation strategies are treated as operators of a planning system. The basic idea behind the planning process is as follows: For each presentation goal, try to find strategies which are either specified by the header or whose effect matches the presentation goal. Check for which variable bindings the applicability conditions of the strategies hold. All strategies whose applicability conditions are satisfied become candidates for achieving the presentation goal. If several strategies are applicable, prioritize them employing metarules. Then select a strategy, instantiate it and post the main and subsidiary acts as new subgoals or - in the case of elementary acts such as 'S-Depict' or 'S-Assert' - write them into the task queues of the mode-specific generators. In case a subgoal cannot be achieved, backtrack. The planning process terminates if all goals are expanded to elementary acts that can be realized by the text or graphics generator. The result of the planning process is a refinement-style plan in the form of a directed acyclic graph (DAG).

To ensure that document fragments in multiple modalities are smoothly tailored to each other in the document to be generated, one also has to consider various dependencies between content determination, mode selection and content realization. As a consequence, the flow of control is more complicated than described above.

5.2 Interleaving Content Planning, Mode Selection and Content Realization

Previous work on natural language generation has shown that content selection and content realization should not be treated independently of each other (see also [Hovy 87] and [Reithinger 91]). A strictly sequential model in which data only flow from the "what to present" to the "how to present" part has proven inappropriate because the components responsible for selecting the contents would have to anticipate all decisions of the realization components. This problem is compounded if, as in our case, content realization is done by separate components (currently a text and a graphics generator) of which the content planner has only limited knowledge.

It seems inappropriate to sequentialize content planning and mode selection even though mode selection is only a very rough decision about content realization. On the one hand, mode selection depends to a large extent on the information to be communicated (cf. section 3). On the other hand, content planning is strongly influenced by previously selected mode combinations. E.g., to graphically refer to a physical object, we need visual information that may be irrelevant to textual references.

A better solution is to interleave content planning, mode selection and content realization. In the WIP system, we interleave content and mode selection using a uniform planning mechanism. This has become possible since the presentation strategies and metarules accessed by the planner contain not only knowledge about what to present, but also knowledge about adequate mode combinations. In contrast to this, presentation planning and content realization are performed by separate components that access disparate knowledge sources. This modularization enables parallel processing, but makes interaction between the single components necessary. As soon as the planner has decided which generator should encode a certain piece of information, this piece should be passed on to the respective generator. Conversely, the planning component should immediately incorporate the results of the generators. Therefore, the processing of all components has to be 'interrupted' at certain decision points to allow other components to react.

However, we cannot presume that the results of the single components are always available at a given time. It might happen that the planner is not able to expand a node because it is still waiting for a generator to supply realization results. If this generator, in turn, is also waiting for the planner or another generator to provide new data, a deadlock occurs. To cope with uncertainties concerning the results of other components, WIP's presentation planner maintains partial descriptions of unspecified variables through the use of constraints. Thus, it is able to continue planning without premature commitment. Furthermore, it does not always expand nodes in a depth-first fashion, but flexibly selects the nodes to be expanded as illustrated in Fig. 2. Assume that the expansion of node B relies on information provided by executing the elementary act A (cf. Fig. 2a). To avoid time delays, C is expanded first (cf. Fig. 2b). After A has been executed, the required information is available and B can be expanded (cf. Fig. 2c).

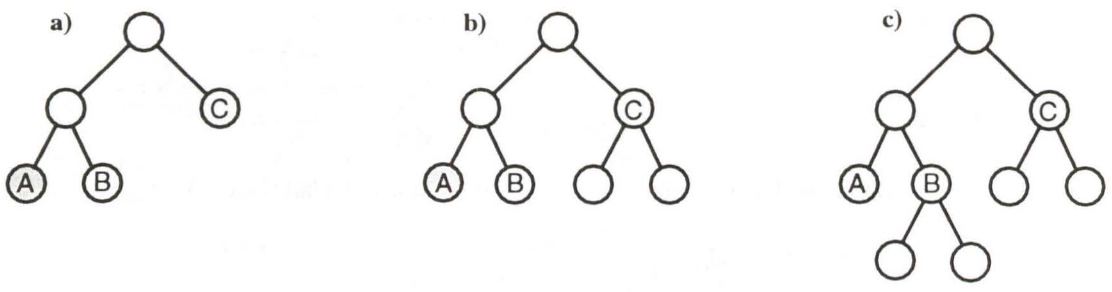

Fig. 2: Opportunistic Node Expansion

Since the generators provide information about (partial) results as soon as possible, situations seldom occur in which information is missing for every plan node to be expanded. In such cases, the planner can select a node considering metrics, e.g., the costs of the assumptions to be made.

5.3 Propagating Data During Presentation Planning

Since every component has only limited knowledge of other components, data have to be passed from one component to the other. E.g., if a generator finds a better solution or is not able to satisfy a task, it has to inform the planner, which has to reorganize its initial plan (see also section 5.4) or to backtrack. To ensure the consistency of the document, all changes have to be propagated to other branches of the plan structure.

Information must not only flow between the content planner and the generators, but also from one generator to the other. Suppose the text generator has generated a referring expression for an object shown in a picture. If the picture is changed due to graphical constraints, it might happen that the referring expression no longer fits. Thus, the planner will have to create a new object description and pass this description on to the text generator, which will have to replace the initial referring expression by a new one.

Furthermore, the need for propagating data during presentation planning arises when dealing with dependencies between presentation strategies. E.g., a decision about mode selection often depends on earlier decisions. Assume the system decides to compare two objects by describing the different values of a common attribute. At this time, the only restriction is that both descriptions should be realized in the same mode. Once the system has decided on the mode for the attribute value of the first object, the result of this decision must be made available for describing the value of the second object. We handle this problem by passing mode information during the planning process both from top to bottom and from bottom to top (cf. Fig. 3).

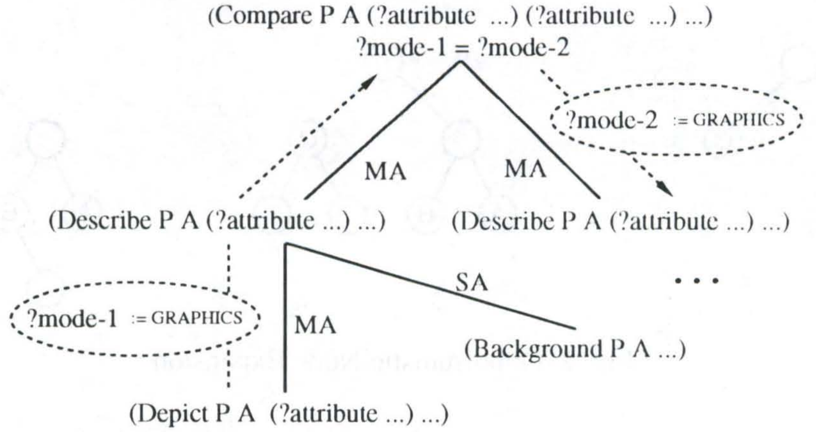

Fig. 3: Passing of information

Mode information is propagated via the header of a strategy. Depending on whether the main acts of a strategy are to be realized in text, graphics or both modes, the values T(ext), G(raphics) or M(ixed) are assigned. The mode remains unspecified until mode decisions are made for the main acts of a strategy. By deferring mode decisions for as long as possible, the planner is able to continue planning without making too specific selections.

5.4 Restructuring after Realization

Since the content planner has no access to realization knowledge of the generators, it cannot consider this knowledge when building up the document structure. As a consequence, it may happen that the results provided by the generators deviate to a certain extent from the initial document plan. Such deviations are reflected in the DAG by output sharing, structure sharing and structure adding. Although in the following examples, restructuring is caused by decisions of the graphics generator, there is no question that restructuring methods are also useful for text generation (e.g., see [Hovy 90]).

Output Sharing

By studying multimodal documents, we found that authors often use one and the same picture or picture part for different purposes. When designing a system for automatic generation, one must determine which component decides when to reuse a picture or picture part. Since the content planner has no knowledge about how information is encoded graphically, the final decision should be left up to the graphics designer. If document parts are reused, this has to be reflected in the document structure as shown in Fig. 4.

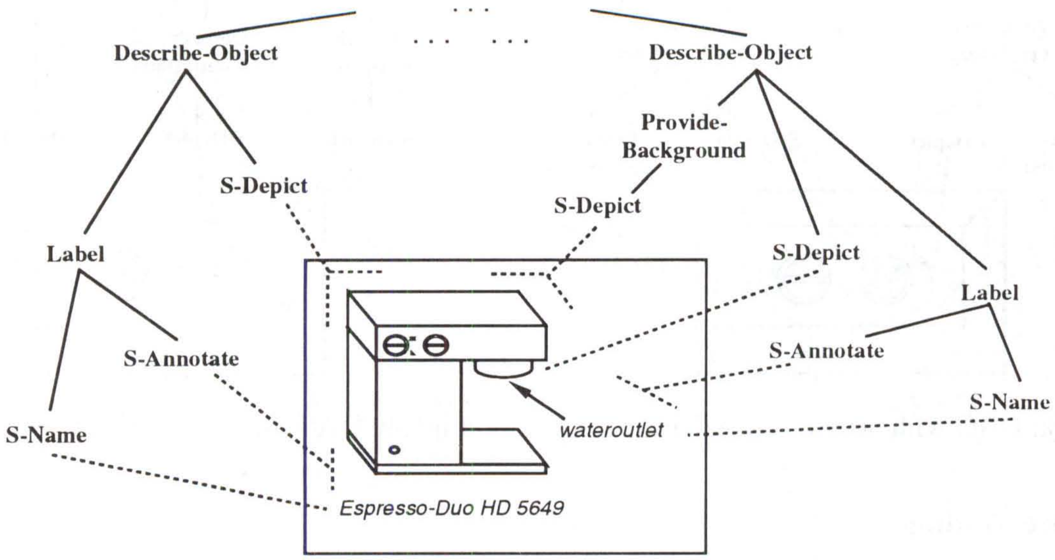

Fig. 4: DAG with Output Sharing

Suppose the planner decides to introduce an object by showing it in a picture and by annotating the corresponding picture part with the name of the object. Let's further assume that some time later it plans to introduce a part of this object in the same way. The graphics designer, however, doesn't generate a new picture since it recognizes that both tasks can be accomplished with a single picture.[7] The planner registers this by linking the corresponding parts of the generated DAG with each other.

Structure Sharing

In the example above, parts of the generated output have been used for different purposes (as background and as part of a label). However, it might also happen that not only the output, but even a more complex part of the DAG can be shared. E.g., assume the presentation planner decides to enable the user to carry out an action by creating two pictures showing the action and its result. To orientate the user, it is planned to show background objects in both pictures (cf. Fig. 5a). If the graphics designer is able to convey the requested information in a single picture, the background for the actions has to be included only once. Consequently, the structure of the document can be simplified by factoring out the background branch (cf. Fig. 5b).

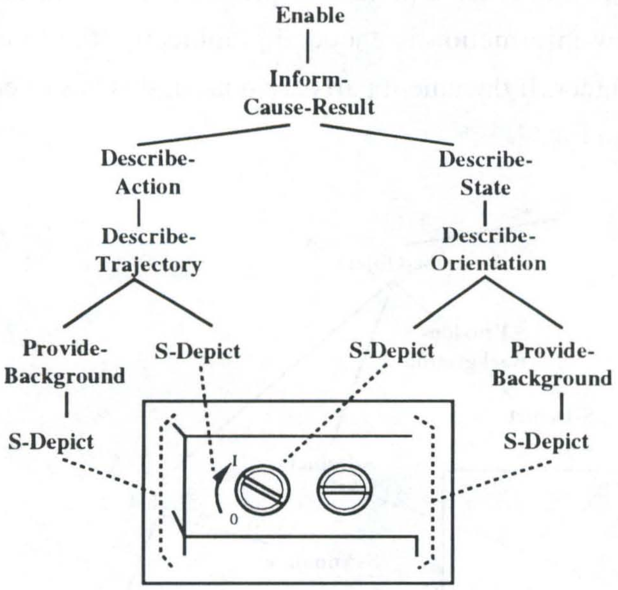

Fig. 5a: DAG without Structure Sharing

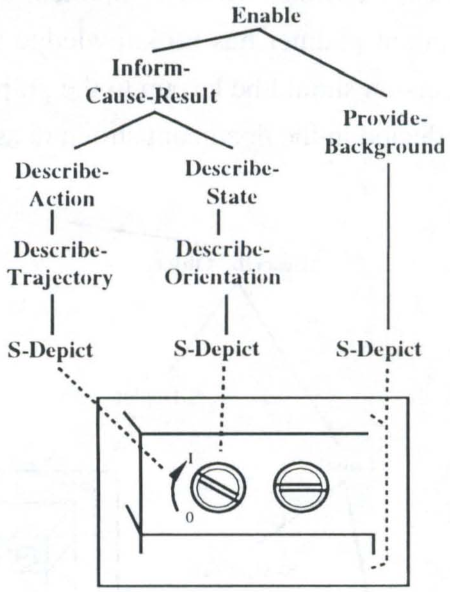

Fig. 5b: DAG after Simplification

Structure Adding

Whereas structure sharing leads to simplifications of the initial document plan, structure adding results in a more complex plan. It occurs if the graphics generator is expected to integrate information in a single picture, but is not able to do so. Let's suppose the planner decides to show the state of the espresso machine in the picture after it has been switched on. Thus, the

[7] This is possible because, during the generation process, the graphics designer builds up an explicit representation of the surface aspects of a picture as well as the semantic mapping between graphical means and the information to be conveyed (for details see [Rist/André 92b]).

graphics designer receives the task of generating a picture showing the current orientation of the on/off switch and the lamp in a burning state. When executing this task, the graphics designer realizes that the labels to the left of the on/off switch are too tiny to be readable if the entire espresso machine is to be shown (cf. Fig. 6a).

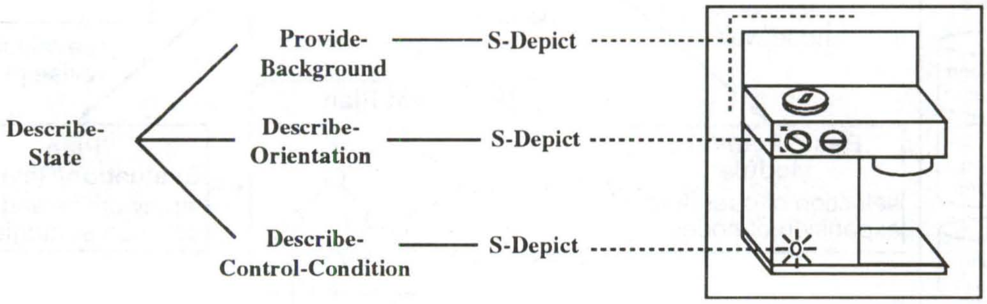

Fig. 6a: Initial DAG

If the graphics designer decides to overcome this problem by creating an inset with a different background, the structure of the document has to be modified as in Fig. 6b.

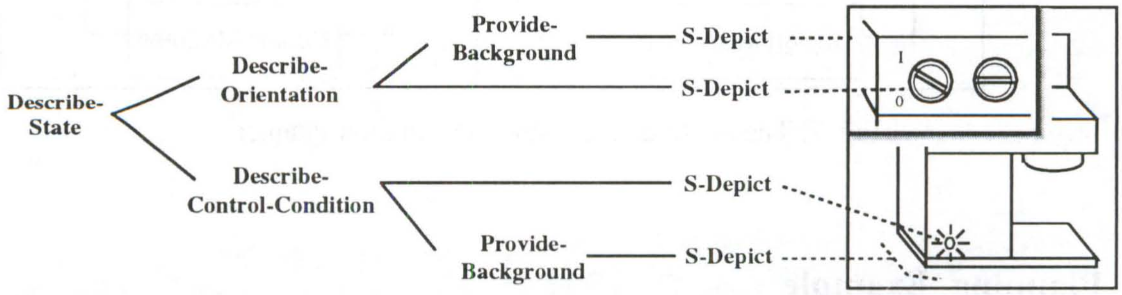

Fig. 6b: DAG after Structure Adding

5.5 Architecture of the Presentation Planner

The considerations above led to the architecture for the presentation planner shown in Fig. 7. The basic planning module selects operators that match the presentation goal and expands the nodes to generate a refinement-style plan in the form of a DAG. The plan evaluation/revision module applies critics and revision strategies. To allow for alternating revision and expansion processes, WIP's presentation planner is controlled by a plan monitor that determines the next action and the next nodes to be expanded. All components of the presentation planner have read/write access to the document plan.

In the overall WIP system (cf. [André et al., this volume]), the presentation planner collaborates with a text generator (cf. [Harbusch et al. 91]), a graphics generator (cf. [Rist/André 92a]) and a layout manager (cf. [Graf 92]). As shown in Fig. 7, the leaves of the document plan are connected to entries in the task queues of the mode-specific generators.

Thus, the document plan serves not only as an interface between the planner and the generators, but also enables a two-way exchange of information between the two generators.

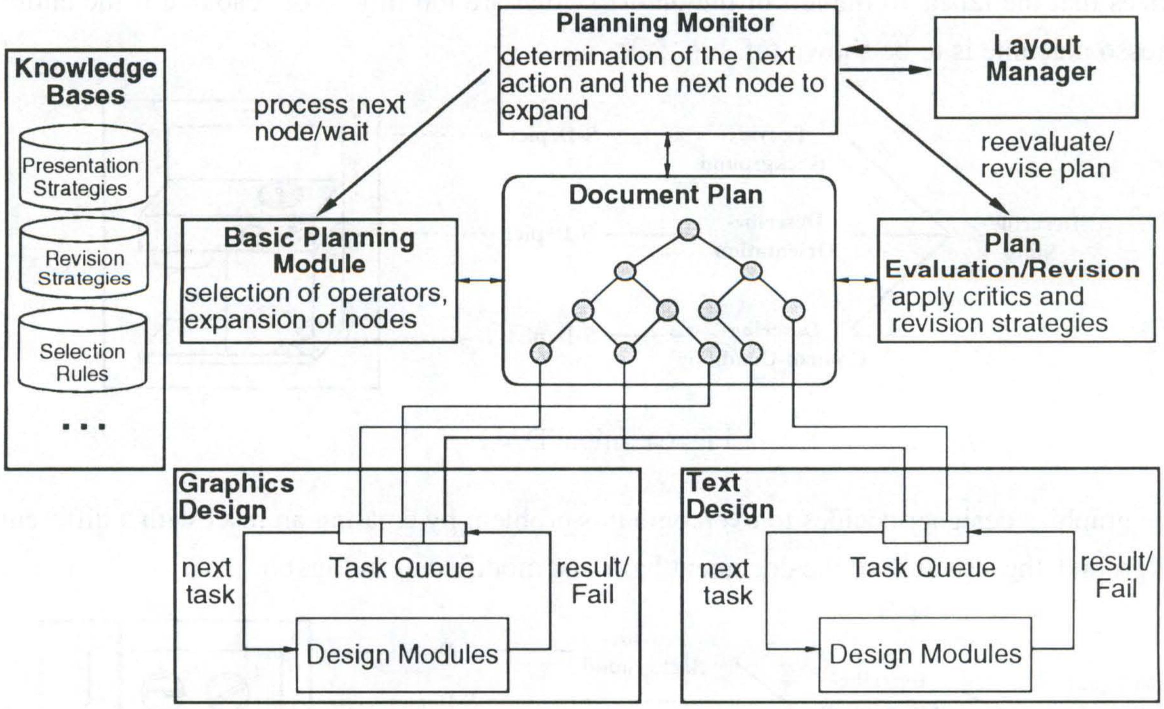

Fig. 7: The Architecture of the Presentation Planner

6 Planning Example

In the following, we give an example that illustrates opportunistic node expansion and revision after graphics generation. Assume the system as the presenter P wants the addressee A to switch on an espressomachine. Thus, it attempts to find plan operators which match the goal:

[1] (BMB P A (Goal P (Done A switch-on-1))).

One plan operator for achieving this goal was shown in section 4. Suppose this plan operator is selected. Then, the main and subsidiary acts are posted as subgoals. In this operator, three kinds of acts occur: two complex communicative acts (Enable and Motivate) which must be further expanded, an elementary speech act (S-Request) which is passed on to the text designer, and several referential acts (Activate) for filling the semantic case roles associated with the 'switch on' action. Assume that the user knows why the action should be carried out. Thus, it is not necessary to motivate him. The expansion of the 'Enable' act leads to a strategy that informs the user via a picture about the trajectory of the object to be manipulated and the result of the manipulation. After further refinement steps, the following subgoals are posted:

[2] (S-Depict P A (Trajectory trajectory-1) ?p-traj ?pic)

[3] (Achieve P (BMB P A (Identifiable A switch-2 ?px ?pic)) ?mode)

[4] (Describe P A (State state-1) ?mode)

At this point the plan monitor has to decide which of these three goals to expand next, so it inspects each one in turn. The first subgoal is an elementary act which is forwarded to the graphics designer. The second represents an intentional goal which is only expanded if it is not yet satisfied. Therefore, the presentation planner requests the graphics designer to evaluate:

[5] (BMB P A (Identifiable A switch-2 ?px ?pic))

For the purpose of this example, assume that the graphics designer has not yet executed [2] and thus is not able to immediately respond to [5]. As a consequence, the presentation planner cannot refine [3]. Instead of waiting for the response, the presentation planner tries to continue with another goal. It expands [4] and posts

[6] (S-Depict P A (Orientation orientation-1) ?p-orientation ?pic) and

[7] (Achieve P (BMB P A (Identifiable A switch-2 ?px ?pic)) ?mode)

as new subgoals. The first subgoal is passed on to the graphics designer. The presentation plan generated so far is shown in Fig. 8.

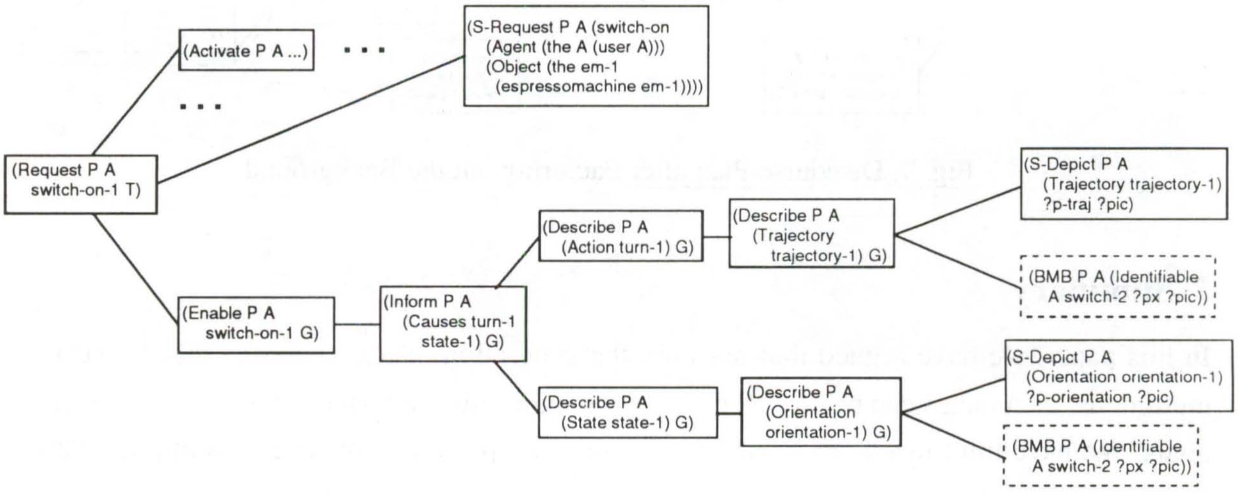

Fig. 8: Initial Discourse Plan

Note that at this time the mode variable occurring in [4] has already been instantiated by bottom-up propagation ot the mode in the header of strategy [S2]. When trying to satisfy the pictorial acts [2] and [6], the graphics designer finds out that it is possible to accomplish these tasks by means of a single picture (namely pic-4). After the goals in [3] and [7] have been instantiated, the planner recognizes that they are identical and that they can be achieved with a shared discourse plan. The planner decides to simplify the discourse plan by factoring out the structures corresponding to the goals in [3] and [7]. After switch-2 has been depicted, the graphics designer is able to evaluate

[8] (BMB P A (Identifiable A switch-2 p-switch-2 pic-4))

where p-switch-2 is the depiction of switch-2 in the picture pic-4. Since the graphics generator assumes that it is unclear to the user which switch is shown, the presentation planner has to find and instantiate[8] a strategy to achieve [8]. Assume it decides to select strategy [S3] and sends the graphics designer the request to depict the espresso machine as a landmark object. The final discourse plan is shown in Fig. 9.

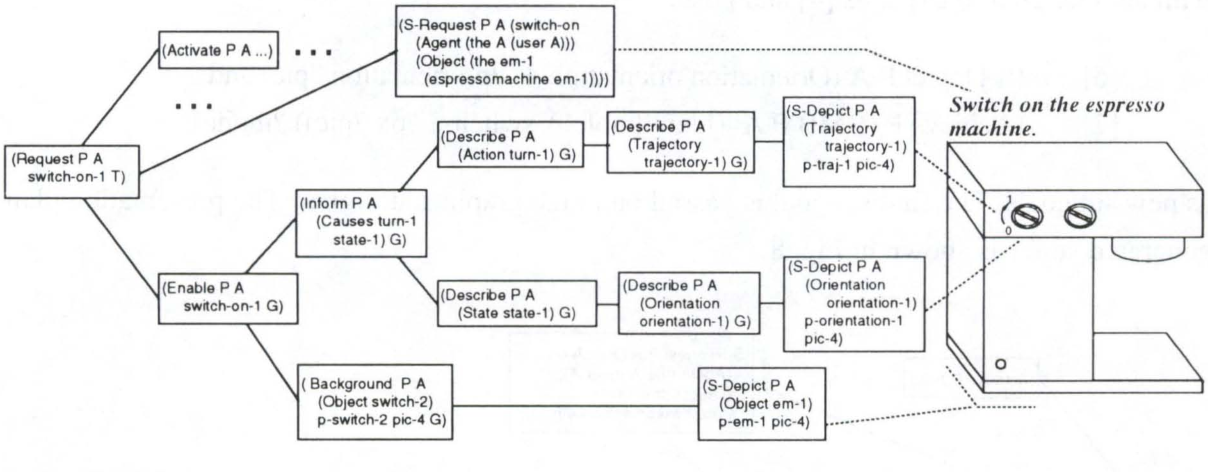

Fig. 9: Discourse Plan after Factoring out the Background

7 Summary

In this paper, we have argued that not only the generation of text, but also the synthesis of multimodal documents can be considered as a communicative act which aims to achieve certain goals. We have introduced presentation strategies to represent knowledge about presentation techniques. In order to decide between several presentation strategies, we have examined how the kind of information to be conveyed influences mode selection and which communicative functions single document parts play in text-picture combinations. In particular, we have argued that most semantic and pragmatic relationships which have been proposed for describing the

[8] Note that acts of the form (Achieve P <goal> <mode>) are treated specially. Whereas <goal> has to match the effect of a strategy, <mode> has to match the mode field in the header of a strategy.

structure of texts can be generalized in such a way that they are also appropriate for describing the structure of pictures and text-picture combinations.

For the realization of a system able to automatically generate illustrated documents, we have proposed a plan-based approach which supports data transfer between the content planner and the mode-specific generators and which allows for global plan evaluation after each plan step. A problem with modularizing presentation planning and mode-specific generation is that the results provided by the generators may deviate from the initial presentation plan. Since such deviations have to be reflected in the presentation plan, the planning scheme also comprises restructuring methods.

8 Implementation

The presentation planner has been implemented in Symbolics Common Lisp under Genera 8.0 running on a Symbolics XL1200 and MacIvory workstations. It has been integrated into the WIP system (cf. [André et al., this volume]). A stand-alone version of the planner is also available. It is embedded in a comfortable test-environment that includes an incremental plan displayer and provides various debugging facilities.

The planner is able to build up document structures as in the examples presented in this paper. However, in some examples we used graphics (e.g., the inset in Fig. 6b) that currently exceed the capacities of the implemented version of our graphics generator.

Acknowledgements:

The work presented here is supported by the German Ministry for Research and Technology (BMFT) under grant ITW8901 8. We would like to thank Wolfgang Wahlster, Andrew Csinger and the anonymous reviewers for their helpful comments.

References

[André/Rist 90a] E. **André** and T. **Rist**, Towards a Plan-Based Synthesis of Illustrated Documents, In: Proceedings of ECAI-90, Stockholm, Sweden, pp. 25-30, 1990.

[André/Rist 90b] E. **André** and T. **Rist**. Generating Illustrated Documents: A Plan-Based Approach. In: InfoJapan 90, Vol. 2, pp. 163-170, 1990.

[André et al. 92] E. **André**, W. **Finkler**, W. **Graf**, A. **Schauder**, T. **Rist**, W. **Wahlster**. WIP: The Automatic Synthesis of Multimodal Presentations, this volume.

[Arens et al. 92] Y. **Arens**, E. **Hovy** and M. **Vossers**. The Knowledge Underlying Multimedia Presentations, this volume.

[Badler et al. 91] N.I. **Badler**, B.L. **Webber**, J. **Kalita** and J. **Esakov**. Animation from Instructions. In: N.I. Badler, B.A. Barsky and D. Zeltzer (eds.), Making them Move: Mechanics, Control, and Animation of Articulated Figures, Morgan Kaufmann Publishers: San Mateo, pp. 51-93, 1991.

[Bandyopadhyay 90] S. **Bandyopadhyay**. Towards an Understanding of Coherence in Multimodal Discourse. Technical Memo DFKI-TM-90-01, Deutsches Forschungszentrum für Künstliche Intelligenz, Saarbrücken, 1990.

[Bieger/Glock 84] G.R. **Bieger** and M.D. **Glock**. The Information Content of Picture-Text Instructions. The Journal of Experimental Education 53(2), pp. 68-76, 1984.

[Bieger/Glock 86] G.R. **Bieger** and M.D. **Glock**. Comprehending Spatial and Contextual Information in Picture-Text Instructions. The Journal of Experimental Education 54(4), pp. 181-188, 1986.

[Brandt et al. 83] M. **Brandt**, W. **Koch**, W. **Motsch** and I. **Rosengren**. Der Einfluß der kommunikativen Strategie auf die Textstruktur - dargestellt am Beispiel des Geschäftsbriefes. In: I. Rosengren (Hrsg.), Sprache und Pragmatik, Lunder Symposium 1982. Almqvist & Wiksell: Stockholm, pp. 105-135, 1983.

[Van Dijk 80] T. A. **van Dijk**. Textwissenschaft. dtv: München, 1980.

[Feiner/McKeown 91] S.K. **Feiner** and K.R. **McKeown**. Automating the Generation of Coordinated Multimedia Explanations. In: IEEE Computer, 24(10), pp. 33-41, 1991.

[Graf 92] W. **Graf**. Constraint-Based Graphical Layout of Multimodal Presentations, to appear in: Proceedings of AVI-92, Rome, Italy, 1992.

[Grimes 75] J.E. **Grimes**. The Thread of Discourse. Mouton: The Hague, Paris, 1975.

[Harbusch et al. 91] K. **Harbusch**, W. **Finkler** and A. **Schauder**, Incremental Syntax Generation with Tree Adjoining Grammars, in: Proceedings Fourth International GI Congress, Munich, Germany, pp. 363-374, 1991.

[Hobbs 78] J. **Hobbs**. Why is a discourse coherent? Technical Report 176, SRI, Menlo Park, CA, 1978.

[Hovy 87] E. H. **Hovy**. Generating Natural Language under Pragmatic Constraints. PhD thesis, Department of Computer Science, Yale University, New Haven, CT., 1987.

[Hovy 88] E. H. **Hovy**. Planning Coherent Multisentential Text. In: Proc. of the 26th Annual Meeting of the ACL, New York, 1988.

[Hovy 90] E. H. **Hovy**. Unresolved Issues in Paragraph Planning, In: R. Dale, C. Mellish and M. Zock (eds.), Current Research in Natural Language Generation, Academic Press, pp. 17-41, 1990.

[Hunter et al. 87] B. **Hunter**, A. **Crismore** and P.D. **Pearson**. Visual Displays in Basal Readers and Social Studies Textbooks. In: D.M. Willows and H. A. Houghton (eds.), The Psychology of Illustration, Basic Research, Vol. 2, Springer: New York, Berlin, Heidelberg, London, Paris, Tokyo, pp. 116-135, 1987.

[Kjorup 78] S. **Kjorup**. Pictorial Speech Acts. In: Erkenntnis 12, pp. 55-71, 1978.

[Levin et al. 87] J.R. **Levin**, G.J. **Anglin** and R. N. **Carney**. On Empirically Validating Functions of Pictures in Prose. In: D.M. Willows and H. A. Houghton (eds.), The Psychology of Illustration, Basic Research, Vol. 1, Springer: New York, Berlin, Heidelberg, London, Paris, Tokyo, pp. 51-85, 1987.

[Mann/Thompson 87] W.C. **Mann** and S.A. **Thompson**. Rhetorical Structure Theory: Description and Construction of Text Structures. In: G. Kempen (eds.), Natural Language Generation: New Results in Artificial Intelligence, Psychology, and Linguistics, Nijhoff: Dordrecht, Boston, Lancaster, pp. 85-95, 1987.

[Marks/Reiter 90] J. **Marks** and E. **Reiter**. Avoiding Unwanted Conversational Implicatures in Text and Graphics. In: 8th AAAI, pp. 450-455, 1990.

[Maybury 92] M. **Maybury**. Planning Multimedia Explanations Using Communicative Acts, this volume.

[Molitor et al. 89] S. **Molitor**, S.-P. **Ballstaedt** and H. **Mandl**. Problems in Knowledge Acquisition from Text and Pictures. In: H. Mandl and J.R. Levin (eds.): Knowledge Acquisition from Text and Pictures. North Holland: Amsterdam, New York, Oxford, Tokyo, pp. 3-35, 1989.

[Moore/Paris 89] J.D. **Moore** and C.L. **Paris**. Planning Text for Advisory Dialogues. In: Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics, 1989.

[Moore/Swartout 89] J.D. **Moore** and W.R. **Swartout**. A Reactive Approach to Explanation. In: Proceedings of the 11th International Joint Conference on Artificial Intelligence, 1989.

[Muckenhaupt 86] M. **Muckenhaupt**. Text und Bild. Gunter Narr: Tübingen, 1986.

[Novitz 77] D. **Novitz**. Pictures and their Use in Communication: A Philosophical Essay. The Hague, 1977.

[Reithinger 91] N. **Reithinger**. POPEL - An Incremental and Parallel Natural Language Generation System. In: C.L. Paris, W.R. Swartout, and W.C. Mann (eds.), Natural Language Generation in Artificial Intelligence and Computational Linguistics, pp. 179-200, Kluwer: Norwell, 1991.

[Rist/André 92a] T. **Rist** and E. **André**. From Presentation Tasks to Pictures: Towards a Computational Approach to Graphics Design, to appear in: Proceedings of ECAI-92, Vienna, Austria, 1992.

[Rist/André 92b] T. **Rist** and E. **André**. Incorporating Graphics Design and Realization into the Multimodal Presentation System WIP, to appear in: Proceedings of AVI-92, Rome, Italy, 1992.

[Roth et al. 91] S.F. **Roth**, J. **Mattis** and X. **Mesnard**. Graphics and Natural Language as Components of Automatic Explanation. In: J. Sullivan and S. Tyler (eds.), Intelligent User Interfaces: Elements and Prototypes, Addison-Wesley, pp. 207-239, 1991.

[Roth/Mattis 90] S.F. **Roth** and J. **Mattis**. Data Characterization for Intelligent Graphics Presentation. In: Proc. of CHI 90, 1990.

[Searle 69] J.R. **Searle**. Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press: Cambridge, MA, 1969.

[Smith/Smith 66] K.U. **Smith** and M.F. **Smith**. Cybernetic Principles of Learning and Educational Design. Holt, Rinehart & Winston: New York, 1966.

[Strothotte/Schmid 90] T. **Strothotte** and C. **Schmid**. Semiformale Darstellungen in wissensbasierten Systemen. In: K. Kansy and P. Wißkirchen (eds.), Graphik und KI, GI-Fachgespräch, Springer: Berlin, Heidelberg, pp. 1-9, 1990.

[Wahlster et al. 91] W. **Wahlster**, E. **André**, W. **Graf**, and T. **Rist**. Designing Illustrated Texts: How Language Production Is Influenced by Graphics Generation. In: Proc. of the 5th Conference of the European Chapter of the ACL, Berlin: Springer, pp. 8-14, 1991.

[Whittaker/Walker 91] S. **Whittaker** and M. **Walker**. Towards a Theory of Multimodal Interaction. In: Proc. of the AAAI-91 Workshop on Intelligent Multimedia Interfaces, pp. 78-85, 1991.

Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

DFKI
-Bibliothek-
PF 2080
D-6750 Kaiserslautern
FRG

DFKI Publikationen

Die folgenden DFKI Veröffentlichungen sowie die aktuelle Liste von allen bisher erschienenen Publikationen können von der oben angegebenen Adresse bezogen werden.
Die Berichte werden, wenn nicht anders gekennzeichnet, kostenlos abgegeben.

DFKI Publications

The following DFKI publications or the list of all publisched papers so far can be ordered from the above address.
The reports are distributed free of charge except if otherwise indicated.

DFKI Research Reports

RR-91-28
Rolf Backofen, Harald Trost, Hans Uszkoreit:
Linking Typed Feature Formalisms and Terminological Knowledge Representation Languages in Natural Language Front-Ends
11 pages

RR-91-29
Hans Uszkoreit: Strategies for Adding Control Information to Declarative Grammars
17 pages

RR-91-30
Dan Flickinger, John Nerbonne:
Inheritance and Complementation: A Case Study of *Easy* Adjectives and Related Nouns
39 pages

RR-91-31
H.-U. Krieger, J. Nerbonne:
Feature-Based Inheritance Networks for Computational Lexicons
11 pages

RR-91-32
Rolf Backofen, Lutz Euler, Günther Görz:
Towards the Integration of Functions, Relations and Types in an AI Programming Language
14 pages

RR-91-33
Franz Baader, Klaus Schulz:
Unification in the Union of Disjoint Equational Theories: Combining Decision Procedures
33 pages

RR-91-34
Bernhard Nebel, Christer Bäckström:
On the Computational Complexity of Temporal Projection and some related Problems
35 pages

RR-91-35
Winfried Graf, Wolfgang Maaß: Constraint-basierte Verarbeitung graphischen Wissens
14 Seiten

RR-92-01
Werner Nutt: Unification in Monoidal Theories is Solving Linear Equations over Semirings
57 pages

RR-92-02
Andreas Dengel, Rainer Bleisinger, Rainer Hoch, Frank Hönes, Frank Fein, Michael Malburg:
Π_{ODA}: The Paper Interface to ODA
53 pages

RR-92-03
Harold Boley:
Extended Logic-plus-Functional Programming
28 pages

RR-92-04
John Nerbonne: Feature-Based Lexicons:
An Example and a Comparison to DATR
15 pages

RR-92-05
Ansgar Bernardi, Christoph Klauck, Ralf Legleitner, Michael Schulte, Rainer Stark:
Feature based Integration of CAD and CAPP
19 pages

RR-92-06
Achim Schupetea: Main Topics of DAI: A Review
38 pages

RR-92-07
Michael Beetz:
Decision-theoretic Transformational Planning
22 pages

RR-92-08
Gabriele Merziger: Approaches to Abductive Reasoning - An Overview -
46 pages

RR-92-09
Winfried Graf, Markus A. Thies:
Perspektiven zur Kombination von automatischem Animationsdesign und planbasierter Hilfe
15 Seiten

RR-92-10
M. Bauer: An Interval-based Temporal Logic in a Multivalued Setting
17 pages

RR-92-11
Susane Biundo, Dietmar Dengler, Jana Koehler: Deductive Planning and Plan Reuse in a Command Language Environment
13 pages

RR-92-13
Markus A. Thies, Frank Berger:
Planbasierte graphische Hilfe in objektorientierten Benutzungsoberflächen
13 Seiten

RR-92-14
Intelligent User Support in Graphical User Interfaces:
1. InCome: A System to Navigate through Interactions and Plans
Thomas Fehrle, Markus A. Thies
2. Plan-Based Graphical Help in Object-Oriented User Interfaces
Markus A. Thies, Frank Berger

22 pages

RR-92-15
Winfried Graf: Constraint-Based Graphical Layout of Multimodal Presentations
23 pages

RR-92-16
Jochen Heinsohn, Daniel Kudenko, Berhard Nebel, Hans-Jürgen Profitlich: An Empirical Analysis of Terminological Representation Systems
38 pages

RR-92-17
Hassan Aït-Kaci, Andreas Podelski, Gert Smolka: A Feature-based Constraint System for Logic Programming with Entailment
23 pages

RR-92-18
John Nerbonne: Constraint-Based Semantics
21 pages

RR-92-19
Ralf Legleitner, Ansgar Bernardi, Christoph Klauck
PIM: Planning In Manufacturing using Skeletal Plans and Features
17 pages

RR-92-20
John Nerbonne: Representing Grammar, Meaning and Knowledge
18 pages

RR-92-21
Jörg-Peter Mohren, Jürgen Müller
Representing Spatial Relations (Part II) -The Geometrical Approach
25 pages

RR-92-22
Jörg Würtz: Unifying Cycles
24 pages

RR-92-23
Gert Smolka, Ralf Treinen:
Records for Logic Programming
38 pages

RR-92-24
Gabriele Schmidt: Knowledge Acquisition from Text in a Complex Domain
20 pages

RR-92-25
Franz Schmalhofer, Ralf Bergmann, Otto Kühn, Gabriele Schmidt: Using integrated knowledge acquisition to prepare sophisticated expert plans for their re-use in novel situations
12 pages

RR-92-26
Franz Schmalhofer, Thomas Reinartz, Bidjan Tschaitschian: Intelligent documentation as a catalyst for developing cooperative knowledge-based systems
16 pages

RR-92-27
Franz Schmalhofer, Jörg Thoben: The model-based construction of a case-oriented expert system
18 pages

RR-92-29
Zhaohur Wu, Ansgar Bernardi, Christoph Klauck:
Skeletel Plans Reuse: A Restricted Conceptual Graph Classification Approach
13 pages

RR-92-33
Franz Baader:
Unification Theory
22 pages

RR-92-34
Philipp Hanschke:
Terminological Reasoning and Partial Inductive Definitions
23 pages

RR-92-35
Manfred Meyer:
Using Hierarchical Constraint Satisfaction for Lathe-Tool Selection in a CIM Environment
18 pages

RR-92-36
Franz Baader, Philipp Hanschke:
Extensions of Concept Languages for a Mechanical Engineering Application
15 pages

RR-92-37
Philipp Hanschke:
Specifying Role Interaction in Concept Languages
26 pages

RR-92-38
Philipp Hanschke, Manfred Meyer:
An Alternative to Θ-Subsumption Based on Terminological Reasoning
9 pages

RR-92-43
Christoph Klauck, Jakob Mauss: A Heuristic driven Parser for Attributed Node Labeled Graph Grammars and its Application to Feature Recognition in CIM
17 pages

RR-92-44
Thomas Rist, Elisabeth André: Incorporating Graphics Design and Realization into the Multimodal Presentation System WIP
15 pages

RR-92-45
Elisabeth André, Thomas Rist: The Design of Illustrated Documents as a Planning Task
21 pages

RR-92-46
Elisabeth André, Wolfgang Finkler, Winfried Graf, Thomas Rist, Anne Schauder, Wolfgang Wahlster: WIP: The Automatic Synthesis of Multimodal Presentations
19 pages

DFKI Technical Memos

TM-91-11
Peter Wazinski: Generating Spatial Descriptions for Cross-modal References
21 pages

TM-91-12
Klaus Becker, Christoph Klauck, Johannes Schwagereit: FEAT-PATR: Eine Erweiterung des D-PATR zur Feature-Erkennung in CAD/CAM
33 Seiten

TM-91-13
Knut Hinkelmann:
Forward Logic Evaluation: Developing a Compiler from a Partially Evaluated Meta Interpreter
16 pages

TM-91-14
Rainer Bleisinger, Rainer Hoch, Andreas Dengel:
ODA-based modeling for document analysis
14 pages

TM-91-15
Stefan Bussmann: Prototypical Concept Formation
An Alternative Approach to Knowledge Representation
28 pages

TM-92-01
Lijuan Zhang:
Entwurf und Implementierung eines Compilers zur Transformation von Werkstückrepräsentationen
34 Seiten

TM-92-02
Achim Schupeta: Organizing Communication and Introspection in a Multi-Agent Blocksworld
32 pages

TM-92-03
Mona Singh
A Cognitiv Analysis of Event Structure
21 pages

TM-92-04
Jürgen Müller, Jörg Müller, Markus Pischel, Ralf Scheidhauer:
On the Representation of Temporal Knowledge
61 pages

TM-92-05
Franz Schmalhofer, Christoph Globig, Jörg Thoben
The refitting of plans by a human expert
10 pages

TM-92-06
Otto Kühn, Franz Schmalhofer: Hierarchical skeletal plan refinement: Task- and inference structures
14 pages

TM-92-08
Anne Kilger: Realization of Tree Adjoining Grammars with Unification
27pages

DFKI Documents

D-92-02
Wolfgang Maaß: Constraint-basierte Plazierung in multimodalen Dokumenten am Beispiel des Layout-Managers in WIP
111 Seiten

D-92-03
Wolfgan Maaß, Thomas Schiffmann, Dudung Soetopo, Winfried Graf: LAYLAB: Ein System zur automatischen Plazierung von Text-Bild-Kombinationen in multimodalen Dokumenten
41 Seiten

D-92-04
Judith Klein, Ludwig Dickmann: DiTo-Datenbank - Datendokumentation zu Verbrektion und Koordination
55 Seiten

D-92-06
Hans Werner Höper: Systematik zur Beschreibung von Werkstücken in der Terminologie der Featuresprache
392 Seiten

D-92-07
Susanne Biundo, Franz Schmalhofer (Eds.): Proceedings of the DFKI Workshop on Planning
65 pages

D-92-08
Jochen Heinsohn, Bernhard Hollunder (Eds.): DFKI Workshop on Taxonomic Reasoning Proceedings
56 pages

D-92-09
Gernod P. Laufkötter: Implementierungsmöglichkeiten der integrativen Wissensakquisitionsmethode des ARC-TEC-Projektes
86 Seiten

D-92-10
Jakob Mauss: Ein heuristisch gesteuerter Chart-Parser für attributierte Graph-Grammatiken
87 Seiten

D-92-11
Kerstin Becker: Möglichkeiten der Wissensmodellierung für technische Diagnose-Expertensysteme
92 Seiten

D-92-12
Otto Kühn, Franz Schmalhofer, Gabriele Schmidt: Integrated Knowledge Acquisition for Lathe Production Planning: a Picture Gallery (Integrierte Wissensakquisition zur Fertigungsplanung für Drehteile: eine Bildergalerie)
27 pages

D-92-13
Holger Peine: An Investigation of the Applicability of Terminological Reasoning to Application-Independent Software-Analysis
55 pages

D-92-14
Johannes Schwagereit: Integration von Graph-Grammatiken und Taxonomien zur Repräsentation von Features in CIM
98 Seiten

D-92-15
DFKI Wissenschaftlich-Technischer Jahresbericht 1991
130 Seiten

D-92-16
Judith Engelkamp (Hrsg.): Verzeichnis von Softwarekomponenten für natürlichsprachliche Systeme
189 Seiten

D-92-17
Elisabeth André, Robin Cohen, Winfried Graf, Bob Kass, Cécile Paris, Wolfgang Wahlster (Eds.): UM92: Third International Workshop on User Modeling, Proceedings
254 pages
Note: This document is available only for a nominal charge of 25 DM (or 15 US-$).

D-92-18
Klaus Becker: Verfahren der automatisierten Diagnose technischer Systeme
109 Seiten

D-92-19
Stefan Dittrich, Rainer Hoch: Automatische, Deskriptor-basierte Unterstützung der Dokumentanalyse zur Fokussierung und Klassifizierung von Geschäftsbriefen
107 Seiten

D-92-21
Anne Schauder: Incremental Syntactic Generation of Natural Language with Tree Adjoining Grammars
57 pages

D-92-26
Enno Tolzmann:
Realisierung eines Werkzeugauswahlmoduls mit Hilfe des Constraint-Systems CONTAX
28 Seiten

D-92-25
Martin Buchheit: Klassische Kommunikations- und Koordinationsmodelle
31 Seiten

The Design of Illustrated Documents as a Planning Task

Elisabeth André, Thomas Rist

RR-92-45

Research Report