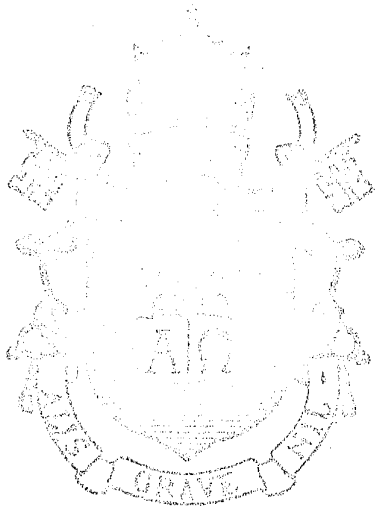


PUC

Series: Monografias em Ciência da Computação
Nº13/88

CONCILIATORY PLANNING FOR EXTENDED DESCRIPTIVE TEXTS

Donia R. Scott
Clarisse S. Souza

Departamento de Informática

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
MARQUÊS DE SÃO VICENTE, 225 — CEP 22453
RIO DE JANEIRO — BRASIL

PUC/RJ - Departamento de Informática

Series: Monografias em Ciência da Computação, nº 15/88
Editor: Paulo A. S. Veloso dezembro, 1988

CONCILIATORY PLANNING FOR EXTENDED DESCRIPTIVE TEXTS*

Donia R. Scott**
Clarisse S. Souza

* Work partially sponsored by MCT.
** Visiting professor Philips Research Laboratory, U.K.

In charge of publications:
PUC/RJ - Depto. de Informática
Rosane T. L. Castilho
Assessoria de Biblioteca, Documentação e Informação
Rua Marquês de São Vicente, 225 - Gávea
22453 Rio de Janeiro, RJ
BRASIL

ABSTRACT

Good text requires good planning. This is even more crucial when the text is long and therefore more sensitive to stylistic blunders that can affect cognitive processing. Descriptions of complex data objects are an instance of such a text.

GEMA is a language generation system for producing descriptions of complex data objects. These descriptions will usually contain several paragraphs. Texts of this length demand more stylistic harmonization than is typically required of text generators, and producing them commands a great deal of interaction between planning and realizing components of the generator. This brings with it some computational problems that are not very well handled by existing planning approaches.

We describe in this paper a planning approach that has been developed for GEMA that combines some of the more desireable features of traditional text-planning approaches and which achieves good stylistic effects. We have called this conciliatory planning.

CONCILIATORY PLANNING FOR EXTENDED DESCRIPTIVE TEXTS

Scott, Donia R.(1,2) and Souza, Clarisse S.(2)

(1) Philips Research Laboratory,
Cross Oak Lane,
Redhill, RH1 5HA,
U.K.

(2) Departamento de Informática,
Pontifícia Universidade Católica do Rio de Janeiro,
Rua Marquês de São Vicente 225,
22453 Rio de Janeiro,
Brasil.

1. INTRODUCTION

Description of complex objects present several challenges to text generation systems. The quality and quantity of the information contained in the description, along with the structure within which it is conveyed, should contribute to the case with which the reader is able to construct an accurate and appropriate mental model of the object being described. This means that the text must necessarily take account of psycho-linguistic factors.

GEMA is a natural language generation program for producing textual descriptions of complex systems in Portuguese. A graphic description of the type of systems GEMA must describe textually is shown in figure 1. Both descriptions are representations given to the user of the machine-internal model of the system (the data model). They differ in that the textual descriptions are very much more complicated to produce since they must include the semantics of the model (without, however, aiming at descriptions of the real world situation that it partially represents). Semantic information is not available for representation in the graphic medium.

Text of the type with which we are dealing tend to be very much longer than that typically produced by most existing language generation systems. At the very least, GEMA must

generate more paragraphs than the number of processes contained in the data model. This factor, along with the complex nature of the models, puts a heavy load on the planning of the text.

Given GEMA's communicative goal to produce a text which describes data models in a manner which maximizes the ease with which the reader can construct accurate mental models of them, and a set of heuristics for determining what factors will contribute to this goal, its task is to decide:

- what slant to put on the text, ie. from the point-of-view of which external entity should the model be presented
- what information from the model should be included in the text
- in what order should the information be presented
- what aspects of the model should be focussed
- what should be the structure of the text
- what should be the surface linguistic form of the text, ie. decisions about lexical, syntactic and stylistic forms.

Existing text generators fall into three basic classes depending on the nature in which the planning and realizing decisions are made: sequential, integrated and mixed approaches. In the sequential approach (McDonald and Pustejovsky,1985; McKeown,1985; Paris,1987), all decisions about text content and structure are made before decisions about linguistic form (ie. all planning occurs before realization begins). In the integrated approach (Appelt,1983,1985), these decisions are made concurrently. The mixed approach, as its name implies, encompasses both approaches. An instance of a mixed planner is PAULINE (Hovy,1988) where planning and realization are interleaved.

The relative merits of the three alternative approaches tend to be bound to computational factors (such as the amount and cost of back-tracking that may be required before a reasonable text is generated), and not to theoretical motivations. The type of text to be produced will often motivate the choice of architecture. For example, a sequential approach is unsuitable for texts which include goals that are realized not through the type of information conveyed but through the way in which it is conveyed (e.g. perlocutionary goals such as 'impress the reader').

The production of descriptive texts in GEMA contemplates the interplay between cognitive goals and linguistic structures: information should not be embedded in complicated sentential configurations or conveyed through the use of lexically ambiguous items. Clearly, some combination of the sequential and integrated approaches is required. Trying to get the best of both worlds, GEMA adopts a mixed approach with the same global

objectives of Hovy's interleaved model and with a 'situated action' planning philosophy suggested by Suchman (1987). GEMA's planner architecture departs from Hovy's, however, in that it uses a traditional two-step sequential model, obtaining combination of planning and realizing via knowledge resources that are available to both sub-components.

GEMA is tuned to account for stylistic rules such as the avoidance of close repetition of words and marked syntactic constructions. These rules are even more important for text in Brazilian Portuguese, the language GEMA produces, than they are for English. In addition, as an inflected language, the use of ellipsis for subject NP's has to be carefully planned. For texts of the length that GEMA is required to produce (on average, 7 paragraphs), integrated cognitive and stylistic planning is therefore of crucial importance.

2. PLANNING ALTERNATIVES FOR THE DESCRIPTION OF COMPLEX DATA-OBJECTS

It is clear that a strictly sequential approach is unsuitable to the needs of GEMA owing to the considerable difficulty with which it can handle stylistic factors. As an example, suppose that the planner decides, for cognitive reasons to do with the importance of entities that receive data, that the following information is to be realized in the following order:

- Process1 sends A to B
- Process1 deletes D from C
- Process1 inserts A in C

A sequential approach will result in a description of the type:

"Process1 sends A to B. It also deletes D from C and inserts A in it"

The style of this text is clearly poor and little can be done to improve on it since the model does not allow for negotiation between sub-components. The opportunity to generate something like:

"Process1 sends A to B and also inserts it in C, whereas it deletes D from C"

the structure of which, for reasons of achieving stylic effects only partially respects the priorities, will not arise. The cognitive superiority of the latter text over the former is evident: psycholinguistically speaking, the syntactic activity of coordination produces more easily processable text when it is applied to direct objects than when applied to indirect ones.

Similarly, coordination among NP's is more easily processable than among VP's (Frazier et. al., 1984).

On the other hand, an integrated approach is prepared to handle the above conflict of interest, since the opportunity to produce the more desireable type of text will be considered during the planning process. But this will be done at the cost of a tight control of syntactic and semantic constraints. In addition, the amount of knowledge that the generator will have to carry along in its operation and the number of computations that will need to be performed will be expectedly large. These are the main disadvantages of a fully integrated planner. Given the heavy load that is put on planning in GEMA by these and by stylistic factors, an integrated planning model too will be unsuitable.

Hovy's proposal is one way of combining the computational efficiency of a sequential approach with the cognitively desireable effects of an integrated approach. His choice of an interleaved model is essentially motivated by pragmatic considerations, since the narratives to be produced by PAULINE must be slanted towards sympathy with one of the parties involved in a particular event. This means that there will often be conflicting communicative goals to be achieved, frequently through the adoption of different lexical and syntactic choices. Thus, interleaving is achieved through the alternation of planning and realization sub-functions in an iterative fashion.

Given the nature of its generation task, GEMA is unlikely, if ever, to encounter conflicting communicative goals. Instead, the preeminent objective of generation is cognitive, and this is reflected in every decision that is made. For this reason, identification of sub-functions for purposes of resolving competing communicative goals would be unnatural for the application. Conflicts at a lower level may, however, occur and these will call for some integration between planning and realization. This we achieve by what we call conciliatory planning.

3. CONCILIATORY PLANNING

GEMA adopts a decision-making strategy for planning and realization activities that draws on combined knowledge. The planner carries out its tasks in view of general syntactic knowledge, and the realizer produces the final text from a basic structure that reflects the minimal syntactic structures the planner has operated on. The general schema of GEMA's planning and realization activities is shown in Figure 2.

The bridge between the two sub-components of GEMA is represented by two basic knowledge representations: paragraph

structures and kernel sentences. Paragraph structures are selected from among six alternatives, depending on the type and quantity of information to be conveyed. These structures account for matters of focus and sentence constituency, which are in direct connection with syntax. Final syntactic features can be predicted by paragraph structures and thus planning is provided with some sort of look-ahead capacity. This is discussed in greater detail in Section 4.

Kernel sentences are a low-level realization of predicates used to represent the original data-flow model. They are the result of a unification grammar of lexical origin that provides for the semantic integrity of text and for basic stylic opportunities (see Section 5). Kernel sentences are a reflection of the semantic structure the planner has produced and, since the final text is the result of a series of meaning-preserving transformations applied to them, such sentences are a constant look-back resource available to the realizer.

Through the maintenance of a basic sequential architecture and the introduction of combined knowledge at different stages of generation, GEMA is able to achieve the primary goals of the interleaved model. Top-down planning is accounted for by its look-ahead capacity, while bottom-up planning is achieved through its look-back capacity. The overall effect is to minimize back-tracking, the main objective of interleaving.

A potential disadvantage of conciliation compared to full integration or interleaving may be that of fine lexical style. As mentioned before, the production of synonyms is especially important for good Brazilian Portuguese text since the language is particularly averse to lexical repetition within what can be fairly large sections of text (of a paragraph or greater). Synonyms, however, may not have the same syntactic sub-specifications and this factor may call for some restructuring of sentences. The impact of such changes is difficult to control since there is no guarantee that only inter-sentential alternations will be required.

The solution we have chosen for this problem in GEMA has been one of simplification by restricting synonyms to those which share both semantic and syntactic sub-specifications, thereby reducing the issue of producing synonyms to one of mere replacement. The result of this decision is a strong lexical constraint and a reduced scope of choice at final stages of text production. Although not ideal, the computation payoff in terms of reduced back-tracking is great, and the final effect is a text that is not unpleasing.

4. THE OPERATION OF THE CONCILIATORY PLANNER

In this section we describe in more detail the operation of the conciliatory planner in GEMA and the manner in which planning and realizing activities interact. As mentioned above, the interactions are aided by our choice of representation schema for the data model, which reflects syntactic and semantic properties. It is this knowledge that the planner acts upon and which it reproduces in an augmented form for the realizer.

Propositions in the data model follow the form:

1. ACTION-TYPE,AGENT,INPUT,OUTPUT
2. ACTION-TYPE,AGENT,ACTION-OBJECT,___

In linguistic terms, these correspond to:

1. VERB,SUBJECT,INDIRECT OBJECT,DIRECT OBJECT(1)
2. VERB,SUBJECT,DIRECT OBJECT, REST(2)

Action-types are primitives for the activies of processing, receiving and sending and are represented as do-acts, p-acts, r-acts and s-acts. As an example of what this will look like to the planner, the representation of Process 6 in Figure 1 will include the following:

do-act,produce,process6,reports,for the manager
r-act,consult,process6,payments,payments received
p-act,transform,process6,payments made,payments report
s-act,send,process,manager,payments report

This type of information is used by the planner to construct a tree that is a macro text structure. The leaves of the tree are the propositions to which tags indicating ordering priorities are added. The root of each major sub-tree is a process in the data model (represented as the do-act for which that process is AGENT) and the root of sub-sub-trees are one of the variables corresponding to action,input and output. All sub-trees are marked for focus, and focus can be nested within sub-trees.

---

1 The exception to this is processing activities that involve a simple passing-on of data. In this case the direct object is input.

2 Propositions of this form apply only to do-acts. These describe the general function of a process. They can be realized as, for example, "Produce reports",or "Produce reports for managers", or even "Produce reports for managers of associate branches". For purposes of generation, only the verb and direct object need be identified.

Furthermore, the structure generated for the lower portion of text-spanning trees are focus-sensitive.

GEMA's descriptions have a fixed general structure which is:

1. Introduce the model by saying what its main function is, what are the major activities involved in performing this function and what are the objects that play a role in the data-processing model.
2. Describe the activities of each of the processes in turn.

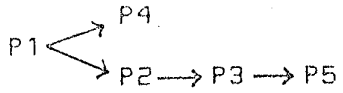
Given this schema, the first task of the planner is to determine what the main structure of the tree should be. This decision is unilaterally taken and is non-negotiable. It is driven by strictly non-linguistic considerations of the data model and of the intended reader. These affect decisions on what aspect of the activity of processes should be focussed, on the viewpoint from which to present the model and on the best path through the model for the chosen viewpoint. These decisions vary greatly in terms of the amount of work the planner must engage in.

Based on the fact that input to process temporally precede output from it, and that the data model can be viewed on an abstract level as the operation of one process which receives data from, and produces data for, external entities, GEMA assigns focus to the input data of each process in the model except the one that is discussed last. The last process always has focus on its output.

Decisions on viewpoint are even more trivial to compute. Requests to GEMA to produce descriptions always precipitate a request to the user to select from the external entities that with which he or she identifies or is most able to. This information (which we refer to as the identity of EE') is available to the planner and is used by it to select the viewpoint of the description: the data model will always be described through the "eyes" of EE'. This is as close as GEMA need get to what Hovy refers to as problems of text slant.

Choosing the best path through the model is considerably more difficult and is in fact the most time-consuming activity of GEMA. For each data model, the number of potential paths through it will be a factorial function of the number of processes it contains. Given GEMA's cognitive goals, only a small number of these will result in a good description. The planner chooses the best path, not by selecting one among the entire set, but by generating only those paths which will satisfy the cognitive goals and by selecting from that set, the best candidate. Good paths are those which:

- start with a process who receives data from EE'
- end with a process who sends data to an external entity
- give priority to pairs of processes that communicate directly with each other over those who communicate only indirectly (via archives, external entities or other processes), which in turn are given priority over process pairs who do not communicate among themselves.
- respect the temporal primacy of input over output
- respect the need to minimize the reader's memory load related to mental-model construction. For example, the best presentation order for the connections



will be [P1 P4 P2 P3 P5].

The best path will be that which conforms most strongly to the above criteria. For the data model of SISVIDEO, the best path from the viewpoint of the client will be through the processes:

Register members or borrowers
Register films
Produce film catalogues
Process returns
Receive payments
Effect loans
Produce reports for the manager

One may well argue for a different ordering of processes, but it is worth re-emphasising that GEMA's task is to describe the data model and not its real-world equivalent. Its decisions are based solely on information that is available in the model.

The final decision of a best path is used to order the sub-trees in the text structure whose heads are individual processes . It is also used to add order-signalling tags the do-acts for processes in the introduction part of the text schema. These, unlike the ordering of the process sub-trees, can be overridden by the realizer for the purpose of achieving a linguistic effect.

The next step the planner takes is to append order-signalling tags to the p-,r- and s-acts associated with each process. These tags suggest the relative priority that should be given to each -act within its class and these too can be overridden by the realizer. The priority values given are sensitive to the context in which the process occurs and to the cognitively-related primacy of the objects that are the source and destination of data into and out of the process.

The last task of the planner is to assign the final levels of the tree that correspond to paragraph structure. This is achieved by (deterministically) selecting from among six alternatives, depending on the focus of the paragraph (input or output) and on the number of p- and r-acts associated with the process. The terminal elements of this structure will be the leaves of the text-spanning tree and are the augmented data-model propositions.

The tree produced by the planner contains all the information that will be required by the realizer to produce a text which is cohesive and which reflects the integrity of the data model.

5. FROM PLANNING TO REALIZING PORTUGUESE

The generation of basic sentences is achieved through the use of a unification grammar that operates on the lexicon. The semantics of each -act is the key for lexical search for the equivalent verb. Associated with each entry is a generative rule of the following type for deriving a kernel sentence:

```
semantics:      s-act1,process6,manager,report on payments

lexical entry: SEND
                  morphology:.............
                  semantics: sact1,AGENT,DESTINATION,OUTPUT
                  syntax:    AGENT send    OUTPUT  to-DESTINATION
                                 |    |       |           |

                             (subj  (verb  (dir-obj   (prep-comp
                              np     pres   np          pp
                              sing   ind    number      number)
                              gen    sing   gen)
                              3rd p) 3rd p)
                  constraints: ..............
```

The kernel sentence that is generated for the above proposition is:

"Process6(+features) send(+features) report on payments(+features) to manager(+features)"

Process6 in this case is only a pointer to a do-act and will be realized as either a nominalization, a pronoun or an equivalent expression. As the example shows, features are added to lexical items in the kernel sentence. These will be utilized in later stages of the realization process.

The planner's choice of paragraph structure will dictate which transformational rules should apply and when. This is the first important place of looking-back. The transformations that

can be applied at this stage (and which are selected by cognitive criteria) are coordination, nominalization, passivization, relativization, pronominalization, topicalization and ellipsis. Of these, coordination, relativization, pronominalization and ellipsis all involve the analysis of more than one kernel sentence. The preconditions for these transformations involve the matching of elements between kernel sentences. The conditions for the application of transformations are matched against patterns of pairs of kernel sentences and the type of pattern found determines how the rule must be applied.

Patterns are derived from 2x3 matrices of which there are 3, one each for p-, r-, and s-acts. These are shown in Figure 3. The parameters of the matrices are: verb vs. np vs. pp and identical vs. different entries. These matrices are interpreted by GEMA as in the following: ADE and ACF command simple coordinations (at NP level) whereas B__ commands complex coordination (at VP level). These patterns also indicate the choice of the coordinating conjunction: ACF signals the need for "or", whereas the others signal "and".

The consultation of matrices in a network fashion also signals other important linguistic opportunities. For example, an ACF pattern in the p-act matrix and an ADE pattern in the r-act matrix together indicate an opportunity for compound coordination and relativization, rendering text like:

"With data on payment and credit notifications, which come from the process for authorizing credit, the information-distributing process compiles a report for managers and supervisors"

instead of

"With data on payment, which come from the process for authorizing credit, and credit notifications, which also come from the same process, the information-distributing process compiles a report for managers and supervisors".

Matrices have proved to be an efficient decision-support tool for GEMA: not only do they provide considerable insight in terms of style (look-ahead), but they also allow GEMA to maintain integrity with the original knowledge structure via the parameters that are used (look-back).

The patterns that are found in the various matrices, however, provide only a general syntactic guideline. Fine-grained syntactic adjustments can only be achieved by examining lexical entry features. This is done by a morpho-syntactic sub-component of the realizer. A typical example of one such situation is that of prepositional compatibility in PP coordination, a situation that provides special problems for Portuguese since prepositions are gender-marked. Another case is

that of anaphoric realization through the use of either personal pronouns or some more complex referring expressions. This is illustrated by the following text:

"With data on payment and credit notifications, the process for distributing information compiles reports and sends them to managers and supervisors. With the same input, the process computes the number of associates per district and sends this information to the director."

Notice that boldface items refer to previously-mentioned elements. However, not all are realized as pronouns. Consecutive items are not anaphorized in the same way, and this decision is made in terms of how much potential ambiguity would be involved in the final text. This is essentially a matter of the global morphology of the closest NP at the final steps of linearization. Had GEMA used "them" instead of "the same input", the resulting text could lead the user to erroneously interpret the antecedent as "managers and supervisors" (which together make up a NP with compatible number and gender).

6. CONCLUSIONS

The conciliatory approach combines the advantages of computational efficiency of sequential planning and of text-quality of integrated planning. It is more efficient than an interleaved approach for the type of texts with which GEMA deals since it is sensitive to the peculiarities of the objects to be described and of the description language. This is achieved by giving the planner look-ahead, and the realizer look-back, capabilities.

The texts produced by conciliatory planning are usually much more readable than those produced by human writers, which tend to be difficult to comprehend (ie. difficult for the reader to visualize what the data model is like) and to suffer from glaring ommissions and additions.

7. BIBLIOGRAPHY

Appelt, D.G. (1983) "Telegram: A Grammar Formalism for Language Planning", Proceedings of the Eighth IJCAI Conference, 595-599.

Appelt, D.G. (1985) Planning English Sentences. Cambridge University Press, Cambridge.

Frazier, L., Taft, L., Roeper, T., Clifton, C., Ehrlich, K.(1984) "Parallel Structure: a Source of Facilitation in Sentence Comprehension", Memory and Cognition, 12, 421-430.

Hovy, E. (1988) "Two Types of Planning for Text Generation", Proceedings of the 26th Meeting of the ACL.

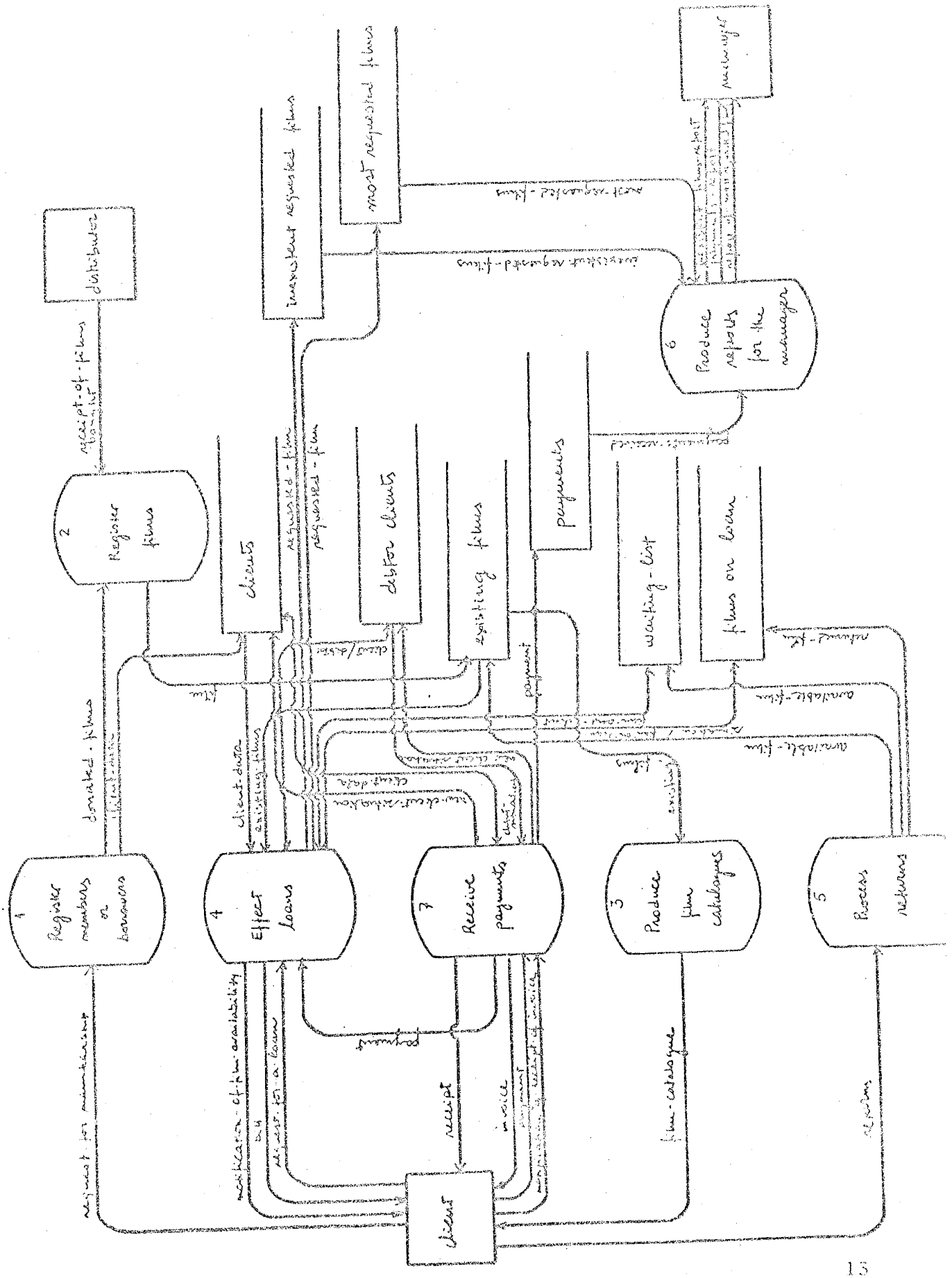
Paris, C. (1987) "Combining Discourse Strategies to Generate Descriptions to Users along a Naive/Expert Spectrum", Proceedings of Tenth IJCAI Conference, 626-632.

McDonald, D.D. and Pustejovsky, J.D. (1985) "Description-directed Natural Language Generation", Proceedings of Ninth IJCAI Conference, 799-805.

McKeown, K.R. (1985) Text Generation: Using Discourse Strategies and focus constraints to generate natural language text, Cambridge University Press, Cambridge.

Suchman, L.A. (1987) Plans and Situated Actions: The Problem of Human Machine Communication, Cambridge University Press, Cambridge.

Figure 1- Data flow diagram of SISVIDEO



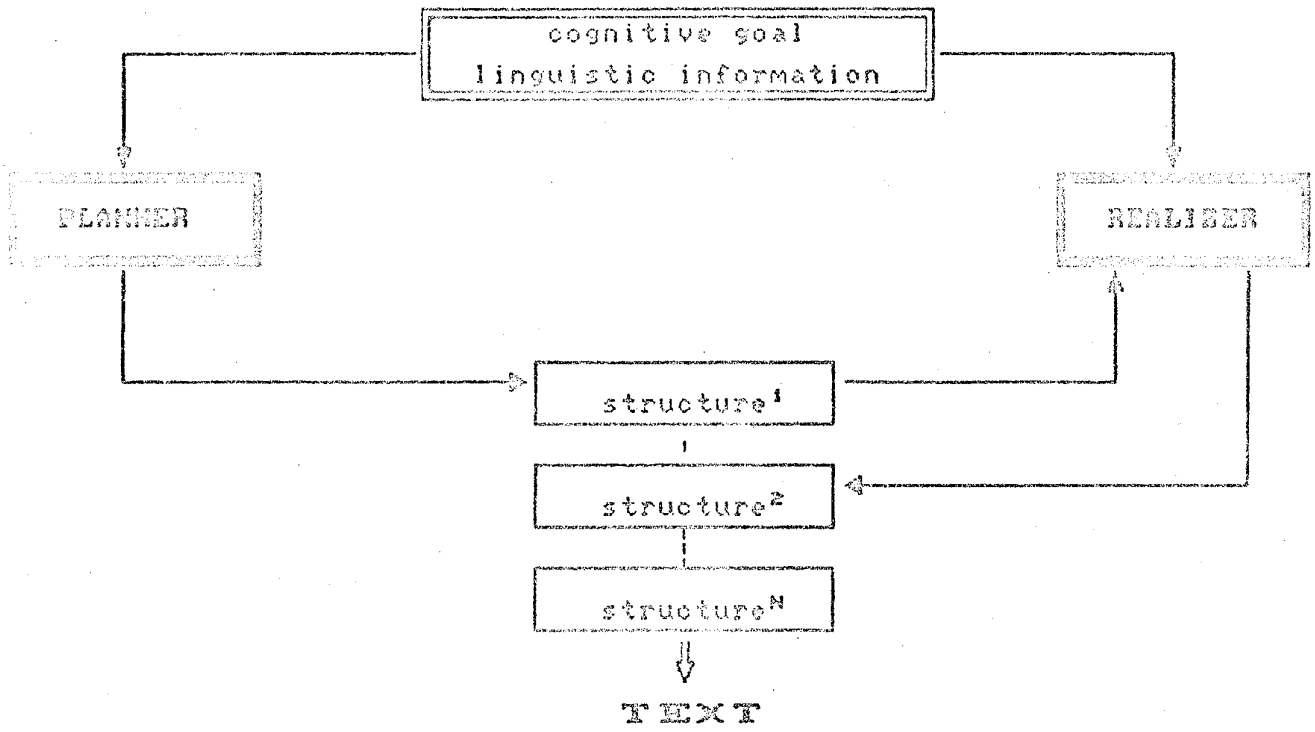


Figure 2: General view of GEMA's planning and realizing activities

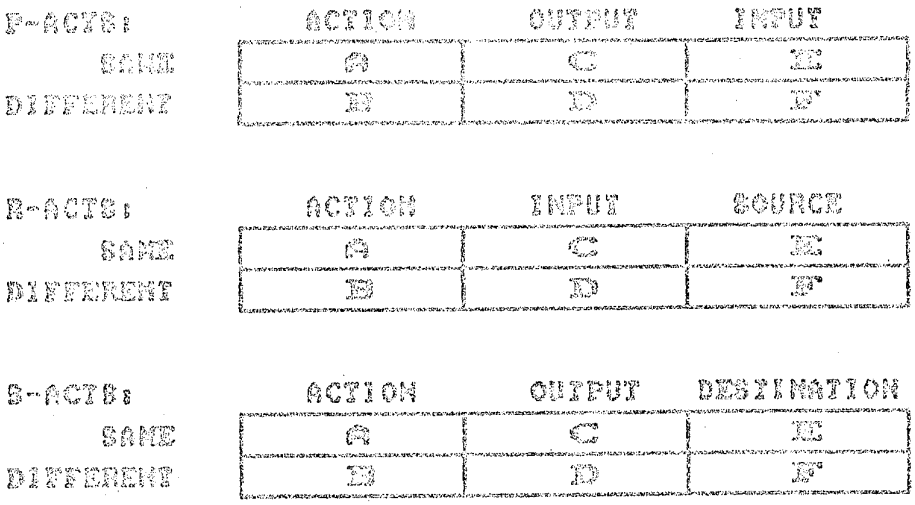


Figure 3: Matrices for identifying transformational opportunities