ELSEVIER

Artificial Intelligence in Medicine 8 (1996) 123–145

Artificial Intelligence in Medicine

Generating recipient-centered explanations about drug prescription

Berardina De Carolis [a], Fiorella de Rosis [a,*], Floriana Grasso [a], Anna Rossiello [a], Dianne C. Berry [b,1], Tony Gillie [b]

[a] *Dipartimento di Informatica, Università di Bari, Via Amendola 173, 70126 Bari, Italy*
[b] *Department of Psychology, University of Reading, Earley Gate, Whiteknights, Reading RG6 2AL, UK*

Received 1 February 1995; revised 1 July 1995; accepted 16 July 1995

Abstract

In this paper we describe how we generated written explanations to 'indirect users' of a knowledge-based system in the domain of drug prescription. We call 'indirect users' the intended recipients of explanations, to distinguish them from the prescriber (the 'direct' user) who interacts with the system. The Explanation Generator was designed after several studies about indirect users' information needs and physicians' explanatory attitudes in this domain. It integrates text planning techniques with ATN-based surface generation. A double modeling component enables adapting the information content, order and style to the indirect user to whom explanation is addressed. Several examples of computer-generated texts are provided, and they are contrasted with the physicians' explanations to discuss advantages and limits of the approach adopted.

Keywords: Drug prescription; Explanation; User modeling; Text planning; Natural language generation

1. Introduction

Explanation systems must now be able to do more than simply justify decisions. In many areas, new types of users are emerging who require very different types of explanations. In the medical domain, for examples, doctors, nurses and patients may all be recipients of medical explanations about a particular case, but the type of explanation required by the three classes of user is likely to be different. The recent interest for

* Corresponding author. Tel.: (39-80) 544-3284; Fax: (39-80) 544-3196; e-mail: derosis@gauss.uniba.it
[1] e-mail: sxsberry@susssys1.reading.ac.uk

SSDI 0933-3657(95)00029-1

increasing population compliance with health care resulted in a number of studies aimed at integrating knowledge-based systems with a patient-oriented explanation component [2,3,8]. In these projects, patients typically interact directly with the system to browse through the medical record. In some cases, they receive clarifications of terms that might be unclear to them [3]. In others, they can look at a patient education pamphlet that contains information about the disease and instructions for prescribed drugs [8]. In this latter case, general information is integrated with user-specific parameters employed in the decision process. Finally, the explanation text can be generated dynamically by exploiting a user model that shapes the message to the patient's clinical history [2].

The methods employed to generate the text in the above examples are not the same: hypertext with canned texts was used by Jimison et al. [8], and planning or rule based text generation was used by Buchanan et al. [2] and Cawsey et al. [3]. In all cases, however, the patient interacts directly with the system to search the information needed. This approach limits, in our view, the likely diffusion of such systems; in several cases, people do not have the possibility nor the background to use a computer, and would prefer to receive a written explanation which can be taken away with them.

In addition, several types of user other than the patients might like to have access to the explanation facility. Typically, in hospitals various health professionals cooperate to ensure the compliance with treatment; the hospital team who cares for the patient and, once the patient has left the hospital, the general practitioner. Traditionally, some of these people receive verbal explanations about what has to be done: this happens, for instance, to nurses. A discharge letter plays a similar role for the general practitioner. The content of explanations addressed to these users (patient, nurse, general practitioner) is not the same, although the decision process behind them is the same. These explanations vary as a function of a number of factors; the role of their intended recipients in the health care cycle, their competence, their relationship with the physician who was responsible for the initial decision, and so on.

In this paper, we describe how we have studied the problem of generating written explanations to 'indirect users' of a knowledge-based system in the domain of drug prescription (OPADE) [2]. We call 'indirect users' those people (patients, nurses or general practioners) who are the intended recipients of the explanation, although they do not directly interact with the system themselves. We use this term to distinguish these users from the prescriber (the 'direct user') who interacts with the system in order to be guided in the decision process, and who and is also responsible for delivering the explanations to the indirect users. The role of the Explanation Module is then to support the prescriber's explanation function by generating a text that he or she will possibly modify before giving it to the indirect user.

[2] The OPADE Project was partially funded by the Advanced Informatics in Medicine initiative of the EEC, under Contract A2027. It was a three-year Project which started in January 1992. The Consortium was composed by the following partners: BIM (B), Bari University (I), City University of London and Reading University (UK), ECLIMED and SETEC Informatique (F) and SWEDIS Dev Center (S). It was a cooperative venture; we are therefore indebted to our colleagues for fruitful discussions. In particular, Bertrand Séné and Alain Venot cooperated to specify the explanation component of the system.

2. Which explanations about the drug prescription

In the context of OPADE, we were given the responsibility of designing, implementing and evaluating the module for generating explanations to indirect users. We had previously carried out experimental studies to assess information needs of the intended users of these explanations. These questionnaire studies were conducted by the Department of Psychology, University of Reading. The first of them involved 243 people [1], the second one 73 nurses. To examine how doctors would explain the prescription to indirect users, we carried out a further study where hospital doctors were asked to make a prescription based on a clinical 'scenario'. They were then asked to produce explanations addressed to a patient, a nurse and a colleague. Five doctors participated to the study, and each of them was required to examine two scenarios. The four scenarios employed corresponded to two cases of angina, one case of hypertyroidism and one case of tuberculosis (see Appendix). Therefore, we could examine 30 explanation texts overall. Analysis of results of the three studies, and their comparison with results of similar studies at international level, enabled us to design the adaptive explanation component of OPADE.

2.1. Physician's explanations

Let us synthetize the results of analysis of the explanation texts provided by doctors, in terms of their pattern of communicative goals and of the related information content.

2.1.1. Purpose

The purpose of explanation is to ensure that the prescription is followed, and is followed correctly. To this aim, the Hearer is informed about the prescription, persuaded that the prescription is correct and instructed on how to perform it. The relevance of these communicative goals is different for the different persons to whom the text is addressed. Persuading the Hearer is less important if he or she is 'benevolent' towards the doctor, that is, if the Hearer is inclined to accept the doctor decisions without discussing them. Instruction is less important to someone who is competent on drug administration, and so on. Each of these goals can be further refined in terms of new subgoals. For instance, persuading the Hearer about the prescription requires, first of all, highlighting its positive aspects: for example, efficacy against the disease. Subsequently, the possible negative aspects (such as side effects and contraindications), have to be illustrated with a reassuring attitude, for instance by saying that they are not frequent, not serious or do not apply to the patient's specific case. The 'intentional structure' of the explanation text can therefore be represented in a tree structure whose nodes are communicative goals and whose leaves are primitive actions of the communication process (to inform, to recommend, and so on).

2.1.2. Information content

The frequency distribution of information items in the explanation texts examined is shown in Fig. 1. The overall structure of the text does not vary much in the examined cases: the text starts by describing the rationale for the prescription (patient characteris-

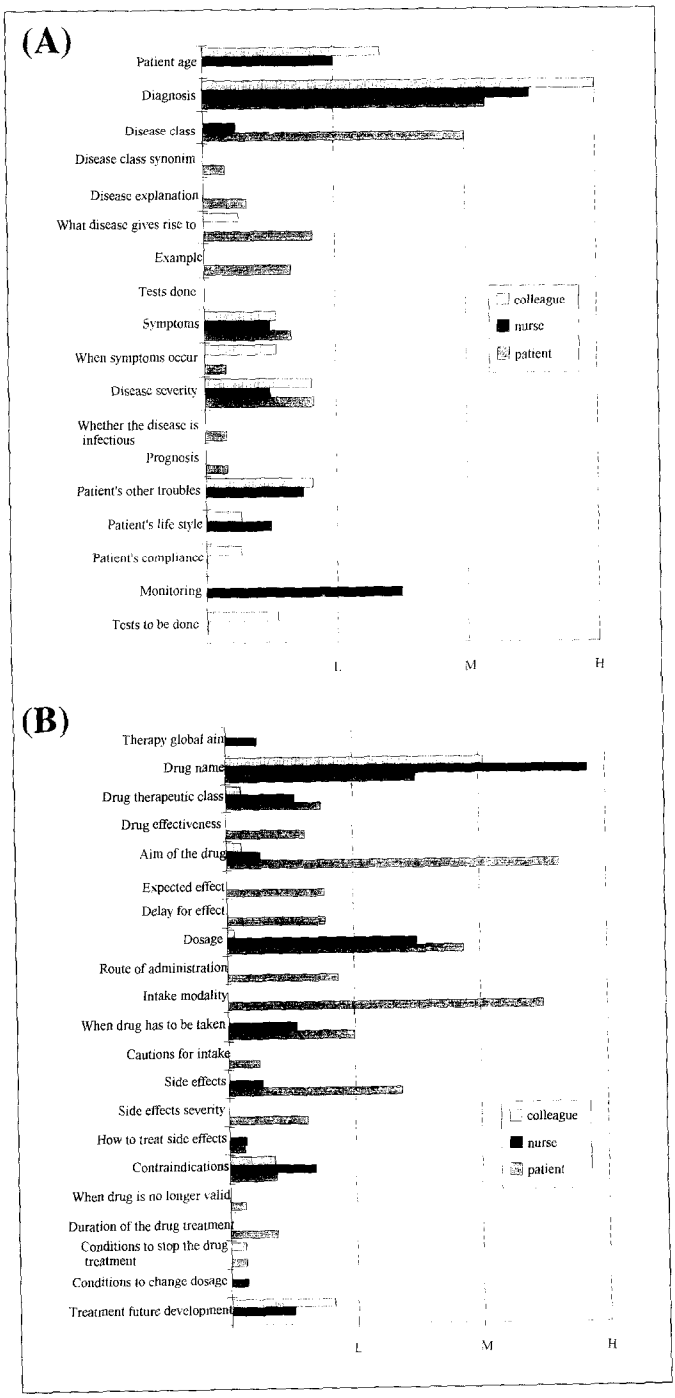

Fig. 1. (a) Frequency distribution of items concerning patient health status in doctors explanations addressed to colleagues, nurses and patients, classified according to designated level of importance in the user model (low, medium, high). (b) frequency distribution of items concerning the prescription in doctors explanations addressed to colleagues, nurses and patients, classified according to designated level of importance in the user model (low, medium, high).

tics and health status) and proceeds with a description of the treatment, in general and with details: several aspects of the prescribed drug are illustrated, with administration details. Precautions in taking or administering it follow (side effects and contraindications) and finally some statement is made about the future treatment planning. The order of presentation of these information items may vary, from text to text. In addition, the frequency with which each item is mentioned is different for explanations addressed to nurses, to patients and to doctor colleagues. More details about the disease, aim of the drug, drug administration and side effects are given to the patient. Some more details about patient's life style, therapy aim, contraindications and especially monitoring are given to the nurse. More details about the patient status and the rationale for the therapeutic decision (diagnosis, severity, other troubles, etc.) and less explanations about the therapy, except its future developments, are given to the colleague.

2.2. Indirect users' information needs

Many studies have shown that patients are not satisfied with simply obtaining a list of drugs they are prescribed. Almost all of them would like to know the reasons for the drugs they are using, as well as the risks involved; the vast majority of them prefer receiving written information. Several studies have shown that the satisfaction of this need affects both patients' knowledge and compliance [15]. Recently, Diana Forsythe [5] combined observations and interviews to examine the information needs of migraine patients. She observed that patients rarely spontaneously demand explanations from their doctors, and that physicians do not necessarily give long and coherent explanations: patients have, therefore, information needs of their own that sometimes go unsatisfied. The main result of her studies is that patients need, at the same time, a formal, general knowledge of the sort found in text books (anatomy, physiology and pharmacology of the disease) and a more informal, specific knowledge about prognosis, etiology, severity of the disease. This specific knowledge may vary according to the context and to the person. Therefore, explanations generated need to be adapted to patient characteristics such as gender, age, educational level, ethnic background and even emotional concerns. Forsythe claims that wording and style of explanations should be tailored, as well, to the user situation, and that, in particular, enlistment of the patient by the physician is promoted by the use of what she calls 'inclusive language', that is language that explicitly recognizes patients' competence and thus treats them with respect. Another finding of this research which influenced our work is the observation that what patients want to know is not the same as what providers believe they need to know: consequently, Forsythe's suggestion is that both perspectives should be incorporated into the design of the explanation system and should contribute to generate the text.

Less knowledge seems to exist about the kind of information that other health professionals, such as nurses and general practitioners, might be interested in receiving. It is frequently argued that mistakes in drug prescriptions which occur in hospitals and lack of coordination in the health care efforts might be due, at least in part, to an inadequate information exchange within the health care team. Several studies have concluded, in particular, that new means are necessary to relate unbiased, authoritative information pertaining to drugs or drug therapy to prescribing physicians [19]. However,

few proposals have been made about how information systems might address this problem. Nurses should undoubtedly be enabled to access directly relevant sections of the Hospital information system: but, in addition to this, receiving specific written explanation material about the drug prescription would probably contribute to the goal.

For a detailed description of the results of our studies examining the needs of indirect users, one should refer to the two cited papers [1,4,15]. In this paper, we will just contrast these results with the results of the study about doctors' explanations, to find out similarities and differences between the two points of view.

2.3. Conflicts between physicians and indirect users' perspectives

By comparing results of analysis of physicians explanations with results of the two questionnaire studies about nurses and patients information needs, we were able to discover a number of differences which confirm the results of Forsythe's study about migraine patients, and also extend them to the case of nurses. Unfortunately, we could not do the same comparison for general practitioners, because no study was made about their information needs.

2.3.1. Conflicts with patients information needs

The major difference between physician explanations and patient needs is in the importance attached to items and in the level of detail required for each of them. Patients want to know first 'what is the diagnosis', in ordinary terms, 'what causes it' and, in some cases, 'how certain is the doctor about the diagnosis' and if there is an alternative, more serious hypothesis (for example: leukaemia, stomach cancer, etc.) which can be excluded. As for drug description, they want to know 'what is the drug' and 'what it does' (whether it cures or only relieves the pain), which is its efficacy and the expected delay for effect, and, in some cases, if it involves a risk of fatality. Precautions about the drug are the most frequently required items: (i) side effects, in general and specifically for drowsiness; (ii) lifestyle changes, especially alcohol, driving, food to be avoided, work on machinery, breast feeding; (iii) interactions between prescribed and common drugs such as contraceptives or Aspirin. It should be noted that interactions between drugs and lifestyle changes were not mentioned in doctors texts and that many doctors tended to skip over side effects and contraindications in order to avoid frightening the patient. The few questions about drug administration deal with treatment duration and how and when to take the drug, in combination or separately, and their relation with meals (something which is mentioned neither in the prescription nor in the doctor's explanation). More frequent are 'what if' questions about monitoring and control and about future treatment, such as: 'what to do if the symptoms change or do not change', 'when to go back to the doctor', 'what will happen if I do not take the drug', 'what to do if I forget to take one or if I take too much'. Finally, questions that explore the possibility of avoiding the therapy or adopting alternatives, such as: 'what can I do to prevent the pain from coming back', 'is there any alternative, drug or not drug, to this therapy', prove that not all patients accept the suggested prescription as something unquestionable.

2.3.2. Conflicts with nurses' information needs

Nurses do not ask explicitly to be informed about the patient health status, but want to know what is the rationale for the prescription: 'why that drug, 'what it is for'. In addition to general questions such as 'what it is', they make very specific ones about the drug: 'research done and results'; 'extent of use, in general and by that prescriber'; 'level of success, also in relation to cost'; 'how it is metabolised'; 'is there an antidote'; 'is it going to mask any other symptoms'. The same consideration holds for drug administration: dosage range (min,max), frequency, duration. Information needs about precautions have the same importance as for patients but are, again, more specific: side effects and contraindications, in general and with specific mention to some diseases such as diabetes, heart failure, allergy status. Information on interactions is required too. The future treatment planning is focused on observation of possible side effects and adverse reactions rather that on changes to the therapy depending on the patient status. Some interest is shown for administrative or legal issues such as 'am I allowed to give this drug under hospital policy' or 'is it stocked by the pharmacy'. A final consideration: nurses do not seem to limit their information to what the doctor tells them: many of them say that they would possibly consult an information source like BNF or the hospital pharmacist, to elicit more details about the drug.

2.4. Text structure

According to the 'rhetorical structure theory' [6.9], the coherence of a text is improved by making overt the relationships between its elements through the use of 'linguistic markers'. These markers specify how two or more sentences contribute two reach a desired purpose, thus fostering the inference process that the Hearer has to make to understand the text. For example, in the physicians' texts, the patient health status corresponds to the problem to be solved, and the prescription is the solution. The two text spans are linked by a rhetorical relation (RR) of 'Solutionhood', whose nucleus (the problem) is the diagnosis and whose satellite (the solution) is the therapy. Various aspects of the diagnosis (such as etiology, severity, prognosis) are linked to the diagnosis name by a RR of 'Elaboration Object Attribute'. The disease class is linked to the disease name to a RR of 'Elaboration Set Member',... and so on. The linguistic markers used to manifest the rhetorical relation between two text spans vary according to the relation itself, to what the Speaker presumes the Hearer should know and to the Speaker—Hearer relationship. For example, the relation of Solutionhood between Description of Patient Status and Description of Therapy is usually manifested, in the doctor's text, by linking nucleus and satellite sentences by markers such as 'in order to solve this problem', or something like this. A relation of Elaboration Set Member is expressed by a 'is a' marker, and so on.

2.5. Hints to the text generation

We could draw a number of conclusions from results of experimental studies made in OPADE:

The *overall structure of the explanation discourse* can be represented as a tree whose nodes are communicative goal/subgoals and whose leaves are primitive communicative actions;
this tree can be pruned out in specific situations, that is when making explanations to specific types of Hearer. This means that some communicative goals will be omitted in those cases, and that some items will, correspondingly, be omitted from the text;
criteria to decide how to prune the tree are different if one takes the point of view of the Speaker or of the Hearer. As we have seen in the previous paragraph, direct and indirect users do not attach the same importance to each goal or information item. These criteria are also a function of the overall amount of information that can be put in the text, and of individual characteristics such as the Speaker 'verbosity'. For example, if the prescription includes several drugs, less details will be given, to follow a criterion of parsimony in the text length.
primitive communicative actions are expressed through *surface sentences*, clauses or phrases which are a function of the Hearer to whom they are addressed. For example, 'Inform about the patient disease' becomes: 'You have been diagnosed as suffering from' to the patient, or 'This patient has developed (or has got)..' to a nurse;
to insure coherence of the text, and therefore its understanding, sentences have to be linked by linguistic markers depending on the *rhetorical relations* between them.

3. Data and knowledge sources

Most researchers agree that an explanation module should not construct its own ad hoc database but should synthetize the text directly from the same knowledge sources used for problem solving. Moore and Paris give very sound justifications for this requirement [13]: the explanation will be coherent with the decision suggested, will evolve automatically when the knowledge base is updated, will not produce redundant information,... and so on. However, limitations in data and knowledge sources available can produce problems in generated texts. In the case of OPADE, data and knowledge sources are typical of medical information systems. They can be classified into two main categories:

- General (patient-independent) data:
  - a **Thesaurus** of terms employed in interacting with the system;
  - a **Database of drugs** available in the European Community;
- Patient-specific information:
  - the **Patient Record**;
  - the suggested **Prescription**, with some tracing of the critiques made by the system and whether they have been agreed by the prescribing doctor.

OPADE supports the drug prescription process and not the diagnosis. Therefore, a list of the patient's main problems (diagnosis, symptoms, risk factors and so on) can be found in the Patient Record: but, the reasoning process through which these problems motivated the choice of a specific drug is not available. Linkage to this large and well organized set of data and knowledge bases is the strength of the Project and enabled us to overtake the level of a purely demonstrative system. However, it restricted consider-

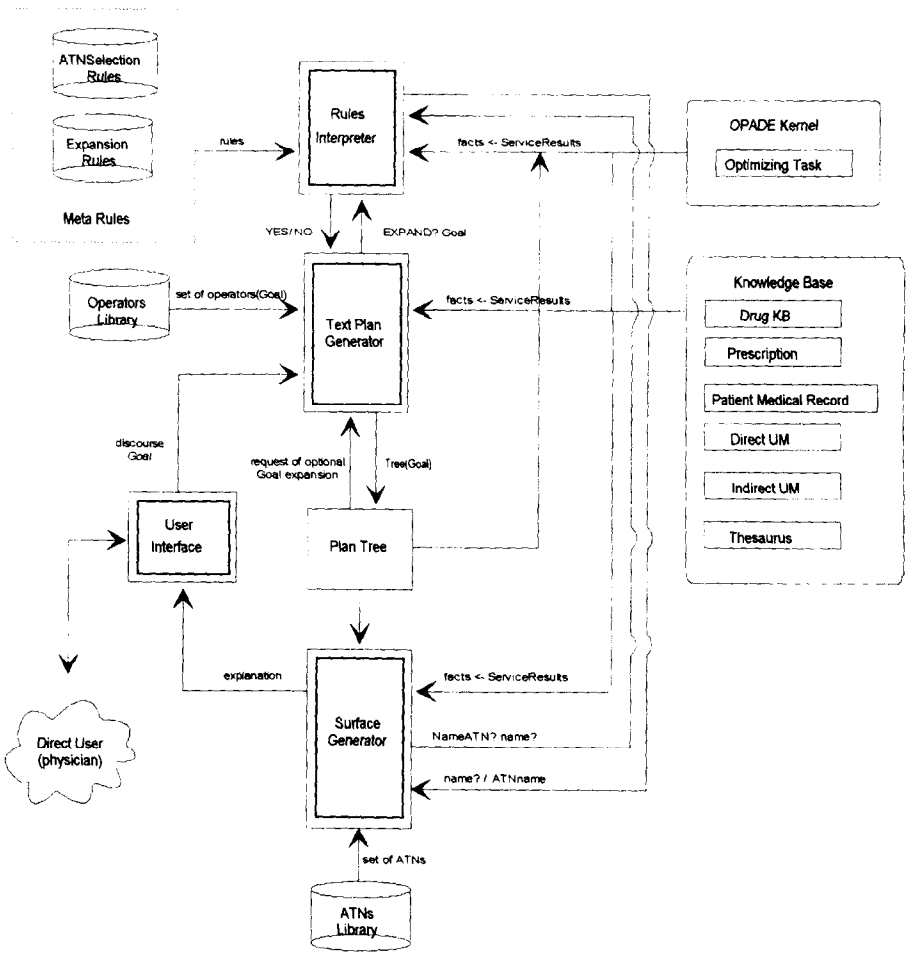

Fig. 2. Architecture of the Explanation Generator.

ably the information content of our texts: if some item was not available in OPADE's knowledge sources, we were not able to include it in our explanations. To solve this problem, we allow the prescriber (direct user) to modify and integrate the computer generated text by presenting it within a text editor.

4. Text generation

We extended ideas of previous text generation systems [10–13] to design a system which supports the physician's explanation activity by considering, at the same time, both the physician's and the indirect user's points of view about the explanation purpose and style. The architecture of this system is described in Fig. 2. It is made up of four main components, that we will describe in detail.

4.1. The user models

Two user models describe the characteristics of the Speaker S and of the Hearer H of explanation (Direct and Indirect users, respectively). The two models hold information needed for planning, conflict resolution and surface generation:

1. About the **Speaker**:
   - **propensity to talk** about subjects, coded in three levels (high, medium or low);
   - **preferences about message styles**: 'Verbose' vs. 'Terse', 'Formal' vs. 'Informal'.
2. About the **Hearer**:
   - **cognitive status**, that is knowledge of subjects that might be included in the text;
   - **information interests**, that is level of Hearer's interest to be informed about each subject, again coded as 'high' or 'medium' or 'low'.

Default values of these data are settled in the stereotype-based models [17] according to the result of experimental studies mentioned in Section 2. However, in any phase of interaction the prescriber can inspect his or her model to modify default values. The 'System Administrator's can do the same for indirect users at the installation phase. Fig. 3a and b show two examples of snapshots of the User Model updating procedure. In the first one, the Direct User looks at the model and discovers that his or her level of propensity to talk to a nurse about the patient's environment conditions is low; the DU can change this level to 'medium' or 'high' by just clicking on the corresponding button. The DU can also change the 'Spoken mode' from 'Formal' to 'Informal' and from 'Verbose' to 'Terse'. In Fig. 3b, the system administrator can make a similar change to the level of interest to be informed about overdosage effects, by the indirect user (i.e., a patient) modeled in this example.

4.2. The text planner

The planner establishes the structure of the text, given a specific Speaker, a specific Hearer and a case to explain. It states information items to be included in the text, their order of presentation and the rhetorical relations among text spans. Text generation criteria are described in a library of *plan operators* which establish how a communicative goal or subgoal can be reached when specific application conditions hold and which effects the operator enables to obtain on the cognitive status of the Hearer. A plan operator is a tuple: (Header, Constraints, IntentionalConstraints, Preconditions, Effects, Decomposition, RhetoricalRelation). The meaning of these slots is the same as in other text planners (see, e.g., Refs. [10,14]). We separate conditions on the domain knowledge base (in the Constraints) from conditions on the Hearer Model (in Intentional Constraints) to allow faster search in the knowledge base. Subgoals in the Decomposition are marked as 'Nucleus' (N) or 'Satellite' (S) of the Rhetorical Relation mentioned in the relevant slot. Some of these subgoals can be optional. The Header and the Rhetorical Relation slots enable discourse segment purposes and intersegment relations to be represented at the same time. Different relations can be attached to the same communicative goal in different contexts, by defining different plan operators having the same Header and different RhetoricalRelation slot values. The action performed by the operator consists of a decomposition of the goal into two or more subgoals, or in one or

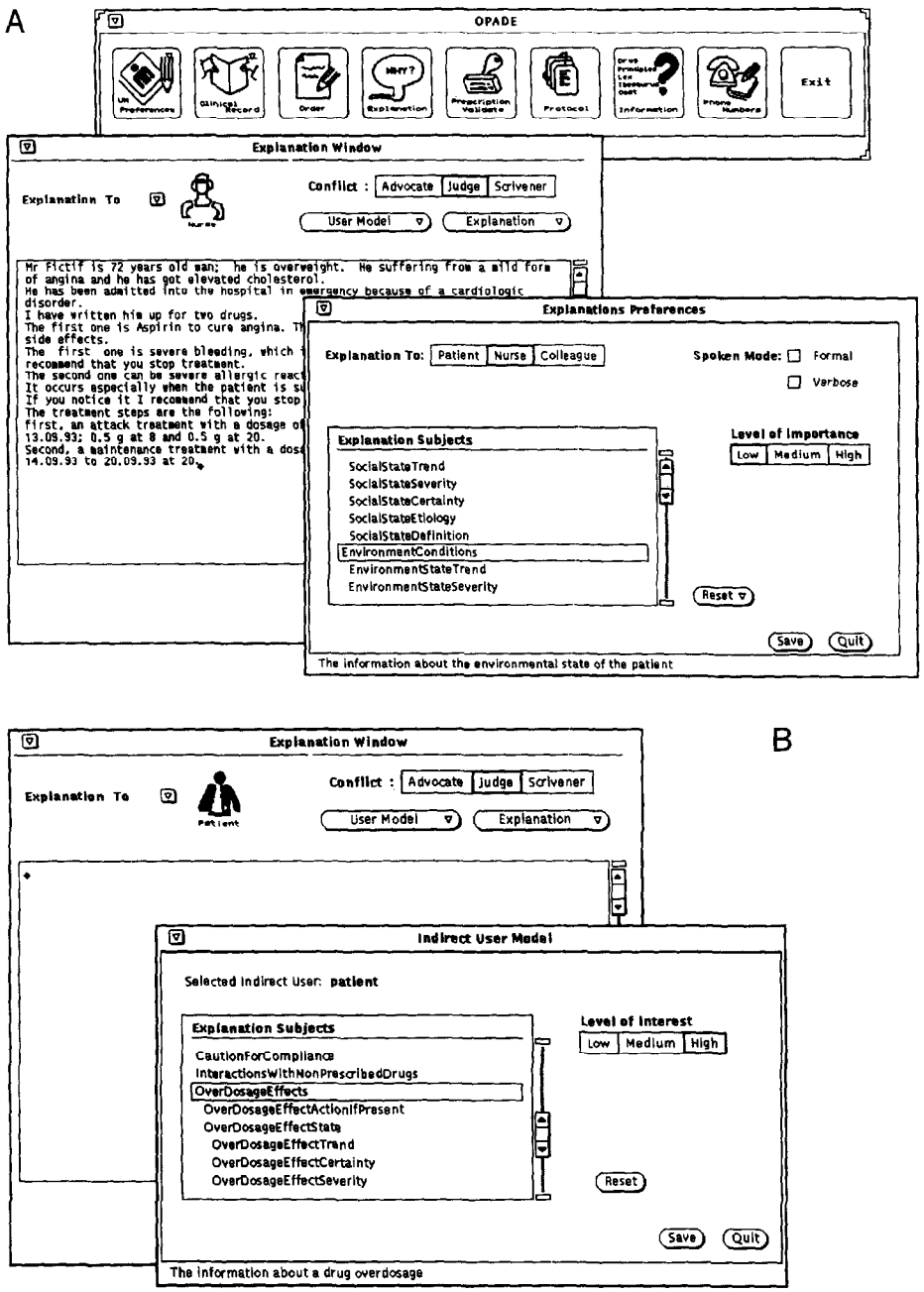

Fig. 3. (a) Direct User Model updating window. (b) Indirect User Model updating window.

Header	Explain S H (Therapy! patient?)	
Constraints	(Patient! patient?)	
Intentional Constraints:	NOT (KNOW-ABOUT? H (Therapy! patient?))	
Preconditions		
Effects	KNOW-ABOUT? H (Therapy! patient?)	
Decomposition	Describe S H (HealthStatus! patient?)	**S opt**
	Expose S H (TreatmentPlan! patient?)	**N**
Rhetorical Relation	Solutionhood	

Fig. 4. A plan operator.

more 'surface speech acts', such as *inform, request*, etc. An example of plan operator is shown in Fig. 4. Planning of a text is made by backward chaining of plan operators through a matching of Header -Decomposition or Precondition-Effect slot values.

4.3. The conflict resolutor

A metalevel mediator finds a compromise between the prescriber and the explanation receiver views when they are discrepant. For example: the prescriber has a low propensity to talk about side effects of a drug, especially if they are serious, and the patient has a high interest to be informed about them. Compromise-finding criteria are represented in the form of a set of *metarules* which are evoked during the hierarchical planning process [18] to decide, at each level of abstraction, whether an 'optional' subgoal has to be expanded. At the first abstraction level, only obligatory subgoals are expanded; at the following abstraction levels, optional subgoals are expanded according to a logical combination of conditions on the following features:

- level of *propensity to talk about the subject* by the Direct User
- level of *interest to be informed about the subject* by the Indirect User
- degree of *complexity* of the resulting *plan* (in the hypothesis of expanding the subgoal)
- *personality of the Speaker*: for instance, his or her 'verbosity'.

The complexity of a candidate plan is measured by a score which is a function of the percentage of the 'expandable' leaves which are expanded in the plan. Expandable leaves are non-obligatory 'successors' of optional nodes. In a low-complexity plan (score = 1), a fraction ranging from zero to one third of all optional leaves is expanded. In a high-complexity plan (score = 3), a fraction from two thirds to all optional leaves is expanded.

Metarules can give different weights to each of these parameters. By defining different metarule sets, one can therefore simulate different behaviours for the System.

```
R 1   (IF ((EQUAL (LevelofAbstraction! plantree) 5)) THEN (STOP? system))

R 2   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 2)
              (WANT_TO_SAY? speaker subject1? (Type! hearer) high) )  THEN (EXPAND? system subject1?))

R 3   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 3)
              (LESSEQUAL (LevelofComplexity! plantree) 2)
              (WANT_TO_SAY? speaker subject2? (Type! hearer) medium) )  THEN (EXPAND? system subject2?))

R 4   (IF (   (EQUAL (LevelofAbstraction! plantree) 4)
              (EQUAL (LevelofComplexity! plantree) 1)
              (WANT_TO_SAY? speaker subject3? (Type! hearer) low)
              (Verbose? speaker TRUE  )        )  THEN (EXPAND? system subject3?))
```

The System as Scrivener

```
R 1   (IF ((EQUAL (LevelofAbstraction! plantree) 5)) THEN (STOP? system))

R 2   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 2)
              (WANT_TO_KNOW? hearer subject1? high)  )  THEN (EXPAND? system subject1?))

R 3   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 3)
              (LESSEQUAL (LevelofComplexity! plantree) 2)
              (WANT_TO_KNOW? hearer subject2? medium)  )  THEN (EXPAND? system subject2?))

      (IF (
         (EQUAL (LevelofAbstraction! plantree) 4)
         (EQUAL (LevelofComplexity! plantree) 1)
         (WANT_TO_KNOW? hearer subject3? low)
         (Curious? hearer TRUE)  )  THEN (EXPAND? system subject3?))
```

The System as Advocate

```
R 1   (IF ((EQUAL (LevelofAbstraction! plantree) 5))        THEN (STOP? system))

R 2   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 2)
              (WANT_TO_SAY? speaker subject1? (Type! hearer) high)
              (WANT_TO_KNOW? hearer subject1? high)          )  THEN (EXPAND? system subject1?))

R 3   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 3)
              (LESSEQUAL (LevelofComplexity! plantree) 2)
              (WANT_TO_SAY? speaker subject2? (Type! hearer) high) )  THEN (EXPAND? system subject2?))

R 4   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 3)
              (LESSEQUAL (LevelofComplexity! plantree) 2)
              (WANT_TO_KNOW? hearer subject3? high)  )  THEN (EXPAND? system subject3?))

R 5   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 4)
              (EQUAL (LevelofComplexity! plantree) 1)
              (WANT_TO_KNOW? hearer subject5? medium)        )  THEN (EXPAND? system subject5?))

R 6   (IF (   (GREATEREQUAL (LevelofAbstraction! plantree) 4)
              (EQUAL (LevelofComplexity! plantree) 1)
              (WANT_TO_SAY? speaker subject8? (Type! hearer) medium))  THEN (EXPAND? system subject8?))
```

The System as Judge

Fig. 5. Three metarule sets.

- a *Scrivener*, who follows strictly the Prescriber's desires;
- an *Advocate of the indirect User's needs*;
- a *Judge*, who mediates between the two points of view.

Fig. 5 shows examples of metarule sets, in the three cases.

When acting as a Scrivener, the system decides whether to expand a subgoal according only to the Speaker's level of propensity to talk about the related subjects (WANT-TO-SAY predicate). The opposite holds when the system acts as an Advocate; in this case, only the Hearer's interest to be informed about subjects (WANT-TO-KNOW predicate) guides the expansion process. The System as a Judge establishes a compro-

mise between the two views: a subgoal is expanded when it is important to either the Speaker or the Hearer. The level of importance needed to expand the subgoal is, in the three cases, a function of the plan complexity. Other features that guide the expansion process are Speaker and Hearer personality traits: for instance, whether S is 'verbose' or H is 'curious'. The conflict resolution criterion can be changed by the system administrator by clicking on the appropriate button in the User Model window.

4.4. The surface generator

A ATN-based Surface Generator [12,20] takes the plan as an input and produces a text in a specified language. Several levels of nested Augmented Transition Networks do this job:

- The ***top-level ATN*** explores the plan-tree from its root, examines the type of node (intermediate or leaf), the associated rhetorical relation or primitive act, and jumps to lower level nodes;
- ***PR-level ATNs*** generate user-tailored linguistic markers when both nucleus and satellite are expanded in the plan;
- ***PR-level ATNs*** generate fragments of sentences according to the communicative act associated to the leaf node, by reading data from Knowledge and Data Base sources.

ATNs enable adaptation of the text phrasing to factors which were not considered in the planning phase. The plan is independent of the language, and also of factors like the Hearer sex and the Speaker formality. These adaptivity factors are introduced as conditions on the ATN's arcs which can be tested by evaluating the value of a function or by applying a set of rules. As in PROSENET and TEXT [11,12], top down and recursive representation of the algorithm to generate the text under various conditions is obtained by labelling arcs with various classes of functions. The first class enables the program to traverse the plan tree by examining the type of node (intermediate or leaf) and the associated rhetorical relation or primitive action. At the same time, the complexity of the portion of the tree below the node is computed. The second class of functions reads the values of items from the KB. The third one sets the names of registers and executes input–output operations from them. Finally, common PUSH, POP, PROSE operations can be made.

5. Texts generated: an example

Fig. 6 shows an example of plan tree that corresponds to the segment of the discourse which is aimed at describing the Treatment Plan. To simplify illustration, several subgoals have been omitted. The plan tree shows, at each node, the communication goal with the associated Rhetorical Relation (in bold):

Expose the Treatment Plan, by linking the two subgoals 'Plan in General' and 'Plan in Particular' with an ElaborationGeneralSpecific;

Expose the Plan in General, by Informing about the Number of Drugs and the Duration of Therapy, in Joint;

Expose the Plan in Particular, by Exposing the Treatment Steps, in OrdinalSequence;

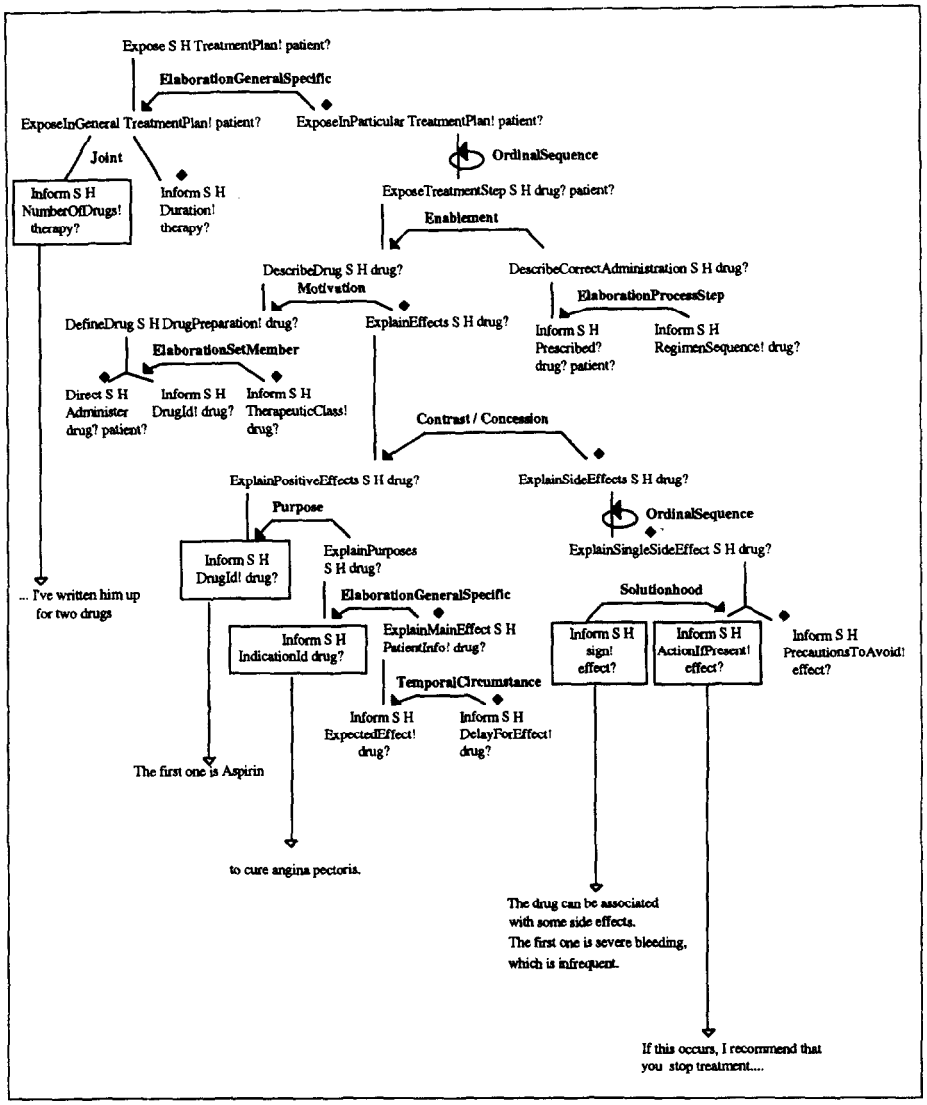

Fig. 6. A portion of the plan tree with the correspondingly generated text.

Expose the Treatment Steps, by Describing the Drug and How to Administer it, with an Enablement;

Describe the Drug, by giving a Description of the Drug Preparation and by Explaining its Effects, with a Motivation,... and so on.

In some cases, two rhetorical relations are associated with the same node (see, for example, 'Contrast/Concession' for 'Explain Effects of Drug'), to indicate that the communication goal can be obtained by two distinct techniques. Fig. 6 also shows the correspondence between communicative actions (for instance: (Inform S H Number Of

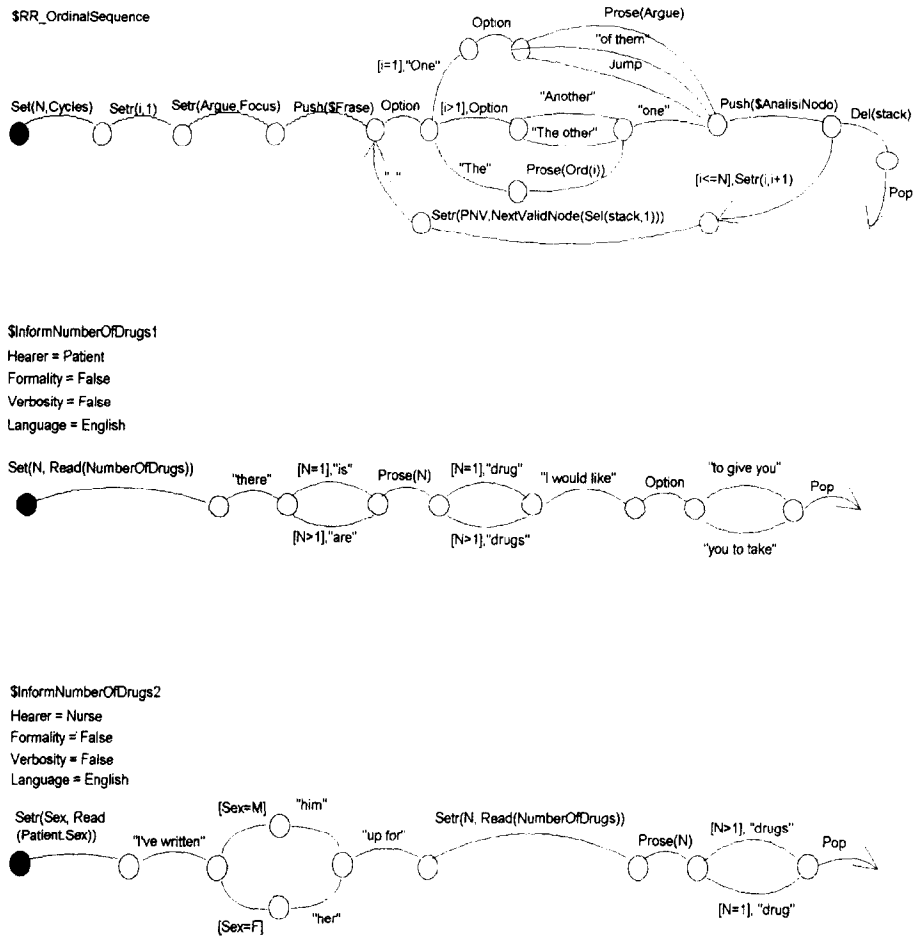

Fig. 7. Some examples of ATNs: RR-ATN of 'Ordinal Sequence', Pr-ATN for 'Inform on Number of Drugs' in texts addressed to patients and Pr-ATN for 'Inform on Number of Drugs' in texts addressed to nurses.

Drugs ! therapy?)) and sentences generated when the text is addressed to a nurse ('I've written him up for two drugs').

Fig. 7 shows the ATNs that are employed to generate a small portion of the text for a patient and a nurse. The RhetoricalRelation ATN of Ordinal Sequence generates the same linguistic markers in both cases. The first arc of this net is labelled with a call to a function which computes how many times the cycle has to be repeated and sets a local register accordingly. The second one initializes the counter of iterations. The third one calls a function which sets the register 'Argue' to the value of the current focus. Subsequent tests enable the program to generate different text fragments in correspondence with different steps of the cycle: 'one of them' at the first step, 'another one' or 'the second one' subsequently. The PUSH function switches the control to the ATN associated with the next valid node.

The PR-level ATN of Inform on Number of Drugs which is shown in the figure is selected when the Hearer is a patient, the Speaker is not formal or verbose and the language is English. The PR-ATN for the same communicative action which is selected when the Hearer is a nurse shows that the text generated depends on the patient sex as well as on the number of drugs prescribed. Finally, Fig. 8 compares three examples of explanations of scenario 4 in the Appendix. The texts are addressed, respectively, to a patient, a nurse and a general practitioner.

The patient text is the most detailed. It includes explanations about the meaning of terms, describes how the drug has to be taken and warns about possible interactions with other drugs. The nurse text mentions the patient's general conditions (age, weight) and the way in which he was admitted into the hospital. Some 'inclusive language' is adopted in the text addressed to the general practitioner, to make it more acceptable.

6. Unresolved problems and future work

The texts that can be obtained with this Generator have several properties worth mentioning. By using the same knowledge sources that are employed for the main, prescriber critiquing, function of OPADE, one can generate explanation texts in a very flexible way. Texts can be adapted to the prescriber's personality and to the person to whom they are addressed. In addition, they can be adapted to criteria established at the site installation level, that fix the relative weights to be given to the Speaker and the Hearer's points of view about the information content of explanation texts. With minor modifications to the module, they can be adapted to the characteristics of the drugs prescribed and to the patient disease problems. By only extending the ATN library, the explanation can be produced in a country-customized language. When they are displayed, the texts generated can be modified by using a Text Editor before being printed.

In spite of this potential, texts currently produced by this first version of the generator still have a lot of problems that need to be solved before it can really be considered to be a useful explanation facility. Let us examine them in detail.

6.1. Gap between strategical and tactical generation

As we said in the previous paragraphs, the *plan* establishes the structure of the text: primitive acts to be communicated and rhetorical relations between text spans. The plan is then passed to the surface generator, to be transformed in a natural language text. The *ATNs* translate each primitive act and each rhetorical relation into a word, a phrase, a clause or a sentence. The consequences of this separation between tactical and strategical phases of generation are the following:

- the text contains several *repetitions* of words or fragments of phrases;
- the *organization of paragraphs* is not always ideal. This does not facilitate understanding of the text.

Some examples of these problems can be found in Fig. 9, where the explanation text provided by a doctor in our preliminary experimental study at the University of Reading is compared with the computer-generated text for the same case.

PATIENT Comments to the drug prescription for Mr Fictif	NURSE Mr Fictif is a 62 years old man; he is overweight.	GENERAL PRACTITIONER As you certainly remember, Mr Fictif is a 62 years old man; he is overweight.
You have been diagnosed as suffering from a mild of what we call 'angina pectoris', that is a spasm of chest resulting from overexertion when heart is diseased.	He is suffering from a mild form of angina	He is suffering from a mild form of angina
In addition, you have elevated cholesterol.	and he has got elevated cholesterol.	and he has got elevated cholesterol.
	He has been admitted into the Hospital in emergency because of a cardiologic disorder.	He has been admitted into the Hospital in emergency because of a cardiologic disorder.
To solve this problem, there are two drugs I would like you to take.	I've written him up for two drugs.	I have written him up for two drugs.
The first one is Aspirin, which is an analgesic, that is it relieves the pain. I have prescribed it to cure your angina.	The first one is Aspirin to cure angina.	The first one is Aspirin to cure angina,
The only problem is that this drug can be associated with some side effects.	This drug can be associated with some side effects.	although, as you know, this drug may have some side effects.
The first one is bleeding, that is a slight haemorrhage;	The first one is severe bleeding	One of them is severe bleeding
it can be a serious side effect and occurs infrequently.	which is infrequent.	which occurs infrequently.
In the unlikely event that this occurs, you should stop treatment.	If this occurs, I recommend that you stop treatment.	In this case, it is advisable to stop treatment.
The second one is allergy; it can be a serious side effect and does not happen commonly.	The second one can be severe allergic reaction	The second one can be severe allergic reaction
It occurs especially when the patient is suffering from what we call 'atopic disease', that is a hypersensitivity to substances having a basis of hereditary predisposition.	which is infrequent. It occurs especially when the patient is suffering from atopic disease	which is infrequent and occurs especially when the patient is suffering from atopic disease.
In case you notice that this problem occurs, you should stop treatment.	If you notice it, I recommend that you stop treatment.	If this occurs, it is advisable to stop treatment.
You have to take this drug by mouth. Take with water at regular intervals. If you miss a dose, do not modify rhythm or quantity of the following ones.		
This drug may interfere with other ones. One interaction is with Vitamin K inhibitors; this could cause increased bleeding risk.		
Regarding administration of this drug, the treatment steps are the following; the first one is an attack treatment.	The treatment steps are the following; first an attack treatment	The treatment steps are the following; first an attack treatment
The dosage is 1 g in 2 intakes for 1 day on 13.9.93. You have to take one tablet of 0.5 g at 8 and one tablet of 0.5 g at 20.	with a dosage of 1 g in 2 intakes for 1 day, the 13.9.93. 0.5 g at 8 and 0.5 g at 20.	with a dosage of 1 g in 2 intakes for 1 day, the 13.9.93. 0.5 g at 8 and 0.5 g at 20.
The second one is a maintenance treatment.	Second, a maintenance treatment	Second, a maintenance treatment
The dosage is 0.5 g in 1 intake for 7 days from 14.9.93 to 20.9.93. You have to take one tablet at 20.	with a dosage of 0.5 g in 1 intake for 7 days from 14.9.93 to 20.9.93 at 20.	with a dosage of 0.5 g in 1 intake for 7 days from 14.9.93 to 20.9.93 at 20.
This are the main treatment I wanted to give you. If you need more information, please do not hesitate to ask me directly.		

Fig. 8. Comparison of texts addressed to a patient, a nurse and a general practitioner.

CASE STUDY N.3 : angina in a 73 years old woman

Doctor's explanation n.2	**Computer generated text**
All the tests show that you have three separate problems.The first is the chest pain, which we know to be angina,	You have been diagnosed as suffering from a mild form of what we call 'angina pectoris',
that is heart pain, which is very similar to the cramp you get in your legs if you walk too far. It's exactly the same with the heart. When it does too much work, it can't cope because it can't get enough oxygen.	that is a spasm of chest resulting from overexertion when heart is diseased.
.....	
This combination of problems necessitates treatment with some special tablets that I'm going to give you, and I'll briefly describe these to you.	To solve this problem, there are three drugs I would like you to take.
The first tablet is simple ASPIRIN,	The first one is ASPIRIN
	which is a analgesics that is, it relieves the pain.
which has been shown to help people with angina and reduce the risk of you having a heart attack.	Therefore, I have prescribed it to cure your angina.
	The only problem is that this drug can be associated with some side effects. The first one is bleedings; it can be a serious side effect and it occurs unfrequently. In the unlikely event that this occurs, I suggest you stop treatment. The second one is allergy; it can be a serious side effect, and doesn't happen commonly. It occurs especially when the patient is suffering from what we call 'atopic disease', that is a hypersensitivity to substances having a basis of hereditary predisposition. In case you notice that this problem occurs, you should stop treatment.
	.....
That's just one tablet in the morning	You have to take this drug by mouth. Take with water at regular intervals. If you miss a dose, do not modify rythm or quantity of the following ones.
	Regarding administration of this drug, you should take one tablet of 150 mg at 8

Fig. 9. Comparison of a physician explanation text with the computer-generated one.

The explanation refers to the case of angina in a 73-year-old woman: the first sentence in paragraph 1 of the doctor's text gives a synthetic description of the whole case, by mentioning the *number* of patient diseases ('You have three separate problems').

The same happens in the paragraph 3, where the treatment is synthetized by saying that the three drugs are all in tablets. The computer generated text mentions the diseases one at a time (the first one in paragraph 1, the second one from paragraph 10 on, etc.) and only evokes the number of drugs in the prescription, at paragraph 3. The description of side effects at paragraph 6 in the computer text was omitted from the doctor explanation, and has been inserted in our text because the metaplanning was acting in the '*Advocate*' mode. However, this part of the text contains several repetitions. Phrases such as: 'it can be a serious side effect', 'it occurs unfrequently', and so on might be said once for both side effects rather than being repeated for each of them. We are currently studying a solution to this problem by adding paragraph-level ATNs, in order to check changes in the focus of discourse and repetitions of communicative acts, and to generate phrases accordingly.

6.2. Use of standard knowledge and data sources

A further limitation of the texts that we can generate is due to the following factors:

- crucial data may not be available in the OPADE KBs;
- crucial links between data may not be available;
- data values may be incompatible with adaptive phrase-building or with the two users' information needs.

Let us look again at Fig. 9. Doctor's explanation about the disease (paragraph 2) is patient-specific, and is based on the technique of 'definition by analogy'. This corresponds to one of the suggestions in Diana Forsythe's work. The computer text tries to do the same, but the disease explanation that we find in the KB is formal and general. The difficulty of using predefined knowledge sources is evident also in the paragraphs 4 and 5 of the computer text. Here, there is a contradiction between the general indication of the Therapeutic Class of Aspirin ('to relieve the pain') and the specific use made of Aspirin for that patient ('to cure your angina'). In the same paragraph, the phrase 'I suggest you stop treatment', is not correct, and it is difficult to find a unique, fixed template that can be combined effectively with different, prestored KB fragments such as 'hospitalisation' and 'stop treatment'.

6.3. Adaptation of RRs

In the present version of the Generator, linguistic markers associated with Rhetorical relations are fixed. However, this does not seem to correspond to what the Speaker does in naturally occurring explanations. According to the analysis of our texts by a linguist, *'if the Speaker judges that the background information is well known to the Hearer, he will only give the proposition which, when combined with what the Hearer already knows, provides an explanation. This is most clearly seen in doctor–colleague and doctor–nurse texts....'* [7,16]. We can give several examples of this adaptivity in the use of linguistic markers, from texts collected in our experiments:

(i) '*and*' is employed frequently as a connective for a '*Solutionhood*' which would normally require a phrase such as 'to solve this problem'. For example: 'I have examined Mrs Smith and I think she has got angina with some congestive cardiac failure **and** I've prescribed some Frusemide...'

(ii) no connective at all is employed, sometimes, for a '*Justification*'. For example: 'He is a bit of a worry, the tests show he has got pulmonary TB', or: 'This 73-year old lady is coming into hospital, she suffers from heart failure, angina and diabetes. I've written her up for Aspirin,....'

(iii) no infinitive clauses and no connectives are used, sometimes, for a '*Purpose*': 'The fourth tablet I want to give you is a special tablet, it's a fairly new tablet, that's just come on to the market in the last few years and it is the most effective tablet that we have for treating heart failure. It's a very powerful tablet and can cause your blood pressure to drop.'

As we mentioned, as Rhetorical Relation markers have the role of fostering the Hearer understanding of the text by increasing coherence between related text spans, omission of markers is a consequence of the fact that the Speaker believes that the Hearer has no difficulty in understanding the text. The use of markers might therefore be conditioned to either the global characteristics of the hearer (a 'patient', a 'nurse' or a 'general practitioner') or, more refinedly, to the specific knowledge of single items, represented in the indirect User Model.

6.4. Evaluation studies

At the present stage, and for the reasons mentioned in the previous paragraphs, the main problems of our explanations rely on the 'tactical' (surface generation) component. Our future work will be devoted mainly to refining the style of the text in order to make it more fluent, less repetitive and more tailored to the indirect user situation. At the same time, we wish to test alternative hypotheses about information content and order in the explanations, that is about the planning process. Evaluation of alternatives will be made by a new set of experimental studies, where samples of potential recipients of these explanations will be required to express preferences about different computer-generated texts.

Appendix A. The scenarios employed in the physician exaplanation study

A.1. Case 1

A 55-year-old woman goes to her general practitioner at her daughter's insistence. The patient herself has not noticed anything wrong, but her daughter thinks she has become generally 'slower'. In addition, her voice has become much deeper and more husky. Amongst other things, examination reveals a slow pulse and almost complete loss of hair from eyebrows. The patient has been diagnosed as suffering from hypothyroidism.

A.2. Case 2

A 44-year-old alcoholic man lives on a poor diet in a cold and damp lodging house. He now complains of a persistent cough, occasionally bringing up blood. His appetite is

poor, and he thinks he is loosing weight. Examination and chest X-rays, followed by other tests, confirm the diagnosis of tuberculosis.

A.3. Case 3

A 73-year-old woman complains of severe central chest pain after walking about a quarter of mile, sooner in cold weather. The pain radiates to both arms and the throat, and is very unpleasant. It gradually subsides after 5–10 minutes rest. In addition, she is breathless on exertion and on lying flat in bed. Her everyday activities are severely restricted and she rarely feels well. Tests reveal that she has high blood cholesterol and diabetes. She is at least 10 kg above her ideal weight. The patient has been diagnosed as suffering from angina, combined with hearth failure and metabolic problems.

A.4. Case 4

A 62-year-old man complains of central chest pain, radiating to the left shoulder and arm. The pain occurs on climbing steep slopes or more than two flights of stairs. he does not have to do either of these frequently, and can walk at least three miles on the flat without difficulty. If the pain does occur, it will disappear after a couple of minutes rest. He is otherwise very well, and is not unduly worried about his symptoms. Tests confirm coronary disease but no other specific clinical syndrome abnormalities. the patient has been diagnosed as suffering from exercise induced angina in a mild form.

References

[1] D.C. Berry, T. Gillie and S. Banbury, What do patients want to know, an empirical approach to explanation generation and validation, *Expert Syst. Appl.* (1995) in press.

[2] B.G. Buchanan, J. Moore, D. Forsythe, G. Banks and S. Ohlson: involving patients in health care, explanation in the clinical setting, in: *Proceedings of American Medical Informatics Association* (1993) 510–514.

[3] A.J. Cawsey, K.A. Binsted and R.B. Jones, An on-line explanation of the medical record to patients via an artificial intelligence approach, *Health Care Comput.* (1995) in press.

[4] F. de Rosis, F. Grasso, D. Berry and T. Gillie, Mediating between hearer's and speaker's views in the generation of adaptive explanations, *Expert Syst. Appl.* (1995) in press.

[5] D.E. Forsythe, Using ethnography in the design of an explanation system, *Expert Syst. Appl.* (1995) in press.

[6] E.H. Hovy, Automated discourse generation using discourse structure relations, *Artif. Intell.* 63 (1993) 341–385.

[7] R. Ingham, *Achieving coherence: code and inference in the analysis of texts*, OPADE Working Paper, 1993.

[8] H.B. Jimison, L.M. Fagan, L.D. Schachter and E.H. Shortliffe, Patient-specific explanation in models of chronic disease, *Artif. Intell. Med.* 4 (1992) 191–205.

[9] W.C. Mann, C.M. Matthiesen and S. Thompson, Rhetorical structure theory and text analysis, in: *Information Science Institute, University of California, Report 89-242* (1989).

[10] M. Maybury, Communicative acts for explanation generation, *Int. J. Man Mach. Stud.* 37 (1992) 135–172.

[11] K. McKeown, Discourse strategies for generating natural language text, *Artif. Intell.* 27 (1985).

[12] P.L. Miller and G.D. Rennels, Prose generation from Expert Systems, *AI Magazine* fall 1988.
[13] J. Moore and C: Paris, Requirements for an expert system explanation facility, *Comput. Intell.* 7 (1991) 367–370.
[14] J. Moore and C. Paris, Exploiting user feedback to compensate for the unreliability of user models, *User Model. User Adapt. Interact.* 2 (1992).
[15] OPADE, Necessary input for a Computerized Drug Prescription System. Deliverable No. 5, *AIM Main Phase Project A2027* (1992).
[16] OPADE. User Interface for Computerized Drug Prescription System. Deliverable No. 4, *AIM Main Phase Project A2027* (1992).
[17] E. Rich, Stereotypes and user modeling, in: A. Kobsa and W. Wahlster, eds., *User Models in Dialog Sytems*. (Springer, Berlin, 1989).
[18] E.D. Sacerdoti, Planning in a hierarchy of abstraction spaces, *Artif. Intell.* 5 (1974) 115.
[19] J.M. Shumway, A.I. Jacknowitz and M.A. Abate, Analysis of physicians', pharmacists' and nurses' attitudes toward the use of computers to access drug information, *Methods Inform. Med.* 29 (1990) 99–103.
[20] W. A. Woods, Transition Network Grammars for natural language analysis, *Commun. ACM* 13 (1970) 591–606.