EXPLANATION CONSTRAINTS IN A MULTI-AGENT ENVIRONMENT

Authors: PIQUEMAL-BALUARD C., GLIZE P.
Address: IRIT - Université Paul Sabatier
118 route de Narbonne
31062- Toulouse cedex
France
Email: Piquemal@irit.fr
Phone: 61556611 poste 7253
Fax: 61556258

ABSTRACT:

The explicit separation of the knowledge and the system that uses them (the inference's engine) offers to expert systems dialog abilities forbidden to programs built on traditional data processing technics.

This capacity of dialog and self-explanation must be used and developed because it represents one of the main advantages of this kind of system. Futhermore the qualities of user friendship involved in the installation of a mechanism of explanation, this allows particularly the systems to be validated. The justification of the results increases in a large way the confidence accorded by user to the solution found.

We present a study of the explanation in the context of a multi-agent system with one unique explanatory module. Because of the constraints of this kind of software, the explanation cannot be based on the addition of specific explanatory knowledge on each form that can be, but only on syntactic analysis of the knowledge and user model. We analyse the consequences of this approach and we give examples on each part in the medical field.

Keywords: explanation, user models, multi-agent systems, man-machine communication, knowledge representation.

I- INTRODUCTION

The addition of an explanatory mechanism into a Knowledge Based System (KBS) is necessary to provide an intelligent description of the expert knowledge and of the reasoning process done with it. Its integration is facilitated by the architecture of KBS where the explicit separation between knowledge and inference engine gives dialog abilities unattainable from classical programs [14]. An explanatory mechanism increases "user friendship" and confidence in the proposed solution. The first works about explanation [6], [16], [17] have provided to the user the inferences leading to a given result. Unfortunately, the lack of discrimination between the message parts is prejudical to the clearness and of inability to explain the terminology.

The following works done on explanation have given much different results. The first reason for that is the adaptation of the explanatory mechanism to the structure of the KBS (its knowledge representation, its inference engine architecture characteristics, and even the expertise field). Finally, it's very difficult to formalise the explanation process, which must answer to the following points [1]:

- To whom is the information addressed?
- What kind of information?
- The granularity of the information?
- In what form of presentation?
- When can it be available?

II- EXPLANATORY MESSAGE ELABORATION

If the elements of an explanatory message are extracted from the knowledge, they must be chosen from some criteria. The sending informations are selected according to their relevance (concepts are more essential than others) and the user nature.

2.1 The knowledge classification by addition of explanatory informations

In this method, the role of each knowledge element is defined by the expert in adding its justification into the knowledge base. The notion of structural knowledge developed by NEOMYCIN consists in the classification

of the knowledge domain in many linked groups. A specific justification method is associated to each group [6].

ADELE [13] put in two separate modules the surface knowledge and the deep knowledge. The deep knowledge module contains theories and resolution knowledge that are specific to the domains. They allow in particular the description of links between objects, between objects and object's classes, leading to the construction of semantic networks.

These "ad hoc" informations are of multiple kinds. They can be written in a message form to facilitate their using, or in belonging to a class.

But, this method set the problems of the memory space and of its updates when the expert knowledge base is modified. Then, explanations in the multi-agent system SYNERGIC [15] are build without adding "explanatory" informations. We analyse further the associated constraints and limits.

2.2 The user modelling by autopresentation

The explanatory message is sent to a particular user that has a given competence when the session starts. His competence level and his expectations are important criteria to build a pertinent answer.

The easier way (and the most frequently used) to know that is to begin a session with a user autopresentation. It may generally consist in asking for his competence domains and the desired degree of explanation development. Inside the Knowledge base, each concept (rules and objects) is evaluated by the expert. This value represents their complexity and understanding degree. The explanatory message contains only rules or objects having a value situated between the domain competence and the explanation degree [17].

BOUZEGHOUD [3] constructs the explanation according to the initial competence degree of the user. For an expert, it consists only in editing the rules without comments, to improve the knowledge if necessary. A designer concerned in the application modelling, receives the closely related rules with comments or graphics. Whereas all the underlying knowledge is edited for a student.

2.3 The method of dynamic user modelling

However, we are never sure of the user judgement: he could underestimate himself. In addition, the same user cannot understand all the concepts of a knowledge base with the same degree. A question to the user about each of them is unrealistic.

For these reasons we have chosen to use a method of dynamic modelling in SYNERGIC. The explanatory mechanism implicitly judges the user competence during the dialog. It has been already employed in another software [4].

III- THE EXPLANATION IN THE MULTI-AGENT SYSTEM SYNERGIC

3.1 Context presentation

SYNERGIC is a multi-agent system where the solution of the problem is the result of a team co-operation. Each team member is an agent, classical procedures considered by SYNERGIC as a black box, or a knowledge module that can be explained.

Each knowledge module is hierarchically ordered in levels. Each level has the metaknowledge of the lower level. The upper level defines the general strategy to solve domain problems (the surface knowledge): it is the patient interview and the clinical examination in medicine. The lowest level contains the detailed informations necessary to solve the high level tasks (deep knowledge). It is the symptom's study realised by the medical expert to find the potential diagnosis [15].

3.2 The explanatory informations about the domain

An expert system shell ought to explain the reasoning process in many diverse domains. In the same way, the multi-agent system shell SYNERGIC could possess many expertise's fields in a same application. Thus, an explanatory mechanism must be able to work in a domain in an independent manner.

Usually, a specific explanatory module is associated with each knowledge module. The explanation built with this method is thus very closed to the domain, but many problems occur. Essentially, the memory space and the update, as we have seen in 2.1. On top of that, when the explanatory knowledge is included in the knowledge base, it is difficult to do the difference between the solving knowledge and the explaining knowledge. Agreing the maxim "all erudite is not a good pedagogue" [13], we think that a good field expert is not necessarily able to create the explanatory knowledge.

Thus, we have chosen in SYNERGIC a method that allows a real domain independent explanation. Because it cannot use the real meaning of the employed terminology, it finds them with a syntactic analysis (parsing) of the knowledge base, to create an explanatory message in natural language.

The answer elaboration imposes to select the informations to send. The main criterion of this pruning is the discovery of the event role inside the module and during the solving process.

An event role is understood with its internal characteristics (or structural informations) and its links with the other events of the domain (or contextual informations).

3.3 The structural informations to classify the knowledge

The first explanatory information is derived from the event decomposition in each of its elementary components: object, attribute, link, value. The citations of **palpation, abdomen, causes** and **pain** are the beginning to explain the event **palpation of abdomen causes pain. Temperature** is a property of the object **patient** for the event **the temperature of patient> 37.2** while **37.2** is a possible value for the entity **temperature of patient.** These indications are helps to define the explanation degree associated to an object.

The following indications are derived from the rules containing the event. An event only situated in premises is an external information obtained with a user dialog. Thus he knows the event existence and its value, and it is not necessary to develop it in the explanatory message.

The last indication is the level of the event. The hierarchical structure of knowledge bases into levels allows to solve the problem of strategic knowledge dilution.

3.4 The contextual informations to classify the knowledge

The event role is completed by an analysis of its links with the remaining knowledge. But the causal links (production rules) can have many meanings: implication, rewriting, procedure call [3].

A directed syntactical analysis is able to furnish these meanings without adding explanatory informations. For instance, common entities between different condition sets inferring an event, is spelled necessary conditions (but not sufficient).

The influence degree of an event on another can thus be evaluated. It is meaningfulless when events are related to the same entity: each value role or attribute's role of an object can be discovered with this method.

3.5 The method of dynamic user modelling

Knowledge about user (interest subjects, domain competence,...) are memorised inside the user model. Even if the explanation goal is not teaching, he can improve his knowledge. Thus, a better understanding of concepts, and the discovery of another must be echoed in the user model: it is the method of dynamic user modelling.

The explanatory mechanism evaluates implicitly the user, without initial autopresentation. This method increases the "user friendliness" and it is a gain of time. The main pieces of information are the questions of the user. An event has a privileged treatment when it is the subject of many questions.

The **dialog trace** is thus necessary for the dynamic modelling. The dialog trace allows also to repeat the same explanations and to detect implicit questions in the user discourse. An increasing interest or a lack of understanding of the explanation could be the underlying reason. The study of the user questions is a good help to find his subjects of interest; and the study of the answers improves their relevance.

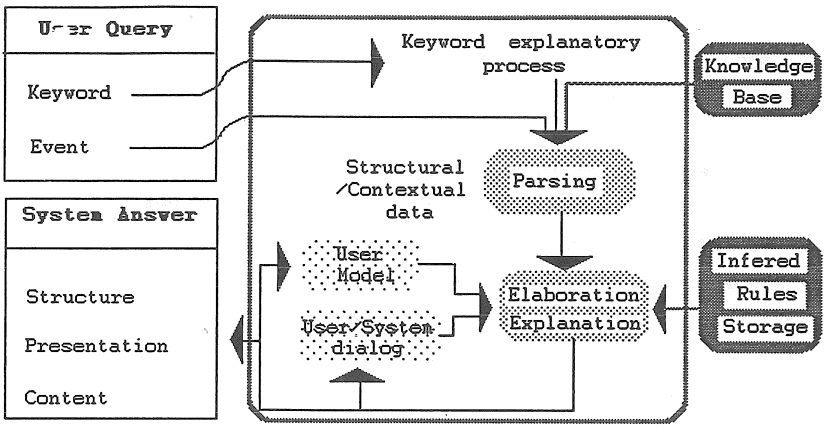
3.6 The analysis granularity

The study of all the links between the events leads usually to a great amount of information and it is time consuming. To reduce this amount of structural and contextual informations, the explanatory mechanism drives the research about an object or an entity. The importance of the event and its links with the other entities of the domain are the criteria used to limit the search.

IV- THE EXPLANATION PROCESS OF SYNERGIC

This process is based on question/answer principle. A question is composed of a keyword and the explanation subject (the event). The user can set three kinds of queries. One is associated to an explanation about the knowledge (WHAT), the others are used to understand the process reasoning during a session (WHY) or at the end (HOW). Each of them owns a specific standard answer, adaptable at user and knowledge.

Fig.1 Interactions between the elements of the explanatory mechanism



4.1 Example of the explanatory method associated to the keyword WHY

In order to show the explanation process, we develop the answer elaboration to the question **WHY an event?**. In our case, the user model is empty because this question starts the dialog.

The event is located in first position in the reasoning process. The answer is added with informations about its real role in the knowledge. Their location in the answer and their granularity depends on the amount of the previously edited informations.

4.1.1 Location of the event in the reasoning process

The short term explanation is deduced from the last fired rule having this event in its premise. It is found in the trace and edited in a simplified natural language form.

The long term explanation consists in finding the context of the inference of this rule. The pruning is made with the two previous criteria: the user model and the place of the event in the knowledge. This pruning allows to catch the user eye on the most significant results to obtain.

4.1.2 The event meaning

The previous answer is completed with short definitions of the event and the entities contained in the justification. An event not deduced by the knowledge (an initial fact or a user fact) is situated in the domain. The

effects of its valuation on the knowledge base are edited (except the last fired rule).

In order to better understand justification, the premise of the rules not fired and containing it are written. The consequence of its determination on the reasoning process are then listed.

4.2 Example of explanation in medecine

The example of a question WHY on TELEMAC is analyzed. TELEMAC [15] is an application of SYNERGIC in the medical domain developed in cooperation with the Centre Hospitalier of Toulouse. The knowledge base currently active is the "thoracic pain" which has 250 rules. Among the 19 possible diagnosis, it tries to find the plausible pathologies for a patient having an acute thoracic pain.

At this stage, the inference engine has no inferable rule, and the man-machine interface asks the user about the pain topography to continue the reasoning process. But the user wants to know the goal of this question before his answer. Thus, his first need of explanation in this session is: **WHY the topography of pain?**

This sentence begining with the keyword WHY leads to the loading of the explanation module. Without any other dialog the system writes:

The search of topography of pain begins the study of features of pain. It should be followed by the determination of irradiation, starting and type.

It allows the deduction of objects to probable or uncertain. An object is member of the set: (colonopathie, acute-appendicitis,..).

In long term, it leads to the initialization of probability of diagnosis allowing the plausible deduction of diagnosis of abdominal = realised.

This result is composed of three parts, precisely ordered. The first part concerns location, in the reasoning process, of the explanation subject (the event) for the landmark of the user. The second part defines the event role independently of the session (the principle of explanation-based example). The last part gives a more general view of the reasoning: the final result and perhaps further questions.

After the message editing the user can restart the dialog. The answer and the questions modify dynamically the user model.

4.3 First message part

It refers to the last fired rule in backward chaining. It locates where the inference engine stops.

```
/* RULE 17 */    If the topography of pain
                 and the irradiation of pain
                 and the  starting of pain
                 and the typeof pain
                 then the features of pain
```

Each premise is sent to define the immediate context of the event. It explain why the system is stopped, and the others informations missing to fire the rule.

Because all the premise concern the same object **pain**, they are written in a shorten form: only the attributes are given. The verb **begins** is employed because the event starts the rule. The terminology used by the explanatory mechanism is therefore particularly chosen to increase the answer precision. The mechanism possesses also a set of predefined sentences to adjust the responses to the knowledge and situations.

4.4 Second message part

In order to locate the event in the knowledge, its function and its impact on the remaining knowledge are presented. A syntactic analysis centered on **topography of pain** allows the explanatory mechanism to find that it is only able to deduce **diagnosis of X = probable or uncertain.** X is a member of the object set having **probability** and **diagnosis** as attributes. These objects are very significant for this reason and justify their presence in the explanatory message.

Nevertheless, the explanation module is unable to discover that each X is one the 19 possible diagnosis. The enumeration of the set overcomes this problem.

4.5 Third message part

The explanation ends with the fact justifying the existence of **topography of pain.** It was selected from the set of rules leading to the conclusion. For instance, all the entities member of this set (as **features of pain**) used to order the resolution are excluded. The conclusion of the last fired rule in backward chaining is also excluded. The complete syntactical analysis of the related rules leads to the great relevant entity **probability of diagnosis.**

The syntactic analysis reveals also that this entity is valuated and has an initial value, obtained at the end of the path starting with the event. For this reason, the explanatory mechanism employs the term initialization.

V- CONCLUSION

The integration of an explanation mechanism in a multi-agent system comme up against the difficulty of knowledge diversity. In effect, the explanation is usually domain dependent. More, an expert uses a specific domain terminology. These reasons leads to think that a knowledge base must be **autosufficient for explain**, allowing a really domain independent explanation mechanism. Our method and its results largely justify this principle.

During the knowledge acquisition process, the expert generally includes justifications. During the transcription phase of the interviews in formalized knowledge by a cognitician, these justifications are generally suppressed because they are not useful for problem solving. We think that the shared knowledge notion in a multi-agent system, allows us to retrieve some of these explanatory knowledge in others modules. When we lack of justification in the currently module, another related module would be able to give supplementary informations. It is our future work for the explanation in a multi-agent system.

BIBLIOGRAPHY:

[1] BOURI M,DIENG R, KASSEL G, SAFAR B: Systèmes à base de connaissances et Explication PRC-GDRIA, 3ième journées nationales, 5-7 Mars 1990

[2] BOURI M,DIENG R, KASSEL G, SAFAR B:- Vers des SE plus explicatifsPRC-GDRIA, 3ième journées nationales, 5-7 Mars 1990

[3] BOUZEGHOUD M, METAIS E: La justification des résultats et l'explication du raisonnement dans le système expert SECSI. CIAM 86

[4] BREZILLON P: Intervention de l'utilisateur dans les explication Deuxièmes journées explication du PRC-GDR-IA du CNRS Sophia-Antipolis 17-19 juin 1992

[5] CHOURAQUI E: Un modèle formel de caractérisation de l'évolution des connaissances. Sixième conférence Canadienne sur L'Intelligence Artificielle. Montréal, 1986

[6] CLANCEY W. J, HASLING D. W., RENNELS G.: Strategic explanations for a diagnostic consultation system.Int. J. Man-Machine studies 20 1984

[7] DAVID J. M, KRIVINE J. P: Explaining Raisoning from Knowledge Level Models.ECAI 1990

[8] DIENG R: Explications à partir d'arbres de raisonnement étiquetés par des relations. Journées Explications du PRC IA,Mars 1989

[9] ERIKSON R, JOHANSSON A. L.: Neat explanation of proof trees IJCAI 85

[10] FININ T, DRAGER D.: GUMS: A general User Modelling System Sixième conférence Canadienne sur L'Intelligence Artificielle. Montréal, 1986

[11] GILBERT N, BUCKLAND S, FROHLICH D, JIROTKA M, LUFF P: Providing Advice through dialog. ACAI 1990

[12] JIMENEZ DOMINGUEZ C: Le problème des connaissances implicites pour l'explication Colloque IA Caen 14, 18 Septembre 1987

[13] KASSEL G: Système Expert de seconde génération et explication du raisonnement. Avignon 1989

[14] LEMAIRE B,SAFAR B.: Some necessary features for explanation planning architectures: study and proposal, in proceedings of AAAI'91 Workshop on Explanations. Anaheim, USA, July 1991

[15] PUYET P, CARPUAT B, GLIZE P, PUJOS M, MAGUES J.P: TELEMAC: a multi-agent system for aided-decision and diagnosis approach. AAAI Springs Symposium of Artificial Intelligence in Medecine, March 1992

[16] SWARTOUT W. R: XPLAIN: a system for creating and Explaining Expert Consulting Programs.AI 21(1983), 285-325

[17] WALLIS J. W, SHORTLIFE E. H.: Customized Explanations Using Causal Knowledge, Rule Based Expert Systems. Addison-Wesley, 1984