WEBEXPERT – TOOL FOR KNOWLEDGE ACQUISITION BY WEB

Scholarship holder:Raphael Luiz Nascimento Advisor:Anita Ma. da Rocha Fernandes, Dr.
raphael@inf.univali.br   anita@inf.univali.br

UNIVALI – University of Itajaí Valley
Uruguai Street 458, Itajaí city , Santa Catarina State - Brazil

1. Introduction

1.1 Contextualization

Knowledge acquisition for systems based on knowledge can be called knowledge engineering, according to Feigenbaum (1980) who used the term to describe the reduction of a wide context of knowledge for a set of facts and rules. The term Knowledge Engineer started to be used to denote the responsible person by development of such systems.

The basic function of Knowledge Engineering is to play the role of mediator between the expert and the Knowledge Base, to extract the knowledge from the expert, to codify the same in contribution with the expert to reach an acceptable performance.

However, Knowledge Engineer in role of intermediary between the expert and the systems based on knowledge can face many problems, concerning the way as each expert reasons (Gaines, 1987). The computer appears as an excellent tool to help the expert to structuralize the knowledge domain and, the researches in knowledge acquisition have centralized the development of knowledge acquisition tools based on computer (Boose and Gaines, 1988; Boose, 1989). Many tools projected to be used directly by the expert with the minimum of intervention of Knowledge Engineer; they have emphasized the easiness for visualization of the domain concepts. The objective of systems such as PLANET (Shaw and Gaines, 1983; 1986; 1987), ETS (Boose, 1984; 1986), MOLE (Eshelman et al, 1987), SALT (Marcus, 1987), Kitten (Shaw and Gaines, 1987b), Knack (Klinker et al, 1987); Kriton (Direch, Ruhmann and May, 1987), KKSSO (Gaines, 1987b. Gaines and Shaw, 1986) is to fulfill the process of knowledge acquisition and to transfer it to the systems based on knowledge.

However, the tools of existing knowledge acquisition are widely based on PC'S and graphical Workstation, and the use of the same in a distributed community involves mobile software. The process of development and test of the models based on knowledge through a distributed community would be used sufficiently if the networks of great areas could be used to coordinate the activities in different sites. The initial objective of the paper that is presented here is to use the web to support the distributed knowledge acquisition by the development that operates in web. Then the technique of Personal Construction Psychology was used.

The option for this technique is due to Kelly's Personal Construction Psychology (1995) has been used for knowledge acquisition since the beginning of the development of systems based on knowledge (Shaw and Gaines, 1983; Boose, 1984) and has been refined throughout the years to support the increase of the complexity of knowledge structures (Gaines and Shaw, 1993).

1.2 The problem

Despite the Knowledge Engineer has little knowledge domain that the expert, the communication problems prevent the transference process of knowledge for a program. The vocabulary used initially by the expert to speak about the domain with a layperson is generally inadequate for solution of problems, and then, the Knowledge Engineer and the expert have to work together to refine it. One of the most difficult aspects for the Knowledge Engineer is to help the expert to structuralize the knowledge domain, to identify and to legalize the concepts of the domain (Haynes – Roth, Waterman, Lenat, 1983).

The development of the systems based on knowledge involves knowledge acquisition starting from various sources generally different geographically. The sources include books, articles, manuals, videos with the performance of the expert, translation of protocols and interviews, and interactions man-machine with the expert. The time of the expert is generally a scarce resort and experts are, generally accessible in different places, particularly in international projects. The methodologies of knowledge acquisition, as well as the tools, have been developed to deal these problems by the use hypermedia systems to manage a wide amount of heterogeneous data, interactive graphical interfaces to present knowledge models in an understandable way to the experts, systems of

fast prototype production to test the models in operation, and systems of models comparison to call attention for the variations among experts.

Another important point in knowledge acquisition is the handling of incoherence. Depending on the form as the knowledge is acquired, errors of acquisition can have. These errors can result of own knowledge nature, as in data gotten through sensory that are subject to noise, or can be generated by the existing human interface between the real world and the representation system. Techniques had been developed to prevent errors of acquisition in what type of waited knowledge is defined. These techniques are common to the knowledge representation systems and to database management systems. On the other hand, a knowledge base can be examined periodically with the purpose to detect incoherence eventually introduced in the acquisition process. This method is limited by the fact of reasonable expressive representation languages do not count on full procedures of known verification. Finally, it must be observed that the adequacy of representation formalism to the type of knowledge of real world to be represented is basically for the efficiency of the acquisition process (Bittencourt, 1998).

The necessity of development of a knowledge acquisition tool is clear based previous displayed on, however several experts have different ideas and understandings for the same domain.

An application area of Personal Construction Psychology, which has been useful in many areas and professions, is social knowledge process modeling through comparison of developments and terminologies of different individuals in the same domain. The comparative analysis of repertory grid to become such similarities and differences exposed was one of the first computational applications of Personal Construction Psychology (Shaw, 1979, 1980) and the involved techniques have been presented throughout the years to supply more efficient tools of analysis (Shaw and Gaines, 1991a, 1991b, 1993).

Thus, this article presents the development of a knowledge acquisition tool for web, based on Kelly's Personal Construction Psychology, and it has as case study the areas of oceanography, biotechnology and environmental engineering.

So, some basic stages must be fulfilled: the study of Kelly's Personal Construction Psychology to establish the criteria to be used in repertory grid; establishment of the structure of database that will give support to the system; evaluation of interface diagramming in order to be friendly to the user originating from any area of knowledge; establishment of criteria for evaluation of prototype performance, during its case study in some domains: oceanography, biotechnology and environmental engineering.

1   Theoretical basis

A bibliography search about Personal Construction Psychology, knowledge acquisition, software utility engineering, PHP and MySQL was necessary to develop the Web Expert.

George Kelly created the Personal Construction Psychology in 1955 (Kelly, 1995). In a computational context, Kelly's approach is attractive because it directly deals with representations and algorithm, and its representation in terms of modern intentional logic leads to a model of human psychology process.

The methodology used by Kelly for the extraction of concept structures is the "repertory grid". Gaines and Shaw (1980) suggested that the repertory grid would be a useful technique for the development of systems based on knowledge. Many tools for knowledge acquisition have incorporated repertory grid as an extraction technique (Boose and Brash, 1987; Dierich, Rulmann and May, 1987; Garg-Janardan and Salvendy, 1987; Shaw and Gaines, 1987; Ford et al, 1990).

Regarding to the knowledge acquisition, it is not a phase of the systems development based on knowledge, but a component that enters in all phases, and since the definition of the problem it has persisted in the maintenance. Rook and Croghan (1989) show how the acquisition can be faced as being a component of engineering structuralized in knowledge systems. Therefore, it is also a very delayed task, which demands constant devotion of the expert and the knowledge engineer.

The knowledge acquisition typically involves actions to congregate information of one or more human experts and/or of documentary sources, arranging this information in some way and, then, translating it for one clear form by machine. Second (Breuker and Wielinga), knowledge acquisition is the process to transform data of expertise into implementation formalism.

The usability engineering was studied due to necessity of developing a friendly interface for the user. The development of Web Expert interface is based on Basteim & Scarpeen's software ergonomics concepts (1994).

To support the implementation of database and knowledge, as well as the interface by web, it was necessary a bibliographical study, beyond the study about PHP, MySQL and Flash.

PHP is a programming language called serve-side script to create dynamic sites. Using PHP, the direct interaction of the user with the site can be made, through forms, counters of access, statisticians, or creating

It is also possible to interact with database and existing applications in the server, with the advantage of no display the source code for the customer. This can be useful when the program is dealing with passwords or any type of confidential information.

Another advantage of PHP referring to other languages of this type is that, it is totally free, or better, we only must have an apache server to use it PHP makes interactions with several database and one of them is MySQL that is also free.

MySQL is a server of multi user database; multitask that works with one of the most popularized languages of manipulation of data of the world. According to Fernandes (2000), it is had as features of MySQL: a) it supports different platforms: win32, Linux, FreBSD, Unix, and others; b) it supports API'S of the following languages: PHP, Perl, C, C++, Java, Pynthon, and others; c) it has support to multiprocessors; d) sophisticated system of flexible and safe cryptographic passwords; e) it has support to DDBC- the user can easily connect the Access to a database MySQL; f) it supports up to 16 index for table; g) source code is written in C and C++ and tested with a variety of different compilers.

Because it is a database totally compatible with language SQL Ansi, practical, fast, efficient and of easy handling and extremely reliable, maybe this is the reason that it is very used and it is being transformed quickly into the favorite server of many developers.

Flash graphical tool was also used because it is a tool of authorship and edition of vectorial pictures with animation, sound and interactive. Based on vectorial pictures, turning possible the creation of advanced effect in very small archives.

Beyond vectorial pictures, to Flash content can be added bitmap archives, digitalized sounds in .au and .wav, video formats, and even though archives with sequences of pictures – the animated GIFs.

The main advantage of Flash is the easy learning, and an enormous amount of tools enclosed in software, facilitating the programming.

2   The System

The basic architecture of the system consists of several experts connected in the system by Internet, set up in the base of experts, feeding the knowledge bases in order that the users in Internet can consult them. There will be two modules, the expert module and the user module. As show in figure 1.
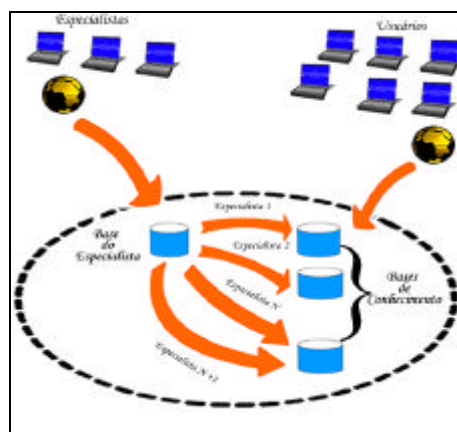
Picture1 – system structure

The current system is composed only by expert module. In this module the expert will be possible to set up in the base informing data referring his/her person, the expert can also inform his/her specialty and a brief professional curriculum. After the register is made the expert will be able to access the system using his/her login and his/her password. As shown in picture 2.

Expert's module possesses the following items: new Base: Where the Expert will be able to initiate his/her Knowledge base; my register: It will be able to make alterations in the register; accesses: Its actions of the system will be able all inside; chat: It will be able to debate subjects with other experts set up in register of the system; forum: It will be able to add questions or to answer questions about definitive subject; alter Password: It will be able to modify his/her password. The picture 3 shows Expert's menu interface.

Picture2 – Part of the screen to set up

Picture 3 – Expert's menu

3.1 New Base

To initiate a new base of knowledge the expert fills a small register informing his/her name, a small comment, his/her knowledge area, his/her sub-area, his/her domain and the purpose to acquire this knowledge. As shown in picture 4.

In this stage the expert informs the main five variable of the context in study and the importance scale of each one, also informing the amount of sub-variable of each variable and its type. As shown in figure 5.
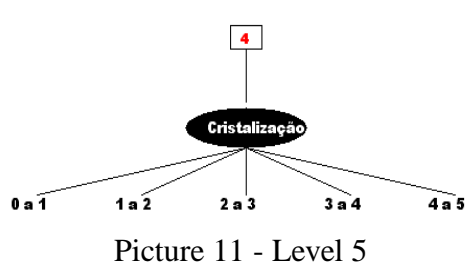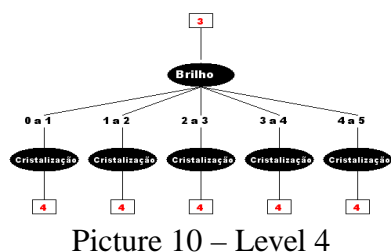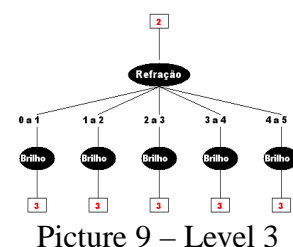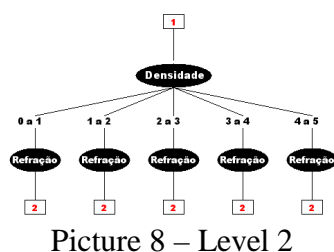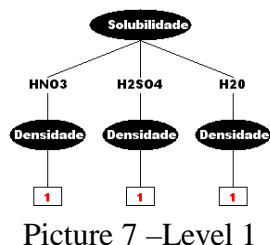

Picture 4 – First screen of new base

Picture 5 – Screen 2 of New Base

The following step is to inform the name of the sub-variables for each variable. The system will go to set up the knowledge tree automatically, representing the connection among the variables. A show in picture 6.
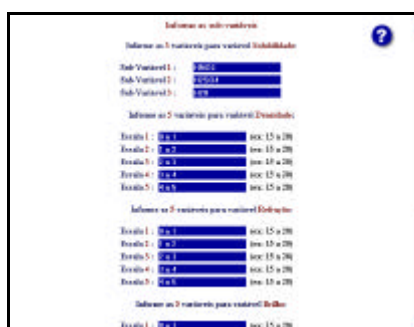
Picture 6 – Screen 3 of New Base

The system will create a directory containing 10 archives, 5 in the format (png) and five in format HTML Each figure will contain a level, and one link for each level except the figure of level 5. To follow the five figures are shown.

Picture 7 –Level 1

Picture 8 – Level 2

Picture 9 – Level 3

Picture 10 – Level 4

Picture 11 - Level 5

The knowledge tree will be shown in the picture format, made with the graphical library of PHP (gd-20.1). It will be set up independent of the number of variables. The root of the tree will contain the variable of bigger scale (5), and below it its sub-variables, after that one variable of scale will come to each one of sub-variables (4), after that the sub-variables will come, thus successively until arriving in the sub-variables of variable of scale (1).

In studies and tests made during the implementation of the tree it was verified that the same one was with its visualization harmed while the number of involved variables magnified. Being thus, for a better visualization of the tree, it was decided to break up it in n parts, where n is the number of involved variables in the process. To go from a level to another the user clicks in the connectors (squared).

3.2 Knowledge Modeling

After the construction of the tree the expert will be able to verify if his/her reasoning is correct through the rules of production generated with base in the tree of decision constructed.

The system has the tools of chat and forum in order to the expert of the same area can change ideas about the best refinement of knowledge in a determined area.

4 Conclusions

The system is still in development phase, being that only the expert module is constructed. The following stages say respect to the elaboration of content confrontation developed by experts in the same subject through Kelly's Personal Construction Psychology. Soon, after the ending of the construction of the production rules the using module will be constructed where students and professionals of the area will be able to visualize the rules and to use the system. And also the use of students of the discipline - Artificial Intelligence, to help in the understanding of the reasoning used for the specialist in determined subject.

5 Bibliographical References

BITTENCOURT, G. **Inteligência Artificial – Ferramentas e Teorias**, Editora da Universidade Federal de Santa Catarina, 1998.

DIEDERICH,J., RUHMANN,I. A knowledge acquisition tool for experts systems.**International Journal of Man-Machine Studies 26** (1) 29-40. 1987

ESHELMAN, L et all. MOLE: A tenacious knowledge acquisition tool. **International Journal of man-Machine Studies 26** (1) 41-54, 1987

GAINES, B.R. Induction of inference rules for expert systems. **Fuzzy Sets and Systems,** 18(3) 315-328,1986

HAYES-ROTH, F.; WATERMAN, D.A. and LENAT, D.B, Eds. **Building Expert Systems.** Reading, Massachusetts: Addison-Wesley. 1983

MARCUS, S. Taking backtracking with a grain of SALT. **International Journal fo Machine Studies 26** (4) 393-398. 1987

KLINKER, G et all. – report-driven knowledge acquisition. **International Journal of Man-Machine Studies 26** (1) 65-79. 1987