An object-oriented framework for virtual diagnosis

Silvia Maria Fonseca Silveira Massruhá [a], Juliano Pastorelli Dutra [b], Sérgio Aparecido Braga Cruz [a], Sandra Sandri [c], Jacques Wainer [d] and Marcelo A B Morandi [e]

[a] Embrapa Agriculture Informatics – Campinas/SP -Brazil , {silvia,sergio}@cnptia.embrapa.br

[b] Brainweb Tecnologia – Campinas/SP -Brazil , juliano@brainweb.com.br

[c] Institut d'Investigación en Intelligéncia Artificial IIIA/CSIC - Belaterra Spain, sandri@iiia.csic.es

[d] Institute of Computing/UNICAMP - Campinas/SP – Brazil - wainer@ic.unicamp.br

[e] Embrapa Environment – Jaguariúna/SP –Brazil, mmorandi@cnpma.embrapa.br

Abstract

This paper presents an object-oriented approach to support, on the web, an integrated framework for diagnosis, investigation and treatment of disorders (diseases, computer faults, etc). In this framework, called Fuzzy Covering Theory (FCT), the knowledge is basically modeled through causal associations between objects and inference is abductive. In this new approach, it was necessary to specify the classes model using object - oriented methodology and the knowledge base model that was a flat file in Prolog. These models are discussed in this paper as well as the advantages and limitations of this new approach.

Key words: diagnosis, abductive inference, fuzzy logic, object-oriented, web application.

1 Introduction

Clinical problem solving has been the subject of intense interest by researchers who intend to model human reasoning. Research in cognitive science, decision theory and computer science has supplied a large view of the cognitive process that is the foundation of the clinical decision-making process. The main segments of a clinical problem are diagnosis, investigation and treatment. The success of clinical problem-solving depends on two main factors: diagnosis correctness and treatment effectiveness. Furthermore, when the investigation task is well done, better diagnoses and treatments are obtained, with a reduction of risks and costs (Durkin, 1994; Henrion et al., 1992; Mahaman et al., 2003; Taboada et al., 1999) .

This work presents an object-oriented approach to support, on the web, an integrated framework for diagnosis, investigation and treatment of disorders (diseases, computer faults, etc) proposed in Massruhá (2003). In this framework, called Fuzzy Covering Theory (FCT), the knowledge is basically modeled through causal associations between objects and inference is abductive (Massruhá et al., 2004; Massruhá et al., 2005). FCT puts together concepts of Parsimonious Covering Theory – PCT (Peng & Reggia, 1990), fuzzy sets logic and decision theory, in order to address the various inherent aspects involved in clinical reasoning, such as the possibility of several disorders to conjunctively cause a series of manifestations, the manipulation of temporal information, the consideration of favorable conditions for the development of a disorder, the dificulty of an expert to yield generalized knowledge devoid of uncertainty/imprecision, the manipulation of crucial factors in decision making in investigation and treatment tasks like cost and risk.

Although the FCT has been a powerful tool to diagnosis, investigation and treatment of diseases, it has some limitations in its implementation. For instance, the programming language used in the implementation of the first version of FCT is Prolog due its adequation for the problem. One limitation is the complexity of generation of the knowledge base. Durkin (1994) brings that the knowledge representation and acquisition are the most critical phases in the development of an expert system. Thus, it is very important to have a tool to generate automatically or semi-automatically the knowledge base from the expert knowledge. Other limitation in the FCT is the integration of the three main modules in Prolog: diagnosis, investigation and treatment.

The object-oriented framework proposed to solve these limitations is presented in this work. The validation of the theoretical model has been performed in phytopathology, specifically, at the diagnosis, investigation and treatment of corn diseases. In this paper, we discuss the interface in JAVA of this application too.

This paper is organized as follows: Section 2 describes the methodology used in this work. Section 3 presents a case study with corn diseases. Finally, Section 4 brings the results obtained so far as well as the conclusions in our research project.

2 Methodology

The model of FCT was revised and a new architecture in levels was implemented using the J2EE specifications (Java 2 Platform, Enterprise Edition)(Sun Mycrosystems, 2005)(Fig. 1). The J2EE model is divided in three main layers: Presentation Level: manipulates the layout, browsing and treatment of actions; Business Level: manipulates the rules and fuzzy logic of the application and Persistency Level: manipulates the operations of database. This new model becomes easier the software evolution phase.

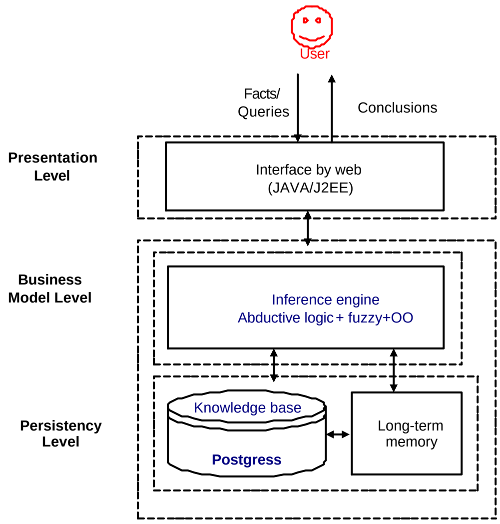

Fig. 1 An object-oriented framework for diagnosis.

To implement this model, we used some free open-source frameworks , such as:

- Struts: is an extensible framework for creating enterprise-ready Java web applications (http://struts.apache.org/);

- Spring: is a layered Java/J2EE application framework. A central focus of Spring is to allow for reusable business and data access objects that are not tied to specific J2EE services. Such objects can be reused across J2EE environments (web or EJB), standalone applications, test environments, etc without any hassle. (http://www.springframework.org/);

- Hibernate: is a powerful, high performance object/relational persistence and query service. Hibernate lets you develop persistent classes following object-oriented idiom - including association, inheritance, polymorphism, composition, and collections. (http://www.hibernate.org/).

Subsequently, we will detail each layer to explain how this new approach supports the FCT model proposed in Massruhá et al.(2005). The persistency layer involves since graphic interface until the database operations such as: insert, update, delete and consult. In this layer, the open-source frameworks (Spring, Struts, Hibernate) were very important because they brought flexibility for the future updates. The auxiliaries predicates such as fuzzy operations become methods in JAVA in this layer. The business layer involves the methods that complain the three phases of clinical process (diagnosis, investigation and treatment). The presentation layer involves the manipulations of the application forms such as the operations that manipulate the layout, browsing and treatment of actions of the forms. New technologies such as AJAX were used in this layer. The technology AJAX integrates Javascripts e XML (eXtensible Markup Language). More details of this structure can be found in http://ajaxtags.sourceforge.net/.

In this new approach, it was necessary to specify the classes model using object - oriented methodology (Rumbaugh, 1999) as well as the database model, because the original knowledge base was one flat file in Prolog.

The class model, generated from this review of the FCT, involves six main classes (Disorder, Manifestation, Exam, Treatment, Problem and Case) and 12 auxiliaries classes (Graph, Node, Edge, Fuzzy, ControlConditions, Categories, Phase, Severity, Question, City, Event and User). These classes incorporate every inference engine developed in Prolog to implement the FCT model.

The Entity-Relationship Model (ERM) was used to represent the database in this work (Fig. 2). The database management system tool used in this implementation is the PostgreSQl. Note in Fig. 2 that the quantity of entities generated in the MER is large. This new knowledge base has 22 tables while in the first version in Prolog we had one flat file only. Thus, we had to divide the structure of this framework in two modules to improve the software performance.

This structure is divided in Expert and User module (Fig. 3). The Expert module generates the knowledge base from the configuration of the expert. The User model determines a diagnosis for a particular case given by the user and a final report is generated, by verifying the consistency between the FCT model, the knowledge base and the particular case.

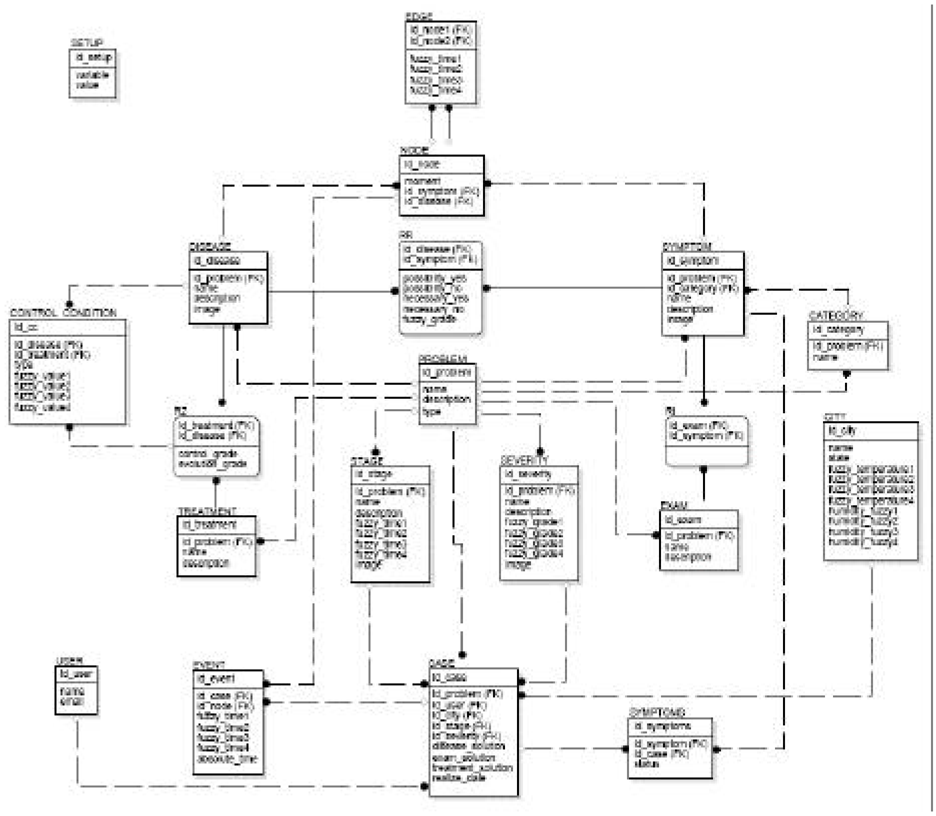

Fig. 2 The database model

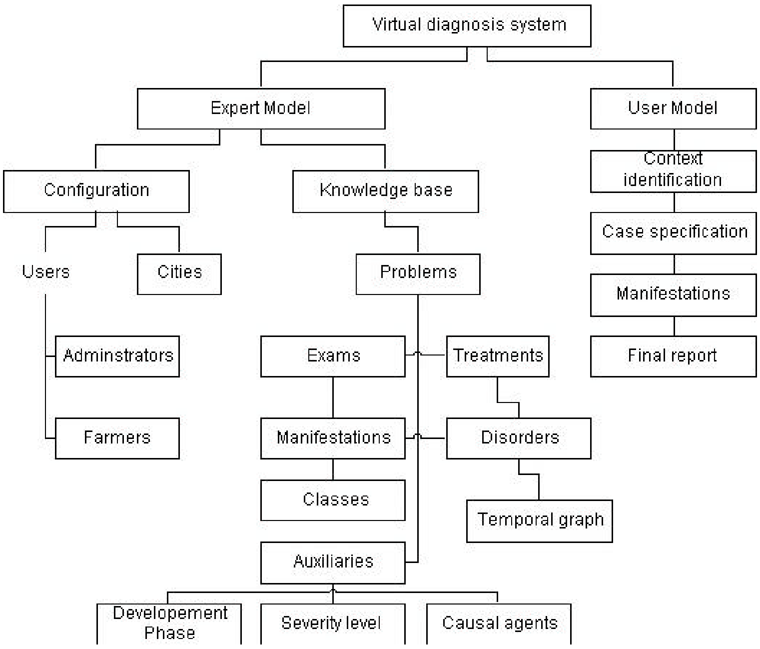

Fig. 3 Two main modules of the object-oriented framework.

The validation of the theoretical model has been performed in phytopathology. In the next section, we will show an example of expert system for diagnosis and treatment of corn diseases that was developed under this structure presented in Fig. 3.

3 An example in corn diseases

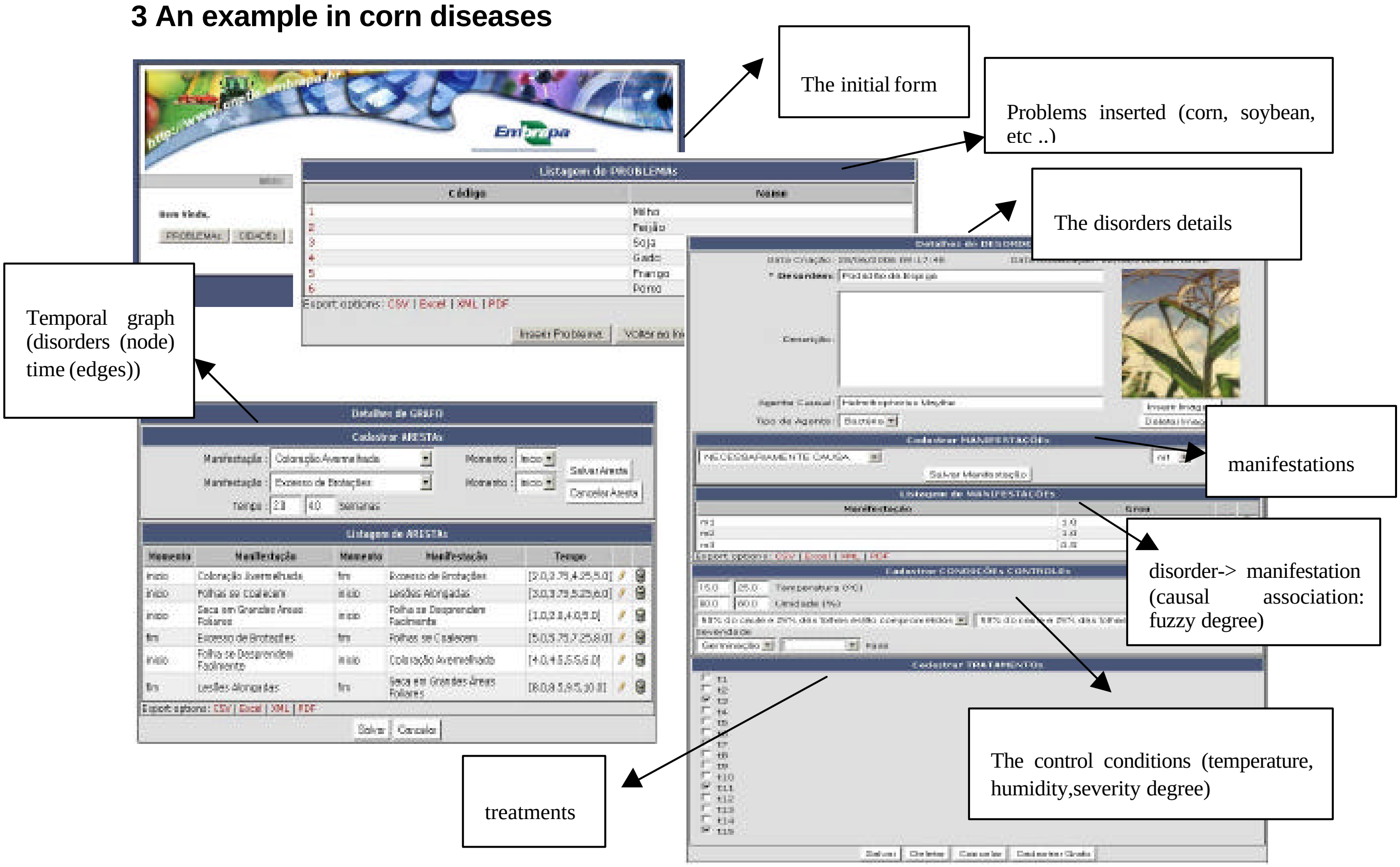

Fig. 4 The Expert module

In our case study, we worked with a knowledge base of 41 disorders that can occur in a corn plantation (Malcoln, 1980) (available in http://diagnose.cnptia.embrapa.br/milho), 12 laboratory exams and 15 fungicides . To simplify our example, this section is split into two subsections: Expert and User module.

3.1 The Expert Module

In this module, the expert access the framework to configure the system. The Expert module generates the knowledge base from the configuration of the expert. In the initial form, the specialist defines the types of disorders he wants to manipulates: vegetal or animal. For example, corn plantation, soybean plantation or cattle diseases. Subsequently, the expert entries the data related to the disorders, symptoms, categories (i.e, leaf, root in plants diseases), development stages, severity degrees, exams or laboratorial tests, treatments and causal agents. An example on how to manipulate the data of the disorders is shown in Fig. 4.

3.2 The User Module

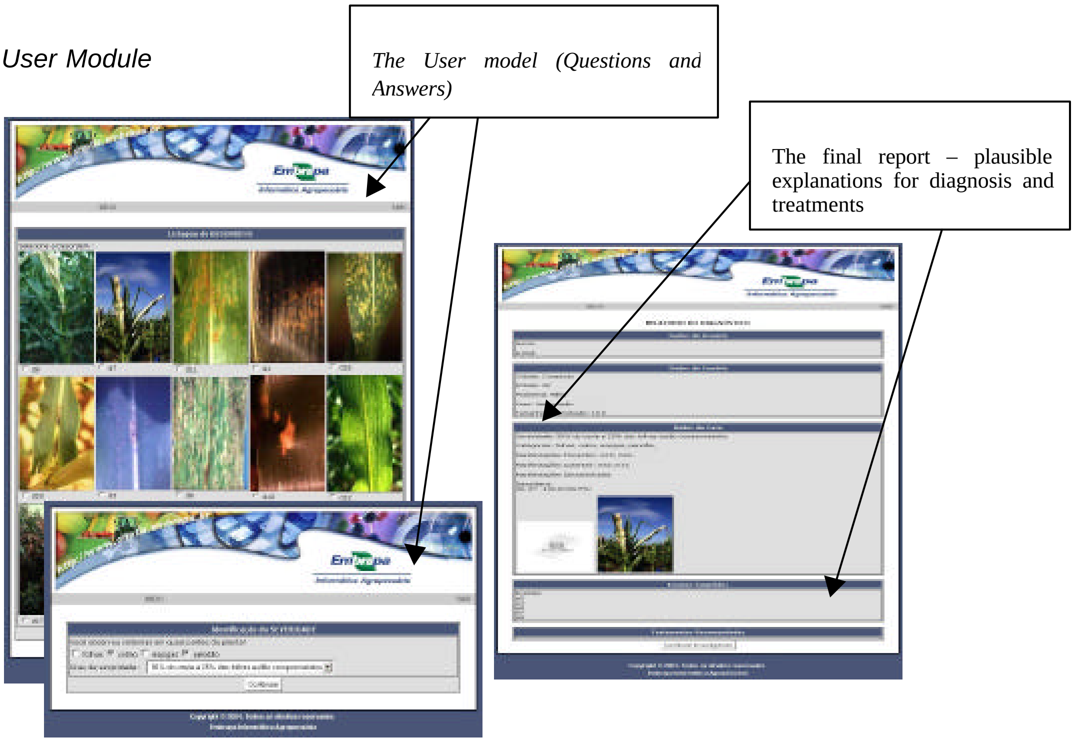

Fig. 5 The User module

The User model determines a diagnosis for a particular case given by the user and a final report is generated, by verifying the consistency between the FCT model, the knowledge base and the particular case. The session starts with a sequence of questions such as “Which location did the symptoms appear?”, “What kind of lesions did it appear?”, “When did the symptoms appear?”, and so on. Then, the system dynamically adjusts the flow-of-control between the rules of the FCT model based on the user's answers to questions and at the end, a diagnosis is provided (Fig. 5).

4 Results and discussion

In this paper we presented a new approach to support, on the web, an integrated framework for diagnosis, investigation and treatment of disorders (diseases, computer faults, etc). as well as the new structure of this framework and the interface in JAVA of this application.

The final results in our research project was successful in relation to the framework as well as the final application. In former, the FCT model proposed in Massruhá,(2003) was revised and a new architecture in layers was implemented using the J2EE specifications (Java 2 Platform, Enterprise Edition). Some of the advantages of this new approach are:

- the J2EE model is more flexible and powerful that Prolog approach because manipulates the three layers (Presentation Level, Business Level and the Persistency Level) independently;
- the automatic generation of the knowledge base by Expert module is more efficient such as described in Massruhá (2003);
- the integration of the three modules (diagnosis, investigation and treatment) of the FCT model became the system with a best performance;
- the interface is more friendly under this framework .
- and finally, this new framework becomes easier the software evolution phase.

In conclusion, the two modules (Expert and User) showed efficiency and powerful in the context of application for decision-making in the diagnosis of plant diseases. However, due to its generic characteristics, it is expected that the theoretical developments obtained in this work can successfully employed to solve diagnosis problems in any animal or vegetal area.

5 References

DURKIN J. 1994. **Expert Systems: Design and Development**. Englewood Cliffs, New Jersey: Prentice Hall, Inc. 800 p.

HENRION, M.; BREEZE, J. S.; HORVITZ, E. J. Decision analysis and expert systems. **Magazine** v. 13, n. 8, p. 64–91, 1992.

MAHAMAN, B. D., PASSAM H. C., SIDERIDIS A. B. and YIALOURIS C. P., 2003. **DIARES-IPM: a diagnostic advisory rule-based expert system for integrated pest management in Solanaceous crop systems**. Agricultural Systems, v. 76 , pp. 1119-1135.

MALCOLM, C. **Compendium of corn diseases**. 2. ed. St Paul: American Phytopathological Society, 1980. 105 p. (APS. Disease Compendium Series).

MASSRUHÁ, S. M. F. S. **Uma teoria de coberturas nebulosas para diagnóstico, investigação e tratamento.** 2003. 251 p. Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos (In Portuguese).

MASSRUHÁ, S. M .F .S., SANDRI S. A. ; WAINER,J., 2004. **Ordering manifestations for investigation in incomplete diagnosis.** In Proc. 10th International Conference IPMU 2004, Perugia,Italy, pp.1153-1160

MASSRUHÁ, S. M. F. S. ; SANDRI, S. ; WAINER, J. ; MORANDI, M. **An integrated framework for clinical problem solving in agriculture.** In: Efita /WCCA 2005 Joint Congress on IT in Agriculture, 2005, Vila Real. Efita /WCCA 2005 Joint Congress on IT in Agriculture, 2005a. p. 1400-1407.

PENG, Y., REGGIA, J. A. **Abductive inference models for diagnostic problem-solving**. New York: Springer Verlag, 1990. 285 p.

RUMBAUGH, J.;JACOBSON, I.; BOOCH, G. **The Unified modeling language reference manual**. New York: Addisson Wesley Longman Inc, 1999, 550p.

SUN MICROSYSTEMS, Inc. ***Java 2 Platform, Enterprise Edition (J2EE)*** Disponível em: <http://java.sun.com/j2ee/index.jsp>. Acesso em: 10 jun. 2005.

TABOADA, M., LAMA, M., BARRO, S., MARIN, R., MIRA, J., PALACIOS, F. A problem-solving method for unprotocolised therapy administration task in medicine. **Artificial Intelligence in Medicine** v. 17, p. 157–180, 1999.