An Overview of the Current State of the MedFrame/CADIAG-IV Project

Dieter Kopecky, Klaus-Peter Adlassnig, and Andrea Rappelsberger

Department of Medical Computer Sciences, Section on Medical Expert and Knowledge-Based Systems
University of Vienna Medical School, Spitalgasse 23, A-1090 Vienna, Austria
e-mail: {dieter.kopecky, kpa, andrea.rappelsberger}@akh-wien.ac.at

Abstract. *CADIAG-IV (Computer-Assisted Diagnosis, system version 4) is a medical consultation system for internal medicine that uses concepts of linguistic variables, fuzzy sets, and fuzzy numbers to deal with uncertainties in medical terminology and medical correlations. Relationships between medical entities such as symptoms, diagnoses, and therapies are formalized as rules and defined by two attributes: (a) strength of confirmation and (b) frequency of occurrence. Reasoning in CADIAG-IV is performed by iteratively applying the compositional rule of fuzzy inference until no further improvement of the generated diagnostic hypotheses is possible. CADIAG-IV is integrated within the medical expert system shell MedFrame, which provides a set of tools and concepts for building homogenous expert systems in an easy manner. MedFrame also provides a set of interfaces allowing the expansion and adaptation of the existing system by an application programmer. In this paper, the principal goals, main components, and concepts of MedFrame and CADIAG-IV are discussed.*

1. Introduction

CADIAG-II [1,2] and CADIAG-III [3] are data-driven fuzzy expert systems built for computer-supported consultation in internal medicine. They provide diagnostic hypotheses and propose further useful examinations in response to the input of a list of symptoms, signs, and test results obtained for a patient. CADIAG-IV is intended to be the successor of these two systems and will intensify the application of fuzzy concepts. In contrast to CADIAG-II and -III, which have been developed for an IBM-host-based system, CADIAG-IV will be part of the PC-based medical expert system shell MedFrame that is currently being developed at our department.

The following sections examine the concepts used in CADIAG-IV with special emphasis on the innovations and modifications in comparison with CADIAG-II. Since the basic, and therefore reusable, methods are an integral part of MedFrame, we will first introduce the key concepts of MedFrame.

In the sequel any reference to CADIAG-II expresses a reference to both CADIAG-II and -III, since the methods used in both systems are basically the same.

2. Methods

2.1. MedFrame

Expert systems are obviously required to contain at least three components: a knowledge base using a particular knowledge representation formalism, an inference engine, and a dialog component for communication between the system and the user. A knowledge acquisition component and a component for explaining the inferred results are desirable. This is especially true for the field of medical decision support.

Expert system shells provide the user with a particular knowledge representation formalism and an adequate inference mechanism for building knowledge bases. Therefore, it is no longer necessary to reimplement or adapt many parts of an existing expert system when a new knowledge base is required. Such systems usually have the following drawbacks:

- only a single knowledge representation formalism (e.g., predicate logic) and a single inference mechanism are provided; therefore only a certain class of problems may be modelled;
- knowledge acquisition components are very seldom targeted to the needs of domain experts, and therefore specially trained knowledge engineers are required;
- the implementation of tightly integrated—and therefore reusable—dialog and explanation components is usually not supported;
- the possibility to provide reference cases for an automatic validation of the knowledge base after a modification is commonly not available; and
- the storage of external information such as patient data is usually not possible.

In addition, expert system shells are poorly integrated within existing, widely used technologies. For instance, as the significance of the Internet is increasing, the missing integration of client/server- and web-based consultation in medical expert system shells is no longer acceptable.

These concerns prompted us to define a set of requirements that a modern expert system shell for use in the medical domain should fulfill. Accordingly, we decided to implement an expert system shell known as MedFrame that will include the following facilities:

- various knowledge representation formalisms to store medical knowledge and adequate inference mechanisms;

- interfaces to add further inference mechanisms;
- concepts for modeling and handling uncertainty in medical terminology and medical relations;
- mechanisms for storing patient data and history;
- a graphical user interface providing the four essential components of expert systems: (a) a knowledge acquisition component, (b) a component allowing the definition of a set of test patients with approved gold-standard diagnoses, (c) an interface for the input of patient administrative and examination data, and (d) a component for displaying the inferred diagnostic hypotheses and proposals for further examination;
- interfaces to adapt the components of the graphical user interface to the requirements of particular medical domains;
- components for implementing client/server- and web-based access to the expert systems in an easy way; and
- various components for the maintenance of the expert system shell.

By offering these facilities, MedFrame provides the end user with a set of powerful tools for developing knowledge bases and also allows the application programmer to extend the possibilities of the expert system shell components, provided by the system programmer, by implementing a set of interfaces and using a collection of libraries. Therefore, MedFrame significantly reduces the time and cost of building new expert systems requiring specialized inference mechanisms, knowledge acquisition and reasoning components, or access from the Internet with an HTML browser.

2.2. CADIAG-IV

CADIAG-IV will be the first expert system completely based on MedFrame. Its predecessor CADIAG-II is a data-driven fuzzy medical expert system providing diagnostic hypotheses and proposing further useful examinations in response to the input of a list of symptoms, signs, and test results obtained for a patient. To deal with uncertainties in medical terminology and medical relationships, CADIAG-II relies on the theory of fuzzy sets, particularly on the concepts of linguistic variables and fuzzy logic [4,5]. A comprehensive discussion of the concepts of CADIAG-II is given in [6–8], details concerning the implementation of the system may be found in [9].

As in CADIAG-II, a clear distinction between patient data on a detailed observational level, i.e., detailed history items, signs from physical examinations, quantitative laboratory test results, etc., and abstract symptoms commonly used in diagnostic discourse is established in CADIAG-IV as well. In the beginning of a consultation, a transformation step known as data-to-symbol conversion, which abstracts and aggregates medical information provided by the physician into this internal representation, is applied.

In CADIAG-II, this transformation process assigns a real number in (0,1], a "degree of compatibility", to every symptom, where a value of 1 means that the corresponding symptom is proven, while values in (0,1) mean that the symptom applies to the patient with some strength of confirmation. Symptoms that can be definitely excluded are assigned the value 0. The transformation process is formally defined by a set of linguistic variables and their corresponding membership functions [1,2].

Unfortunately, this methodology can only express total exclusion of a medical entity, but is unable to provide so-called negative evidence, thus indicating the absence of a particular medical entity only to a certain degree. To overcome this limitation CADIAG-IV assigns two values to every medical entity, (a) strength of evidence and (b) strength of counterevidence. Both values are—in order to overcome the criticism of rather sharp and therefore not very fuzzy point values in CADIAG-II—fuzzy numbers in [0,1]. The interpretation of these values is as follows: a fuzzy number representing 0 means that we have no evidence (or counterevidence) regarding this medical entity, while a fuzzy number representing 1 is interpreted as evidence (or exclusion). Intermediate values denote evidence that is not sufficient to prove or exclude the entity in question. Therefore, the data-to-symbol conversion is adapted so that every symptom is assigned two fuzzy numbers instead of a single point value. In addition, the process has been enhanced to deal with context-sensitive as well as pathophysiologically interdependent data. A more detailed examination of the data-to-symbol conversion in CADIAG-IV may be found in [10].

As in CADIAG-II, relationships between medical entities are represented as rules being defined by (a) the strength of confirmation and (b) the frequency of occurrence in CADIAG-IV as well. The occurrence value describes the certainty with which the left-hand side of the rule will occur in patients already showing the right-hand side, and the confirmation value describes the certainty with which the consequence of the rule occurs in patients already showing the antecedent, that is, how much evidence the antecedent provides for the consequence. While in CADIAG-II a relation is described by a point value, in CADIAG-IV rules are classified into two categories: those giving evidence for and those giving evidence against the right-hand side; the strength of the relation is expressed by a fuzzy number in [0,1].

CADIAG-II supports the following kinds of relations: symptom/symptom, diagnosis/diagnosis, symptom/diagnosis, and symptom-combination/diagnosis (a symptom-combination is a complex fuzzy logical combination of symptoms and/or diagnoses defining pathophysiological states). CADIAG-IV, in contrast, allows a wider range of combinations, also including therapies and medical treatments, both as antecedents and consequences of rules.

The basic concept on which the inference mechanisms of both systems rely is the compositional rule of fuzzy inference [11]. A comprehensive description of the inference in CADIAG-II is given in [1,2,6]. The main improvement in CADIAG-IV is its handling of the newly introduced concept of counterevidence and the computation of evidence and counterevidence with fuzzy numbers. Further improvements are the realization of patient-specific adaptation of the rule base during inference as well as the possibility to use different fuzzy operators for the evaluation of symptom combinations and the operators used for the compositional inference rule. The theoretical considerations regarding inference in CADIAG-IV may be found in [12,13].

3. Results

Based on the results of the formal considerations mentioned in the previous section, we started to implement both MedFrame and CADIAG-IV. In order to accomplish smooth integration with client/server- and web-based access concepts, MedFrame and all the components based on it were realized with Java.

3.1. MedFrame

MedFrame provides a set of tools and concepts for building homogenous expert systems meeting the requirements outlined in the previous section. A comprehensive overview of MedFrame would far exceed the scope of this paper. Therefore, only the most important facts have been examined.

The core component of MedFrame is an object model for storing domain knowledge in various representation formalisms [14], including lookup tables, rules (if-then rules, certainty factor rules, fuzzy control rules, ...), crisp and fuzzy automatons, crisp and fuzzy decision graphs, and fuzzy-neuro systems. In addition, the object model has been extended by a set of classes for storing patient administrative and examination data [13]. The object-oriented database management system Poet 6.1 is used as the underlying permanent storage medium.

In order to support the modeling of uncertainties and to consider uncertain facts in inference, a set of classes based on the theory of fuzzy sets [4,5] have been developed, including the following:

- a large set of type-1 and type-2 membership functions (e.g., Pi shapes, Gamma shapes, …) [15];
- the implementation of the concepts of type-1 and type-2 fuzzy sets; and
- the implementation of fuzzy numbers and tools for calculating with them.

In order to ensure that expert systems implemented on the top of MedFrame support client/server- and web-based access, a set of mechanisms for developing client/server-based software on the top of Remote Method Invocation (RMI) and the Common Object Request Broker Architecture

(CORBA) have been created. Support for browser-based access has been provided by designing a framework for implementing HTML-based interfaces to MedFrame. A set of Java classes designed for being integrated into Java Server Pages have been developed for that purpose. They provide the application programmer with everything needed to realize multi-language HTML interfaces in MedFrame. The resulting communication channels are depicted in Figure 1. For a detailed examination of the communication facilities of MedFrame please refer to [14,16].

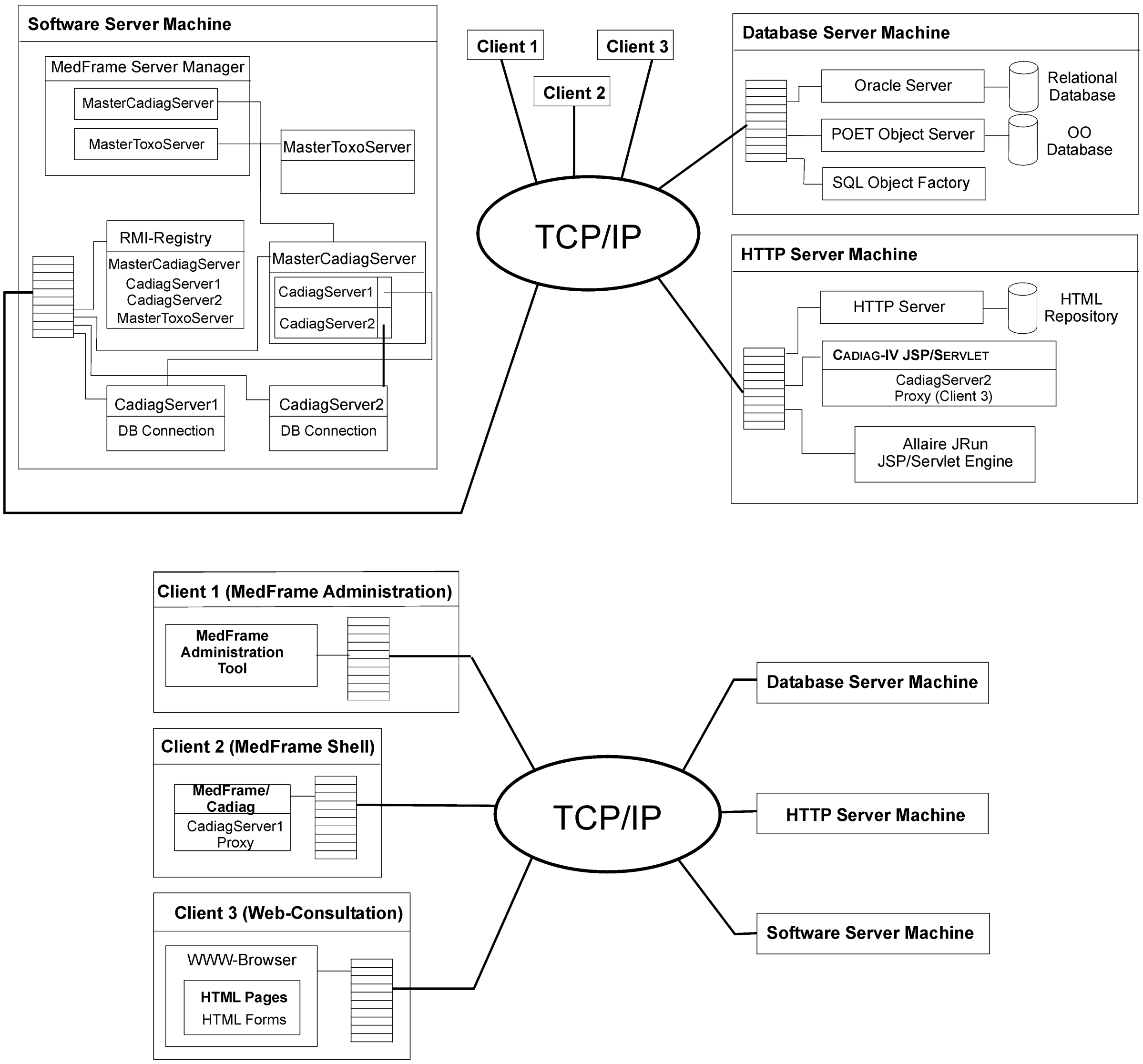

Figure 1: A model of the communication channels in MedFrame.

Finally, substantial effort has been made to develop reusable components for the implementation of practical user interfaces. For this purpose most of the Java Swing GUI components have been extended to include the ability to deal with multiple languages and context-sensitive help that may be reused for the creation of attractive user interface components.

In a further subproject a graphical user interface component providing interfaces for the four essential parts of an expert system mentioned in the previous section has been implemented. The resulting framework enables the application programmer to adapt the user interfaces to the needs of specific fields of internal medicine and to the users' requirements by simply subclassing the existing classes, and to profit from the integrated availability of multiple language and help system support. In addition, a reference implementation for the clinical field of hepatology has been developed.

Although no knowledge acquisition components have currently been integrated into MedFrame and the graphical user interface, two tools developed at our department may be adapted for integration within the expert system shell: the Knowledge Base Builder [17] and the MedFrame browser [18].

At the present time we are working on the definition of the interfaces for the inference mechanisms and the implementation of the default inference that will be provided by MedFrame. This, in fact, will be the inference mechanism of CADIAG-IV.

3.2. CADIAG-IV

The development of CADIAG-IV is currently in its early stage, as the required preparatory work on MedFrame has just been completed. The first and most important task is the transformation of the existing knowledge base and of patient data into the MedFrame database. For this purpose, a converter component is currently being developed.

Furthermore, after adaptation of the syntax of CADIAG-II rules to the new requirements, a parser component was designed with JavaCC. This component is able to convert the textual representation of CADIAG-IV rules into the internal representation of MedFrame, and is not only useful for data conversion but also may be used later when having to parse the textual specification of rules during the knowledge acquisition process.

As previously mentioned, we are currently working on the formal specification and implementation of the inference mechanism of CADIAG-IV.

4. Conclusion

A number of important steps for the proposed expert system shell have been initiated. The reference implementation of the dialog component in clinical hepatology has proven the applicability of the user interface components that have been developed. In the same way, the object model and the communication components for client/server and web-based consultation have passed a first test in the development of ToxoNet [16], which was implemented along with the evolving MedFrame components.

The current state of MedFrame is a suitable foundation for implementing the copious preparatory theoretical work that has been done in the CADIAG-IV project. The data conversion components and the implementation of the inference process are steadily evolving and the first prototype of CADIAG-IV is scheduled to be completed in the beginning of next year.

Acknowledgements

This project was financially supported by Siemens AG, MED-GT, Erlangen, Germany.

References

[1] Adlassnig, K.-P., Kolarz, G., Scheithauer, W., and Grabner, H. (1986) Approach to a Hospital-Based Application of a Medical Expert System. *Medical Informatics* 11, 205–223.

[2] Adlassnig, K.-P., Scheithauer, W., and Kolarz, G. (1986) Fuzzy Medical Diagnosis in a Hospital. In Prade, H. and Negoita, C.V. (eds.) *Fuzzy Logic in Knowledge Engineering*. Verlag TÜV Rheinland, Köln, 275–294.

[3] Schuh, C. (1993) CADIAG-3: *Ein medizinisches Konsultationssystem und seine Anwendung in der internen Medizin*. Master Thesis, Vienna University of Technology, Vienna.

[4] Zadeh, L.A. (1965) Fuzzy Sets. *Information and Control* 8, 338–353.

[5] Zadeh, L.A. (1975) The Concept of a Linguistic Variable and Its Application to Approximate Reasoning-I. *Information Sciences* 8, 199–249.

[6] Adlassnig, K.-P. (1983) *Computerunterstützte medizinische Diagnostik mit Fuzzy-Mengen und Fuzzy-Logik*. PhD Thesis, Vienna University of Technology, Vienna.

[7] Adlassnig, K.-P. (1986) Representation and Semiautomatic Acquisition of Medical Knowledge in CADIAG-1 and CADIAG-2. *Computers and Biomedical Research* 19, 63–79.

[8] Adlassnig, K.-P. (1980) A Fuzzy Logical Model of Computer-Assisted Medical Diagnosis. *Methods of Information in Medicine* 19, 141–148.

[9] Fischler, F. (1994) *Die Wissensbasis und der Inferenzprozeß des medizinischen Expertensystems CADIAG-II/E*. Master Thesis, Vienna University of Technology, Vienna.

[10] Bögl, K., Leitich, H., Kolousek, G., Rothenfluh, T., and Adlassnig, K.-P. (1996) Clinical Data Interpretation in MedFrame/CADIAG-4 Using Fuzzy Sets. *Biomedical Engineering—Applications, Basis & Communications* 8, 488–495.

[11] Zadeh, L.A. (1973) Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Transactions on Systems, Man, and Cybernetics* 1, 28–44.

[12] Brein, L. (1997) *Das medizinische Expertensystem CADIAG-IV: Inferenzprozeß, Erklärungen und Untersuchungsvorschläge*. Master Thesis, Vienna University of Technology, Vienna.

[13] Brein, L., Adlassnig, K.-P., and Kolousek, G. (1998) Rule Base and Inference Process of the Medical Expert System CADIAG-IV. In Trappl, R. (ed.) *Cybernetics and Systems'98*, Austrian Society for Cybernetic Studies, Schottengasse 3, A-1010 Vienna, 155–159.

[14] Kolousek, G. (1997) *The System Architecture of an Integrated Medical Consultation System and Its Implementation Based on Fuzzy Technology*. PhD Thesis, Vienna University of Technology, Vienna.

[15] Driankov, D., Hellendoorn, H., and Reinfrank, M. (1993) *An Introduction to Fuzzy Control*. Springer-Verlag, Berlin.

[16] Kopecky, D. (1999) *Design and Implementation of the Internet-Based Medical Expert System ToxoNet*. Master Thesis, Vienna University of Technology, Vienna.

[17] Bögl, K. (1997) *Design and Implementation of a Web-Based Knowledge Acquisition Toolkit for Medical Expert Consultation Systems*. PhD Thesis, Vienna University of Technology, Vienna.

[18] Heisz, H., Schürz, M., Kopecky, D., and Adlassnig, K.-P. (2001) Cadiag-IV/Rheumatology—An Internet-Based Consultation System for Differential Diagnosis in Rheumatology. In Dubois, D.M. (ed.) *International Journal of Computing Anticipatory Systems* 9, Chaos, Institute of Mathematics, University of Liege, Belgium, 382–392.