Medical Informatics Europe '97
C. Pappas et al. (Eds.)
IOS Press, 1997

Integration of a Knowledge-based System and a Clinical Documentation System via a Data Dictionary

H.-P. Eich, C. Ohmann, E. Keim, K. Lang

Theoretical Surgery Unit, Department of General and Trauma Surgery,
Heinrich – Heine - University, Düsseldorf, Germany,
[eich, ohmannch]@uni-duesseldorf.de

Abstract. This paper describes the design and realisation of a knowledge-based system and a clinical documentation system linked via a data dictionary. The software was developed as a shell with object oriented methods and C++ for IBM-compatible PC's and WINDOWS 3.1 / 95. The data dictionary covers terminology and document objects with relations to external classifications. It controls the terminology in the documentation program with form-based entry of clinical documents and in the knowledge-based system with scores and rules. The software was applied to the clinical field of acute abdominal pain by implementing a data dictionary with 580 terminology objects, 501 document objects, and 2136 links; a documentation module with 8 clinical documents and a knowledge–based system with 10 scores and 7 sets of rules.

1. Introduction

Computer-aided decision-support has rarely been introduced into clinical routine. The existing programs are not user-friendly and often there is no linkage between clinical documentation and knowledge base.

The data dictionary approach is used to guarantee the consistency of clinical and terminological data. This has become a standard to documentation in clinical routine, but not to knowledge-based systems (hereafter referred to as KBS). However, a data dictionary is necessary for the integration of clinical documentation and KBS [1]. KBS are often island systems which can only be linked to a clinic-wide documentation system, if both systems use the same vocabulary. One of the most positive effects of such a linkage is the on-line application of the KBS directly to clinical data.

In this paper we describe design and realisation of a KBS and a clinical documentation system linked via a data dictionary. The system is exemplarily applied to the diagnosis of acute abdominal pain.

2. System description

The aim of our project was to develop a documentation program and diagnostic support module for acute abdominal pain. We used object-oriented analysis, design and programming methods corresponding to the notation of Coad & Yourdon [2,3]. The program modules were developed with Borland C++ and Borland Database Engine under WINDOWS 3.1 and WINDOWS 95 on an IBM-compatible PC. The whole KBS [4] consists of three sub-modules: a data dictionary, a KBS and a documentation program (Fig. 1).

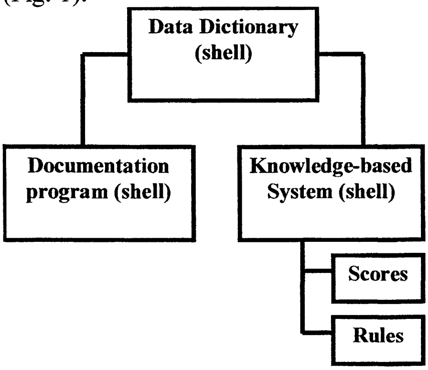

Fig. 1 : Informatical Conception

The data dictionary is a controlled vocabulary by which the integration of a knowledge base and documentation component can be achieved. In our approach it contains terminology objects and document objects, a difference which is not made in the majority of existing systems [1]. On the one hand we have definite clinical terms which meanings do not change, and on the other hand we have clinical parameters which can differ in the way they are measured. Terminology objects are described by their ID, name, definition, by internal relations to other terminology objects and by links to external classifications (e.g. SNOMED). We use five internal relations ("is a", "is part of", "is equivalent to", "is associated with", "is located in") to build up a semantic network. Document objects are terms which are usually used for clinical documentation. These objects are medical terms, possible answers, sub-documents or documents. Each document object is described by its ID, name and definition. In addition to that, a clinical parameter is described by its type, scale and unit. Each document object is linked to a terminology object. The data dictionary is developed as a shell. This conception gives the opportunity to handle terminological knowledge from different medical fields only by changing the data base. The user can easily enter new terminology and document objects.

Our documentation program is designed to collect clinical data. We use a form-based data-entry for each document (e.g. history). To reach a high acceptance rate the documentation was designed in collaboration with physicians of our surgery unit. With this documentation program we have carried out several prospective evaluation studies, in which our aim was to build up a quality-controlled prospective database on which knowledge-based methods can be applied.

The third component of our system is the KBS. It is used for diagnostic support. We have integrated rule-based systems which consists of automatically generated rules from prospective databases and diagnostic scores. Different knowledge acquisition techniques were evaluated [5]. Sets of rules in the clinical field of acute abdominal pain were generated with inductive rule-generating algorithms such as CN2 [6] and C4.5 [7]. Set of rules have a simple structure which gives the physician an inside-view why a specific diagnosis is preferred by the system. Also, different scoring-systems [8] were analysed concerning their structure and way of application. Each set of rules or score is integrated as one knowledge module. This module can be applied to patient data or the user can see its definition. The application of knowledge modules is controlled by the KBS. The user can set a filter with several criteria to find an appropriate module to his problem. These criteria are clinical problem (e.g. diagnosis, prognosis, treatment), specification (e.g. appendicitis), population (e.g. acute abdominal pain, suspected appendicitis) and method (e.g. scores, set of rules). The KBS has its own graphical human interface and can be used as a stand-alone system.

To establish such a system two forms of integration have to be taken into consideration: the terminological integration and the functional integration. The terminological integration can be achieved by the data dictionary. It guarantees that both documentation and KBS use the same vocabulary. Both components handle only IDs of clinical parameters, possible values or documents. For a functional integration the documentation component has to be extended with a trigger function which executes the KBS. Also data-entry or report modules should be adapted to achieve a user-friendly graphical human interface.

The interaction of the components and flow of information is shown in figure 2. The documentation component gets information about IDs of parameters, values and documents from the data dictionary (i). If the physician wants decision support, he calls the KBS (ii) from documentation. By this call information about the current patient (e.g. ID, sex, existing documents and duration of hospital stay) is given to the KBS. The KBS is responsible for selection and display of different knowledge modules. Information how to display terminology comes form the data dictionary, too (i). If the user has selected a specific knowledge module that should be executed, the KBS has information about which parameters are needed and requests their values from the database interface (iii). The KBS has a read-only access to the database.

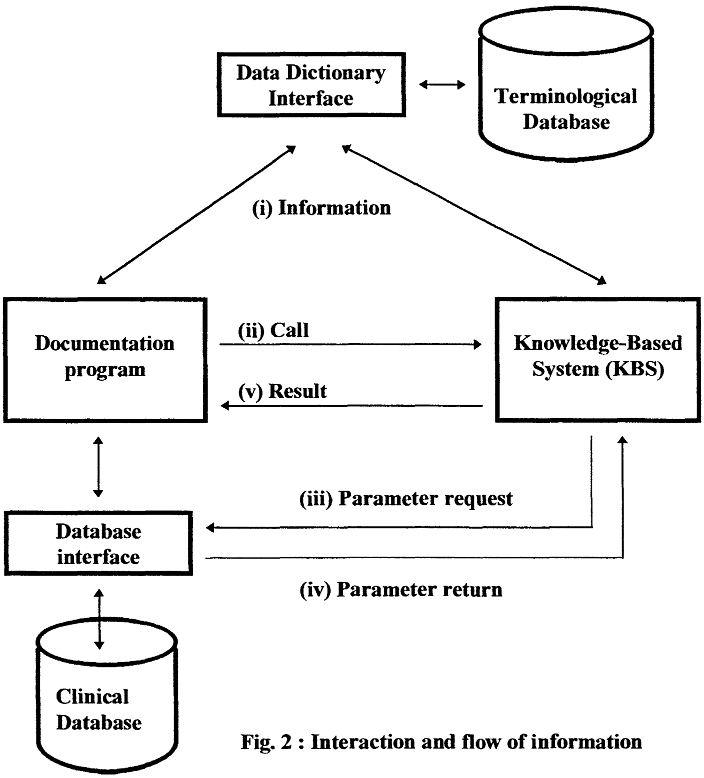

Fig. 2 : Interaction and flow of information

The database interface returns all found parameter values to the KBS (iv). The selected score or set of rules is then executed. The decision support result is sent back from the KBS to documentation (v) where it is displayed to the user. If a result should be stored to patient data, this can be done in the documentation module.

3. Results

According to the methods presented above a data dictionary for the clinical field of acute abdominal pain was created. It consists of 580 terminology objects and 501 document objects. Between these objects 2136 different links were established. Also the interface to KBS and documentation program was developed.

So far, ten diagnostic scores were integrated in the KBS. For the integration of scoring-systems a score description module was developed which allows the definition of a score with an arbitrary structure and the storage to our relational database. The parameters of the score have to be selected from the data dictionary to guarantee a correct application to clinical data. Furthermore, seven sets of rules were generated with C4.5 and integrated in the KBS. The rules were developed from an European database for acute abdominal pain. One of these sets of rules has 13 different diagnostic outcomes. The other sets have two diagnostic outcomes to strengthen or to exclude a specific diagnosis found by the first set.

Eight different documents were integrated in the documentation program so far. These are documents for history, clinical investigation, laboratory investigation, ultrasound, x-ray examination, operation, diagnosis and discharge. To store this data a database interface was programmed.

Figure 3 shows a screen shot of our KBS for acute abdominal pain. In the foreground the KBS with a set of rules is presented. The user has specified his problem, specification, population and method. The KBS found one set of rules of which two rules are displayed in the middle of the screen. In the background the documentation program with a form-based entry for the history of acute abdominal pain is presented.

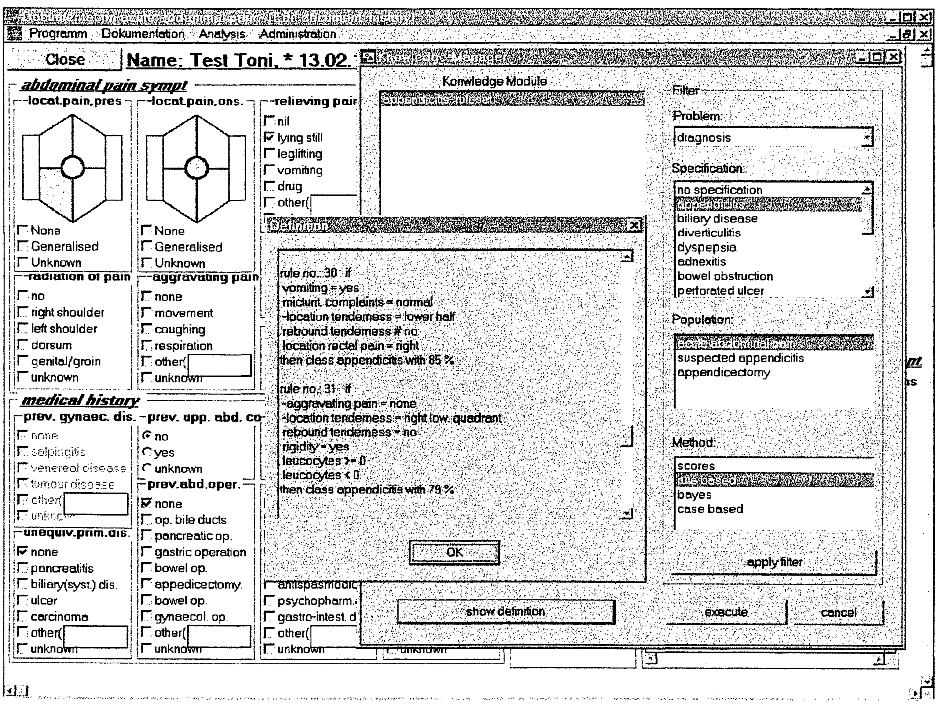

Fig. 3: Documentation program with knowledge manager

4. Discussion

The results above show that an integration of a clinical documentation system and a KBS can be established, if terminological and functional integration is achieved via a data dictionary.

Under consideration is now the linkage of KBS to a commercial clinical information system. In this case the terminological data has to be mapped with data from the information system. The whole process of mapping terminology should be done by a meta-data dictionary. The data dictionary used by the KBS is left untouched not to violate the integrity of the knowledge base. If no data dictionary is present on side of the clinical information system the correct terms with its IDs have to be identified directly out of documentation. For functional integration a trigger is needed to execute the KBS.

5. Acknowledgements

The work was supported by grant of the German Ministry of Education, Science, Research and Technology (MEDWIS – program, project A 70).

6. References

[1] R. Linnarson and O. Wigartz, The data dictionary - a controlled vocabulary for integrating clinical databases and medical knowledge bases, *Methods of Information in Medicine* **28** (1989) 78-85.
[2] P. Coad and E. Yourdon, Object-oriented analysis and object-oriented design. Yourdan Press, Englewood Cliffs, 1991.
[3] P. Coad et al., Object models, strategies, patterns applications. Yourdan Press, Englewood Cliffs, 1995.
[4] C. Ohmann et al., Expertensystem zur Unterstützung von Diagnosestellung und Therapiewahl bei akuten Bauchschmerzen. *Informatik, Biometrie und Epidemiologie* **26** (1995) 262-274.
[5] C. Ohmann et al., Evaluation of automatic knowledge acquisition techniques in the diagnosis of acute abdominal pain. *Artificial Intelligence in Medicine* **8** (1996) 23-26.
[6] The Turing Institute, CN2 Software Systems, George House, 36 North Hanover St., Glasgow G1 2AD, UK, 1993.
[7] J. R. Quinlan, C4.5 Programs for Machine Learning. Morgan Kaufmann, San Mateo, CA, 1993.
[8] C. Ohmann et al., Diagnostic scores for acute appendicitis, *Eur. J. Surg.* **161** (1995) 273-281.