The Use of Ontologies for Medical Guideline Servers

VESSELIN DETSCHEW, ANNE-KATHRIN KAEDING, GERT FUNKAT
Institute of Biomedical Engineering and Informatics
Ilmenau Technical University
P.O.Box 100 565, D-98684 Ilmenau
GERMANY

Abstract: - The correct manipulation of semantic information is essential for the development and implementation of medical guideline server. This is a knowledge based information system for computer-aided processing of clinical practice guidelines. In this study the following core functions are involved in the data processing of guideline server: (1) information transfer from other hospital information systems, (2) knowledge processing and (3) reporting functions. An execution engine is developed for guideline knowledge processing by our group. The ontological structures of the knowledge are the results of a knowledge engineering process. Additionally, the terminological standardization by Unified Medical Language System (UMLS) is integrated into the execution engine. The applied methodical approach for the ontology development is explained as a part of the knowledge engineering process. The developed ontological structures are used as interface descriptions for data and information transfer from electronic health record (e.g. test for plausibility) as well as for report generation and documentation (i. e. generation and explanation of terms and term sequences). The developed concepts are presented in this paper.

Key-Words: - knowledge based information system, medical guideline server, knowledge processing, ontology, UMLS, information transfer, interface description

1 Introduction

The relevance of clinical practice guidelines is growing in an extensive way in clinical environment. Each guideline should provide systematic decision support to appropriate health care for specific clinical circumstances. The German definition of clinical practice guidelines uses the definition of the US Agency for Health Care Policy and Research: "Clinical practice guidelines are systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances [1]."

Guidelines represent the actual state of medical knowledge at the time of its development. They describe the generally recognized standard of treatment, mainly containing results of the so called "Evidence Based Medicine". The use of evidence based medicine is a core demand of public health care caused by quality assurance reasons. Because medical knowledge itself increases constantly, every guideline must be checked and updated continuously.

Printed guidelines are the most used solution. Existing guidelines are unstructured papers, sometimes supported by diagrams or decision trees. Their observance is difficult to check. The daily application is impractical and often not accepted by the staff. Updating of printed guidelines is very expensive.

Guideline server is a knowledge-based software system integrated in the hospital information system. The support of the medical intervention requires automatically generated advices or warnings under critical conditions. This requires a relevant knowledge base for the problem domain as well as the knowledge-processing components. The disadvantage is high expenditure for development and implementation.

Integrating a guideline server into existing information systems requires data transfer from other information systems in the hospital. The heterogeneity and the huge number of data sources make it difficult to use data correctly.

2 Ontologies

The term 'Ontology' is used to refer to a common understanding of a particular domain of interest. The intended goal of ontologies is the solving of communication problems between people, organizations and technical systems. The most cited definition of ontology gave Gruber in [1]: "An ontology is an explicit specification of conceptualization."

The use of ontologies concerns the following aspects [2]:

- Communication between people with different needs and viewpoints arising from their different contexts,
- Compatibility between software systems;
- Re-usability: Ontologies represent a formal encoding for important entities, attributes, processes and their inter-relationships in the domain of interest. The automatic translation of this formal representation results in a re-usable component for a software system.
- Reliability: A formal representation allows the automation of consistency check increasing the reliability of software systems.
- Specification: Ontologies facilitate the process of identifying requirements and defining a specification for software. This is of particular relevance when the requirements cover different groups using different terminology for the same domain, or multiple domains.

In the domain of medicine there are a huge number of different medical vocabularies and classification systems used by medical information systems like electronic patient record and hospital information system. As a result of this situation a certain patient parameter has synonymous names. This fact causes difficulties in using of different information sources.

A general standardization approach arose by the development of the Unified Medical Language System (UMLS) [3]. The aim of this medical ontology is the integration of information from multiple machine-readable biomedical information sources.

UMLS is based on a long-term research of the National Library of Medicine (NLM). The UMLS approach disposes three "Knowledge Sources": the Metathesaurus, the Semantic Network and the SPECIALIST lexicon.

The Metathesaurus represents the central vocabulary of the UMLS. This database contains information from more than 60 thesauri, classifications and coding systems. For each concept the meaning, several attributes and its hierarchical linkages with other concepts are determinated (see Figure 1).

Each concept has an *unique concept identifier* (CUI). Concepts with the same meaning possess the same CUI. Each unique concept name or string in each language in the Metathesaurus has a *unique string identifier* (SUI).

Concept Name:	Atrial Fibrillation
UI:	C0004238
Semantic Type:	Finding ; Pathologic Function
Definition:	Disorder of cardiac rhythm characterized by rapid, irregular atrial impulses and ineffective atrial contractions.
Synonyms:	Auricular Fibrillation | AF - Atrial fibrillation | afib | a fib | af
Sources:	MSH2001, MTH, PSY97, CCPSS99, COS89, CST95, DXP94, CCS99, SNMI98, ...
Other Languages:	eteisvaerinae – Finnish FIBRILLATION Vorhofflimmern – German …

Figure 1: UMLS Metathesaurus entry for concept 'Atrial Fibrillation' from [4].

The UMLS Semantic Network defines categories of all concepts, represented in the Metathesaurus and the important relations between them. These categories are called Semantic Types. Figure 2 shows a part of the Semantic Network for the semantic type *Biologic Function*.

The hierarchies of Semantic Types is realized by using of *Is-a*-relations. Additionally, a set of non-hierarchical relations between the types are defined. These are grouped into five major categories, which are themselves relations:

- *physically related to*,
- *spatially related to*,
- *temporally related to*,
- *functionally related to*
- *conceptually related to*

The SPECIALIST lexicon provides lexical information for each spelling or set of spelling variants of a concept. Each lexical entry consists of a set of different slots and its fillers.

In this approach the UMLS is used to identify and define data from medical information systems to provide an integrative use within an application ontology.

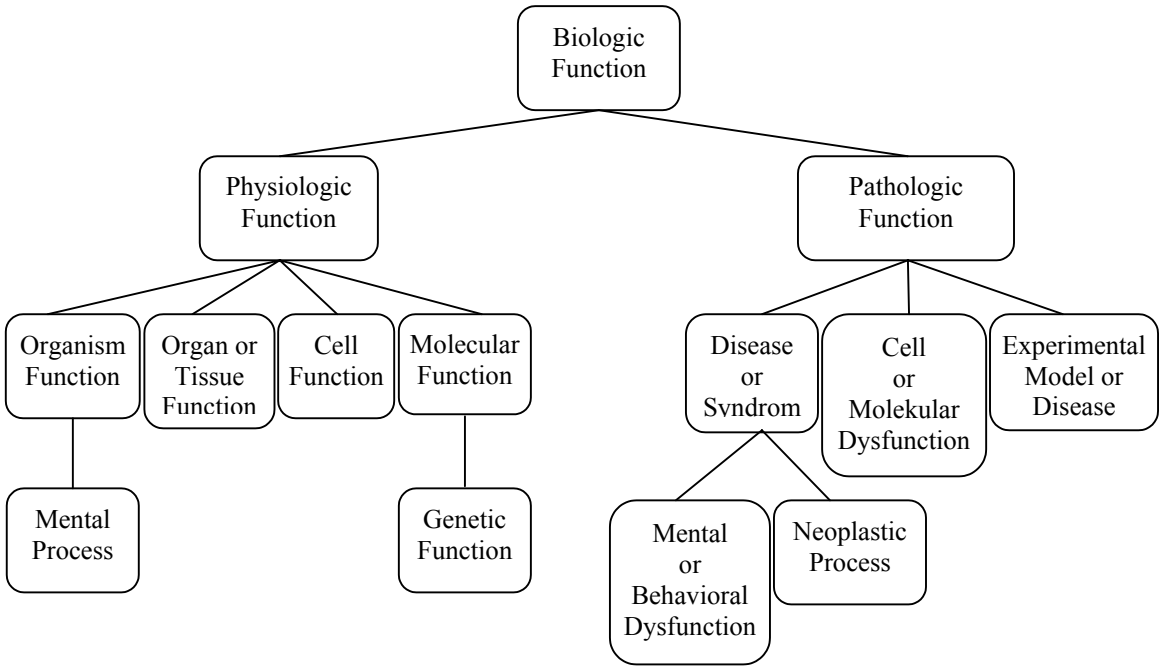

Figure 2: Part of the UMLS Semantic Network for the Semantic Type Biologic Function [3].

3 The Use of Ontologies for Medical Knowledge Information Systems

In this study we use ontologies for supporting the following core functions of a medical guideline server:

- information transfer from other hospital information systems,
- knowledge processing
- report functions.

The ontology for transfer of data and measurement values plays a key role caused by the heterogeneity and the huge number of information sources in hospitals. In further development stages ontologies are used for guideline knowledge processing and report generation and documentation.

The creation of ontologies for a specific domain requires the acquisition and structuring of concepts and their attributes.

From our point of view the following attributes are essential to describe patient data: name and meaning definition as textual description, synonyms, valid data sources, ranges of values, units of measurement, as well as period of validity.

Definitions and synonyms are used to allow editing, changing and extending data for non-physicians. Limiting values, units of measure and period of validity are necessary to check the validity of patient data by the help of ontologies.

For the validity check the actual patient data are matched against the corresponding slot within the ontology. In case of validity the data are redirected to the guideline server. On the other hand, guideline processing by the server will be documented using the vocabulary described in the ontology (see Figure 3).

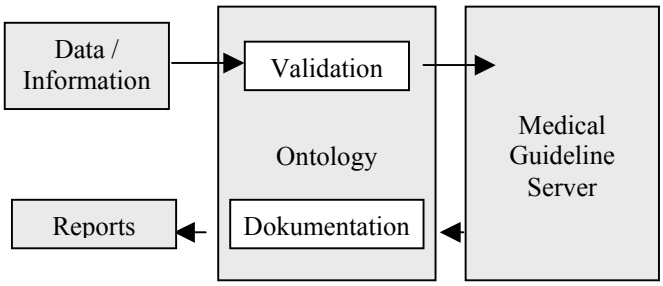

Figure 3: Validity check and documentation using ontologies

In medicine different terms for the same subject are often used. That's why it is necessary do define a standard within the system. This is the precondition for the clinic-independend shareability of an information system.

On the basis of UMLS (Unified Medical Language System) the data from the electronic patient record are classified by a term out of UMLS. This term will be declared as standard-term. The accompanying CUI (Unique Concept Identifier) and the definition from the Metathesaurus will be integrated into the

application ontology, too. The above mentioned regulations allow an adaptation to different hospital information systems by the implementation of a user specific mapping table between electronic patient record and guideline server.

UMLS is not complete at the moment. That's why, within the project there are so called private terms defined by the medical partners. For this terms corresponding UMLS representations and CUI do not exist yet.

4 Realization

Medical terms are subdivided in several areas (see Table 1). All of the terms constitute the ontologie. At the moment the ontology represents only the domain of traumatic brain injury. As a basis for structuring the UMLS Semantic Network was used. Table 1 show a part of the structuring of the concrete problem domain

This facilitates the possibility to add and edit data in an easy way. For the implementation we use XML schemata. XML schemata allow an easy implementation of data description. A XML-parser provides a validity check. In this project we use the XML4J-parser by IBM, based on the Apache Xerex version 1.4.

5 Summary

The introduced project has been carried out by the following steps: (1) Acquiring the medical knowledge; (2) Modeling the relevant guidelines; (3) Knowledge formalization with GLIF; (4) GLIF-model transferring into a computational form; (5) Knowledge processing by the execution engine; and (6) Graphical representation of the results based on the client/server-solution. All reported steps are supported methodically and technologically. Using ontologies has been proved very useful for interface description between guideline server and hospital information systems. We expect a similar effect from implementing the ontology for report function.

References:

[1] T. R. Gruber, A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition,* Vol. 5, 1993, pp. 199-220.

[2] M. Uschold and M. Gruninger, Ontologies: Principles, Methods and Applications. *The Knowledge Engineering Review*, Vol. 11, No. 2, 1996, pp. 93-136.

[3] *2001 UMLS Documentation*, 12th Edition - January Release 2001, http://www.nlm.nih.gov/research/umls/UMLSDOC.HTML

[4] *UMLS Knowledge Source Server*, http://umlsks.nlm.nih.gov (01.11.2001)

Table 1: Structuring of terms
(e.g. traumatic brain injury)

Conceptual Entity
- Glasgow coma scale
- GCS motor response subscore

Phenomenon or Process
- Heart Rate
- Intracranial pressure

Organism Attribute
- INTRAVASCULAR MEAN:PRESSURE:POINT IN TIME:PULMONARY ARTERY:QUANTITATIVE
- PCWP-Pulmonary capillary wedge pressure

Substance
- C-reactive protein
- Fibrinogen
- Hemoglobin
- Extravascular Lung Water
- Lactates

Activity
- Sodium measurement
- Osmolarity measurement
- Glucose measurement

Finding
- Rate of spontaneous respiration
- Arterial partial pressure of carbon dioxide
- arterial pO2
- Arterial oxygen saturation
- Total protein
- Mean blood pressure
- Diastolic blood pressure
- Systolic blood pressure
- Invasive mean arterial pressure
- Cardiac Output
- Venous mean pressure
- 24 hour urine volume level
- diuresis
- urine specific gravity level
- cerebral perfusion pressure