STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **LVI**, Number 2, 2011

CONCEPTUAL KNOWLEDGE PROCESSING GROUNDED LOGICAL INFORMATION SYSTEM FOR ONCOLOGICAL DATABASES - EXTENDED ABSTRACT

A. BRAD (1), L. NEAMŢIU(2), S. RĂUSANU(1), AND C. SĂCĂREA(1)

Abstract. Conceptual Knowledge Processing is a fundamental paradigm in data analysis and knowledge management. We use several methods of Conceptual Knowledge Processing to build a Logical Information System (LIS) for Oncological Databases. This paper describes this approach by analyzing the use of these methods on a cancer registry and discusses the main features of this LIS.

1. Introduction

Conceptual Knowledge Processing is a particular approach to knowledge processing, underlying the constitutive role of thinking, arguing and communicating human being in dealing with knowledge and its processing. The term processing also underlines the fact that gaining or approximating knowledge is a process which should always be conceptual in the above sense. The methods of Conceptual Knowledge Processing have been introduced and discussed by R. Wille in [5], based on the pragmatic philosophy of Ch. S. Peirce, continued by K.-O. Apel and J. Habermas.

R. Wille defined Conceptual Knowledge Processing to be an applied discipline dealing with ambitious knowledge which is constituted by conscious reflexion, discursive argumentation and human communication on the basis of cultural background, social conventions and personal experiences. Its main aim is to develop and maintain methods and instruments for processing information and knowledge which support rational thought, judgment and action of human beings and therewith promote critical discourse (see also [2], [3], [4]).

Our approach on building a Logical Information System for Oncological Databases has been motivated by this understanding of knowledge and its processing. Moreover, the promotion of critical discourse in acquiring, processing, retrieval and/or approximating knowledge is the grounding principle in developing this system. Its main aim is to support human thought, judgment, and action. This implies a certain understanding of what knowledge is. Knowledge is considered to be much more than a

---

Received by the editors: March 12, 2011.

2000 *Mathematics Subject Classification.* 68P15, 03G10.

1998 *CR Categories and Descriptors.* H.4.2 [**Information Systems Applications**]: Types of Systems – *Decision support*; G.2.3 [**Discrete Mathematics**]: Applications – *Applications*.

Key words and phrases. Conceptual Knowledge Processing, Formal Concept Analysis, Knowledge Acquisition, Logical Information System.

collection of facts, rules, and procedures, i.e., a cognitive-instrumental understanding of knowledge. K.-O. Apel advocates in [1] for a pragmatic understanding of knowledge. Hence, as has been stated by Wille in [5], the methods of Conceptual Knowledge Processing can only be successfully applied if discourses can be made possible which allow the users and the persons concerned to understand and even to criticize the methods, their performances, and their effects.

The mathematical theory underlying the methods of Conceptual Knowledge Processing is Formal Concept Analysis, providing a powerful mathematical tool to understand and investigate knowledge, based on a set-theoretical semantics, developing methods for representation, acquiring, retrieval of knowledge, but even for further theory building in several other domains of science.

2. Conceptual Scaling

A **many-valued context** (G, M, W, I) consists of sets G, M and W and a ternary relation $I \subseteq G \times M \times W$ for which holds that

$$(g, m, w) \in I \quad and \quad (g, m, v) \in I \Rightarrow w = v.$$

The elements of G are called *objects*, those of M (*many-valued*) *attributes* and those of W *attribute values*. The fact $(g, m, w) \in I$ is read as *the object g has the attribute m with value w*.

Conceptual Scaling is the process of transforming a many-valued context into a binary one, in order to assign formal concepts to the many-valued context. This gives rise to an *interpretation process*, the concepts of the *derived* binary context are interpreted as concepts of the original many-valued context. This process is not uniquely determined, the concept system of a many-valued context depends on the scaling.

The cancer registry database, in its original form, contains 25 attributes for each patient: Tumor sequence, Total number, Incidence, Topography, Morphology, Cause of death, etc. are just a few of them.

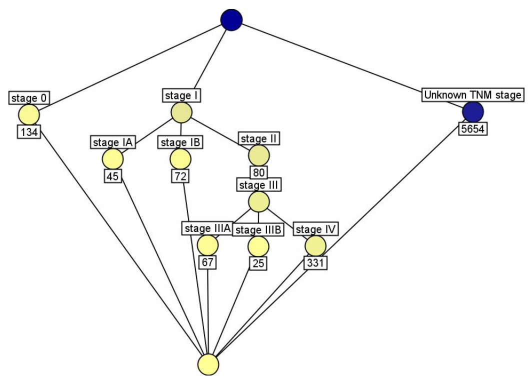

Figure 1. TNM stage scale

3. Analysis of Concept Hierarchies

The scales mentioned in the previous paragraph are included in the knowledge management system Toscana which are used to search structures to the objects under consideration. This is done by using the *conceptual landscape* paradigm. Knowledge is organised by scales and represented by conceptual hierarchies in diagrams, which can be aggregated in order to highlight knowledge structures and concept patterns. We will present some scenarios for browsing the knowledge structure of our database in order to discover concept patterns and to retrieve knowledge.

Graphically represented conceptual hierarchies prove to be a very efficient tool for the discovery and understanding of complex relationships between the concepts in which knowledge is aggregated.

In the following, we will describe one scenario, for more please refer to the main paper. These scenarios are simple aggregations between certain diagrams.

Treatment - Survival - Vitality - Cause of death. In the diagram comprising all the types of treatments, divided by curative and non-curative, we have the option of selecting one curative treatment and check for the survival period. The survival period is not maximized for all the patients, which can help to conclude that the effect of a curative treatment is rather relative. We can move forward to the vitality diagram to check the current status of the patient(alive/dead) as the survival diagram does not provide such information. Indeed, for most of the patients the vitality is alive, however there are some cases in which the patient has died and in most cases the cause of death was the cancer. This aggregation can go on by adding also the topography diagram or the morphology diagram to find out which is the *strongest* type of cancer. Another extension can be considered the age diagram to find a reason why death has occurred (due to the old age).

4. Conceptual Knowledge Inferences

A knowledge structure is not only characterized by its concepts and their hierarchy, but also by inherent inferences. We focus in our research on dependencies, implications, and associations.

For example we could check if there is an association rule that contains age >65 in the premises and does not contain prostate cancer in the conclusions, having a minimal support of 20 and a confidence of at least 60%. This feature allows the user to conduct a more organized browsing of the association rules and thus obtain information that is relevant and structured more easily.

5. LIS at a glance

The central idea of the LIS is the current context, which at the beginning is the entire context comprising all data, and by querying and browsing it can be reduced or extended (the context always contains all the attributes and the objects that constitute the intent of the concept whose extent is specified by the current conditions). All navigation is always performed starting from the current context. This LIS will thus allow the extension of the context by removing conditions (reducing the intent of the sought concept) or its reduction by adding other conditions (growing the intent of the sought concept), either through additional querying or step-by-step browsing. For example, the user might ask to see al patients diagnosed with a digestive apparatus tumor who are still alive. After obtaining the result, he may want to give up the vitality constraint or further add new conditions, like filtering only the patients that have received surgical treatment. The current context will always change accordingly.

References

[1] K.-O. Apel: Begründung. In: H. Seiffert, G. Radninzcky (eds.): Handlexikon der Wissenschaftstheorie. Ehrenwirt, München, 1989, 14-19.
[2] Wille, R.: Plädoyer für eine philosophische Grundlegung der Begrifflichen Wissensverarbeitung. In: R. Wille, M. Zickwolff (eds.) Begriffliche Wissensverarbeitung – Grundfragen und Aufgaben, B.I.-Wissenschaftsverlag, Mannheim 1994, 11–25.
[3] Wille, R.: Conceptual landscapes of knowledge: a pragmatic paradigm for knowledge processing. In: G. Mineau, A. Fall (eds.): Proceedings of the International Symposium on Knowledge Representation, Use, and Storage Efficiency. Simon Fraser University, Vancouver 1997, 2-13.
[4] Wille, R.: Begriffliche Wissensverarbeitung: Theorie und Praxis. Informatik Spektrum 23 (2000), 357 - 369.
[5] Wille, R.: Methods of Conceptual Knowledge Processing, in Proceedings of the 4th International Conference ICFCA 2006, LNAI 3874, Springer Verlag, 2006, pp. 1–29.

(1) Babeş-Bolyai University, Department of Computer Science, Kogălniceanu 1, Cluj-Napoca

E-mail address: brad.alexandru@gmail.com, silvia.rausanu@gmail.com, csacarea@math.ubbcluj.ro

(2) I. Chiricuţă Oncological Institute, 34-36 Republicii Str., Cluj-Napoca

E-mail address: luciana@iocn.ro