Complex Information Management Using a Framework Supported by ECA Rules in XML

Bing Wu, Essam Mansour and Kudakwashe Dube

School of Computing, Dublin Institute of Technology Kevin Street, Dublin 8, Ireland
<firstname>.<surname>@dit.ie

Abstract. It is every organization's desire to incorporate best practice into its enterprise. This incorporation gives rise to the need to maintain information that could be viewed as complex. Managing this complex information poses a major challenge to the area of information management. This paper presents a framework for the incorporation of best practice and subsequent management of the resulting complex information. The paper also presents an approach to supporting this framework by using the ECA rule paradigm with an XML-based language, called AIM, for specifying and querying best practice and the complex information.

Keywords: Information management, ECA rule paradigm, active database, XML, rule markup languages.

1 Introduction

Domains, such as patient care practice and securities trading order management, involve cross-domain activities that require the application of best practice, e.g., knowledge from research and experience, and also require constant monitoring of some vital signals or essential happenings within the domain. Computerizing best practice in the form of such *complex information*, as conceptualized here, could be seen to pose major challenges for information management. The first challenge is not only a matter of disseminating the best practice within application domains, but also a matter of ensuring the computer-interpretation for its specifications. The second challenge is to support the customization of best practice to suit a specific domain user and task. The third challenge is the incorporation of best practice into the computerized execution of tasks that uses the *complex information*, and keeping the execution history in order to review the evolution of the *complex information*. The fourth challenge is to facilitate the manipulation, query, and review by replay operation as well as sharing and distribution for this information.

This paper presents a novel approach and framework that incorporate application best practice into organizations' enterprise, and manage the complex information produced by such incorporation. Our approach and framework is restricted to applications that naturally take the form of reactive applications that monitor events of interest to domain users, and respond to changes in situations by issuing alerts, reminders, requests, and/or observations to domain users. In order to adapt to the

needs of users, markets, and enterprise management, the decentralized management is considered as one of the key aspects of our approach and framework. The method adopted to realize the approach and framework aims at utilizing the generally available highly optimized and easily maintained technologies, such as the *Event-Condition-Action (ECA)* rule paradigm [1] and XML technologies, to provide a Web-enabled tool that assists domain experts and users in managing the *complex information*.

The rest of this paper is organised as follows: Section 2 discusses the problem of managing the complex information; Section 3 outlines related work; Section 4 presents our approach to modelling the complex information; Section 5 outlines the SEM framework for managing complex information through the three planes covering specification, execution and manipulation of the information; Section 6 explains the role played by ECA rule paradigm and XML in supporting the SEM framework; Section 7 presents a conceptual architecture for the proposed prototype system that implements the framework; and finally, Section 8 summarises and concludes this paper.

2 Problem of Managing Complex Information in Computerised Best Practice

Best practice exists in several forms, such as recommendations from recent research results, practical experience, expertise and new methods of doing things in a given domain. Incorporating the best practice into application domain tasks and activities results in the need to manage *complex information*, whose main components include: 1) The domain information or data items that is relevant and is therefore required to be monitored in the application of knowledge and best practice; 2) Knowledge from the best practice and experience that is applied in consideration to the user's preferences or situations; 3) A description or reference material associated with the specific area of focus; and 4) The history and experience arising from daily practice of using the best practice knowledge in the domain. Consider as example the stock market trading domain, in which the customer order is specified by a customer who intends to buy or sell securities at the prevailing market price. The customer orders have c*onditions* and *constrains* that determine whether a transaction will occur and at what price. The customer order could be considered to be a body of *complex information* that consists of: 1) Information on particular stock items making up the order. This information is also a subject of monitoring; 2) Knowledge that represents constrains on the monitored stock items defined by the customer and the best practice that is used to continuously adjust the customer orders, as soon as stock items are changed, 3) The order history that shows the changes in the stock items and the constrains, over the life-cycle of the order. The order evolution history represents experience that assists in making decisions regarding future orders, 4) Descriptive information for the order, such as the order number and date, as well as any relevant didactic information that may be of help in educating stakeholders.

It is required to manage such information through the specification, execution and manipulation as one distinct entity using a high level and declarative method.

3 Related Work

It is generally recognised that best practice need to be incorporated into daily practice and further be provided through computer-based support mechanisms. Several works in different research areas, such as a) the active database; b) workflow and business process; c) AI and decision making, address the problem of incorporating best practice into application activities. Our work is distinguished from other research efforts as explained in the following paragraphs.

In the area of active database, several research efforts have utilized ECA rule paradigm [1] to combine best practice (business logic) and the database together into the DBMS, such as [2, 1]. According to this approach the best practice is represented as individual triggers, and the domain information is represented as relations (tables). Both, the best practice and the domain information are managed separately. This approach suffers from 1) the problems of unexpected interactions among the individual rules as the rule base size increases [2]; and 2) the gap between the end-user's needs of managing specific situations and the low-level representation using triggers. Our approach provides a high level and declarative manner to bridge this gap, and classifies the rules into sets.

In the area of workflow and business process, several research efforts, such as [3, 4], attempt to standardize the applications activities, according to the application best practice, as processes. This approach differs from ours in that it address the specification aspects with little or no support to query and manipulate all aspects of the information in a unified manner. However, our approach aims at automating the *complex information* that includes dynamic and static aspects within a unified framework that provides support to specify, query, and manipulate the dynamic and static aspects of the *complex information.*

Comparing with decision making and AI techniques, our work focuses only on assisting the decision-making process by issuing notifications, reminders, and/or observations regarding situation of interest to the domain user, or by executing predefined changes.

At a general level, existing works provide little or no support for comprehensive unified management of the information and processes derived from best practice in domain applications. Our approach facilitates the incorporation of best practice into application activities requiring constant monitoring, and provides decentralized management support for all information as one distinct entity produced from such incorporation.

Incorporating the ECA rule paradigm into XML has being addressed in several research areas, including *active XML*, *RuleML*, and *semantic Web.* In the area of active XML, the ECA rule paradigm [1] is incorporated into the XML to support active behavior over XML data, such as in [5, 6], which support the reactive applications at the level of rules and triggers only. The RuleML language aims at providing a standard rule language that is interoperability platform [7, 8]. Wanger in [7] classified the ECA rule paradigm as a subtype of the RuleML language. The RuleML language has been utilized to support semantic Web and business applications, such as in [9]. However, the RuleML language formalizes the application best practice as individual rules, not as a unified distinct entity.

4 An Approach for Modelling the Complex Information

Fig. 1 illustrates a model for the complex information by using a UML class diagram. The components of a complex information item for a given application domain are: a knowledge and action component that models aspects such as reactive behaviour, a domain information component that models monitored information, an evolution history component and a descriptive information component. These pieces of information could be classified and viewed as *active* and *passive* information.

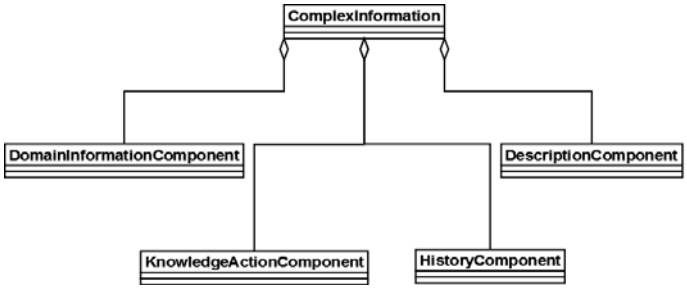

Fig. 1. The model for complex information.

Here, the *active* information is modelled by the knowledge action component that determines the reaction that should be taken as response to changes in the monitored domain information. The knowledge action component could constitute reactive behaviour that could be practically represented as modularized sets of rules. The *passive* information is the evolution history as well as the domain and descriptive information, which are essentially of a factual nature.

The approach adopted in this study emphasizes the management of the *complex information* (CI) as one distinct entity as follows: The CI 1) has a general structure specified according to the best practice of an application; 2) deals with particular situations according to the user's preferences and interest; 3) is executed, as soon as a change of interest happens to the monitored information. The execution of the CI provides notification and suggestions to the user and might modify the knowledge action component; 4) could be manipulated and queried as a first class object, not only as row data; 5) is provided a distributed management that supports the remote users and distributed applications. The approach's unique features include: 1) managing the active and passive information of the CI as one object, not as a separated objects, such as in the active database the passive information represented as relations and the active information managed as individual triggers; 2) a high level declarative language for specifying, manipulating, and querying the CI; 3) utilizing an event-driven mechanism incorporated with XML technologies; 4) support for the decentralized management of the CI.

5 SEM Framework for Managing Complex Information

This Section presents a framework, called SEM, for incorporating best practice into the organization activities and managing the *complex information* produced from such incorporation. The SEM framework extends and enhances the framework developed by Dube in [10] for managing clinical guidelines. A key aspect of information management is the handling of knowledge acquired by one or many disparate sources in a way that optimizes access and use by all who have a share in that knowledge or a right to that knowledge. Information management in this work involves the administration of information, its uses and transmission and spans the entire process of defining, evaluating, protecting, and distributing data and information within an organization. It also includes the planning, budgeting, manipulating, and controlling of information throughout its life cycle. To be comprehensive enough to incorporate this conceptualization of information management, a framework for computer-based support for information management must include the three planes: *specification*, *execution* and *manipulation* as illustrated in Fig. 2.

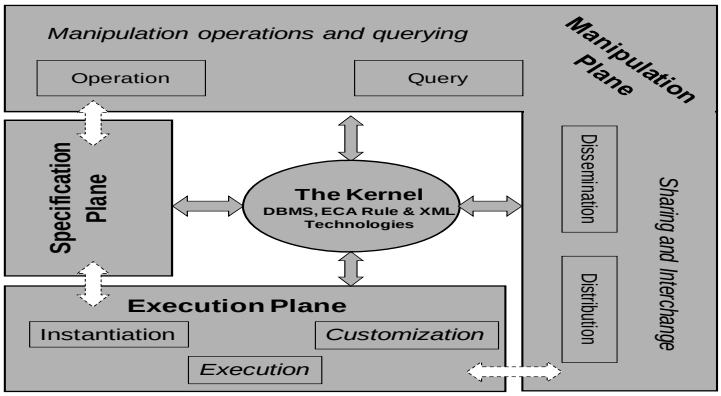

Fig. 2. SEM: The complex information management framework.

The *Specification Plane* provides for the requirement of creating specifications that are essential to capture domain knowledge and information. It also involves specifications that are essential to enable the execution of instances. The *Execution Plane* focuses on meeting the requirement for allowing the execution of the aspects of domain knowledge that is amenable to execution by using mainly an event-driven approach based on the ECA rule paradigm. However, other execution formalisms are also possible. The *Manipulation Plane* provides for the need to be able to manage domain knowledge and information through: performing manipulation operations; issuing various types and forms of queries; enabling the sharing by allowing dissemination and interchange of knowledge and information; and providing for all these within distributed contexts. The dotted white arrows in Fig. 2 illustrate the fact that these three planes also interact in a dynamic fashion and offer services to one another.

6 The Role of ECA Rule Paradigm and XML in Supporting SEM Framework

Realizing SEM as a unified framework for managing the *complex information* (CI), in the reactive applications requires technologies that: 1) can be seamlessly integrated and easily incorporated, 2) support the monitoring process and distributed management. The ECA rule paradigm incorporated into XML technologies are utilized here to support the management of the CI. Furthermore, the modern DBMSs, such as Oracle and DB2, provide support for ECA rule paradigm and XML technologies. Consequently, the event-driven and XML technologies are adopted to play a crucial role in providing the basis for the management of the CI. The benefits of this method include: the flexibility of managing the CI as one unit, and the easy integration of the complex information management system into other systems. This method facilitates the development of the proposed approach for modeling the CI, a management language and a decentralized system for the complex information. Based on this method a language, called AIM [11], is developed such that: 1) it is an XML-based language and enjoys the general benefits of XML, such as parser reuse, incorporation into Web services, query generation; 2) It has an ECA- and XML-based specification component language, called AIM-SL, which formalizes the application best practice into interpretable format; 3) it has a high level XQuery-based component language, AIMQL, that provides support to manipulate, query, and has an extension for replaying the CI. The AIM language provides a CI model, which is developed as an active temporal XML document. This document includes: 1) the knowledge and action component represented as ECA rules; 2) the history component represented using temporal XML mechanism; 3) the domain information and descriptive components represented using conventional XML. The service-oriented architecture incorporated into the modern DBMS is utilized to develop a system, called AIMS. AIMS manages the complex information in distributed environment, as outlined in Section 7

The SEM framework based on the use of ECA rules with specifications in XML offer a number of benefits. Firstly, the framework unifies all aspects of the management of information and fosters a holistic perspective to the management of the information and knowledge. Secondly, a high level of flexibility can be afforded by the ability to make on the fly specification and modifications during execution. Thirdly, the ECA rule paradigm in XML provides the best practical compromise for best practice knowledge incorporation into, and sharing among, information systems and integration with domain databases.

7 Conceptual Architecture for A Complex Information Management System

Fig. 3 illustrates the abstract and conceptual architecture for the proposed prototype system, called AIMS, that supports the *SEM framework* for managing complex information in terms of the three planes. The implementation of the *Manipulation Plane* is achieved through the query and manipulation management. The effect of the

manipulation plane will be reflected in the database as well as the ECA rule mechanism. The implementation of the *Specification Plane* is attained through the information and knowledge management, which allow users to specify knowledge using AIM [11].

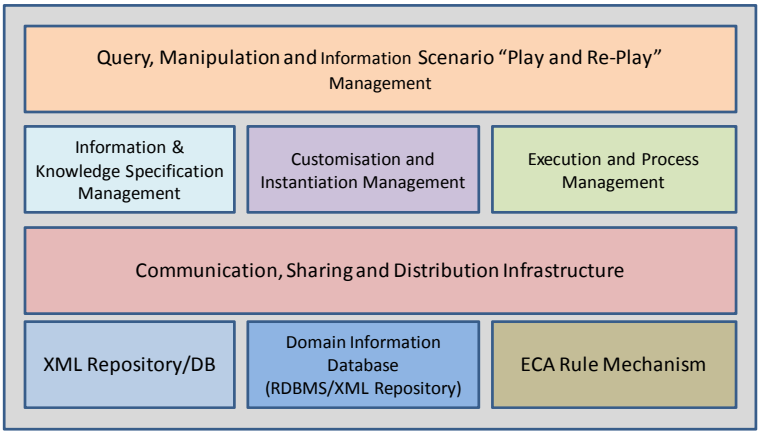

Fig. 3. Conceptual architecture for a system for supporting the SEM framework.

The resulting specifications are stored in the database. The implementation for the *Execution Plane* is attained through the customization and instantiation, and the execution and process management functions as well as the ECA rule mechanism within the database or XML repository. The operations and queries of the manipulation plane should also be able to apply to information in both the execution plane and specification plane.

The SEM framework offers a comprehensive way to manage information such that the specifications and the execution process and results are enabled for the manipulation operations and queries within the same framework. The ECA rule paradigm as implemented within the context of databases or of XML repositories offer a practical way of incorporating best practice into routines used daily.

8 Summary and Conclusion

The incorporation of best practices into the computer-based tasks and activities of an organisation gives rise to the need to maintain information that could be viewed as complex. Managing this complex information poses a major challenge to the area of information management. This paper has presented the SEM framework for the incorporation of best practice and an approach to the subsequent management of the resulting complex information. Unlike others, the SEM framework does not focus only on one aspect of information management but incorporates planes for specification, execution and manipulation. The approach to supporting this framework is based on the use of the ECA rule paradigm within an XML-based language, called AIM, for specifying domain knowledge and information. Specifying the ECA rule

paradigm in XML allows domain specific formal annotations and structuring of information and knowledge that is manageable and sharable across technologies and domains. This work demonstrates that RuleMLs provide a strong basis for realising the comprehensive SEM framework for knowledge and information management. Thus, every organization's desire to incorporate best practice into its enterprise would appear to be made easier.

References

1. Jennifer Widom and Stefano Ceri, *Active Database Systems: Triggers and Rules For Advanced Database Processing*. 1996: Morgan Kaufmann.
2. Dayal Umeshwar, Barbara T. Blaustein, Alejandro P. Buchmann, et al., *The HiPAC Project: Combining Active Databases and Timing Constraints.* SIGMOD Record, 1988. 17: p. 51-70.
3. Hyeong-Seok Hong, Bong-Seok Lee, Kwang-Hoon Kim, et al. *A Web-Based Transactional Workflow Monitoring System*. in *Proceedings of the First International Conference on Web Information Systems Engineering (WISE'00)*. 2000: IEEE Computer Society.
4. A. Tsalgatidou, G. Athanasopoulos, M. Pantazoglou, et al., *Developing scientific workflows from heterogeneous services.* SIGMOD Rec., 2006. 35(2): p. 22-28.
5. Serge Abiteboul, Omar Benjelloun, Ioana Manolescu, et al. *Active XML: A Data-Centric Perspective on Web Services.* in *BDA*. 2002.
6. Angela Bonifati, Daniele Braga, Alessandro Campi, et al. *Active XQuery*. in *Proceedings of the 19th International Conference on Data Engineering ICDE*. 2002. San Jose (California).
7. Gerd Wagner, Grigoris Antoniou, Said Tabet, et al. *The Abstract Syntax of RuleML - Towards a General Web Rule Language Framework*. in *Web Intelligence*. 2004.
8. Gerd Wagner, Adrian Giurca, and Sergey Lukichev. *A General Markup Framework for Integrity and Derivation Rules*. in *Principles and Practices of Semantic Web Reasoning*. 2005.
9. Enrico Pontelli, Tran Cao Son, and Chitta Baral. *A Framework for Composition and Inter-operation of Rules in the Semantic Web*. in *RuleML*. 2006.
10. Kudakwashe Dube, *A Generic Approach to Supporting the Management of Computerised Clinical Guidelines and Protocols. PhD Thesis*. 2004, Dublin Institute of Technology (DIT).
11. Essam Mansour, Kudakwashe Dube, and Bing Wu. *AIM: An XML-Based Temporal and ECA Rule Language for Managing Complex Information*. in *submitted to: International RuleML Symposium on Rule Interchange and Applications (RuleML 2007)*. 2007. Orlando, Florida.