Semantic Interoperability for Financial Information: a Component-based Approach

Bo Hu, Keqing He, Jian Wang
State Key Lab of Software Engineering
Wuhan University
Wuhan, China
hubo_wuhan@foxmail.com

Abstract—Semantic interoperability challenges emerged when comes to financial system integration and information exchange. This paper presents a component-based approach for financial information semantic interoperability. In the approach, taxonomies in XBRL GL have been modeled as core components in CCTS, which act as common understandings for interoperability participants. Additionally, guidelines for selecting and reusing these component-based XBRL GL taxonomies would be introduced to instruct users to assemble contextualized financial information.

Keywords-XBRL; XBRL GL; CCTS; Semantic Interoperability

I. INTRODUCTION

Enterprise resource planning (ERP) is an integrated computer based system used to manage internal and external resources in an organization, including financial resources such as information on accounting, auditing, taxation, etc. Data in an ERP could be migrated flexibly from one integrated system to another. However, not all enterprises can afford the ERP for its gigantic expense on system launching and maintenance, especially to the small and medium enterprises. Most of the financial information systems in sectors of these enterprises today are proprietary and often only designed and deployed for specific use, which brings the interoperability challenges to achieve seamless business transactions and data integration among these systems in an inexpensive manner.

Interoperability is a property referring to the ability of two or more systems or components to exchange information and to use the information that has been exchanged [1]. Issues of interoperability have been concerned over more than two decades and evolved into three levels [2]. The first level of interoperability is occurred in early 1980s, which emphasizes communications among computer systems, databases and other software systems. The issue of system interoperability has been settled by communication protocols such as TCP/IP. The second level of interoperability is syntactic interoperability, specified data formats, middleware (e.g. COBAR) and XML based techniques (e.g. Web service) are effective means to this issue. The real challenge of interoperability now is on the semantic level, where semantic interoperability is the ability to automatically interpret the information exchanged meaningfully and accurately in order to produce useful results as defined by the end users of both systems.

In this context, the authors would present a novel approach about semantic interoperability for financial information. In the approach, information modeling in financial information systems must be done first, the information models enriched with common semantics which would play an important role in enabling semantic interoperability among these financial information systems. Furthermore, a prototype should be developed following the approach. After experimental usage of the prototype, this approach has been proven to be effective in addressing the issue of semantic interoperability for financial information.

The remainder of the paper is structured as follows: In Section II, issues and challenges in semantic interoperability are discussed by providing the state of the art background and related work upon which this work is based. The novel approach for modeling financial information is elaborated on in Section III. Section IV presents the design and implementation of a prototype for semantic interoperability. Section V is the conclusion including the future work.

II. BACKGROUND

In the background of our research, requirements of semantic interoperability on financial information are elicited from a scenario. Related work is also briefly introduced, which act as the base of our approach.

A. Requirements of the Approach

Financial information systems and services could be found almost in every financial specific sectors of an enterprise. It could be an accounting system in Finance Department used for recording and processing accounting transactions such as accounts payable, accounts receivable, payroll and trial balance. Or it may be an auditing information system in Audit Commission used for financial audit. What’s more, it also could be a decision-making service for decision-makers to manage enterprise business. Currently, systems and services aforementioned are now physical deployed separately, staffs from Audit Commission having to ask for necessary information from Finance Department to enable the auditing information system running. Both staffs from Finance Department and Audit Commission should submit their reports to enterprise management for decision making. When financial information exported from a system or a service, they should
principles and rules without taking into account integration within or across their boundaries.

C. XBRL GL

XBRL GL (XBRL Global Ledger) [5] is a XBRL taxonomy framework which intends to enable the efficient handing of financial and business information scattered across disparate financial systems contained within an organization using XBRL specifications. XBRL GL was designed so generic and holistic that it serves as a bridge among all of the related sectors of an organization, allowing taxonomies defined in any sector representation in standardized charts. Figure 1 illustrates the framework of XBRL GL both from the logical and physical perspectives. Though the figure itself is not normative, it is used to illustrate the basic ideas.

![Figure 1. Framework of XBRL GL](image)

A complete taxonomy is consisted by assembling a conceptual element and a specific content model within a specific application. Taxonomies in XBRL GL could be divided into two main parts logically, the element declarations and the content model declarations. The part of element declarations consists of a set of module to define abstract concepts of taxonomy, including core concepts module COR and other extended concept modules such as BUS (Advanced Business Concepts), MUC (Multi Currency), USK (concepts for the US, UK), TAF (tax audit file), etc. Though content models of concepts could be also defined within the part of element declarations, most of them have defined in content model declarations separately. Content models are logical models depending on the combination of element declaration modules that are used in any financial information, where GEN is the general content schema of concepts among all of the concept definition modules, and PLT is a set of content model declarations about various combinations of element declaration modules.

XBRL GL ensures taxonomies shared and managed across sectors in an organization, which enables the semantics of information represented by these taxonomies found in a financial information system could be understood by other systems. However, XML format based XBRL GL couldn’t be a suitable mean to express semantics of
Components Technical Specification is a process of contextualized financial information. Taxonomies would be presented to instruct users on selecting and reusing these component-based XBRL GL understanding for interoperability firstly, and then guidelines for semantic interoperability for financial information is proposed. In this section, a novel approach about semantic interoperability for financial information is proposed. In order to facilitate converging, we would like to design a metamodel-mapping method for aligning concepts. Therefore, the first thing we should do is to find concepts that are similar to each other in these two methodologies, and then construct alignment rules among these similar concepts. Figure 2 depicts the concepts and the concept alignment between XBRL GL and CCTS.

### D. CCTS

Core Component Technical Specification, abbreviated as CCTS [7], is a cornerstone of the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT) standardization activities, which presents a methodology for developing a common set of semantic building blocks to represent the structure and contents of conceptual, logical, and syntax neutral data models, process models, and information exchange models in financial information systems and services. Metamodel in right part of figure 2 provides a glance of the key concepts of the CCTS, which are distinguished into two major parts: the Core Components (CC) and the Business Information Entities (BIE). Core Components are building blocks on a conceptual level. They provide a context independent template used for developing business information entities in a given business context with semantically and logically correct structure and content. Core components could be further divided into three types: Aggregate Core Component (ACC), Basic Core Component (BCC) and Association Core Component (ASCC). Basic concept behind Core Components is the identification of objects, simple object properties (e.g. data, name, age) and complex object properties (references to other objects) [8]. An object is represented using an ACC, where simple object properties are referred to as BCCs and their corresponding data types is known as Core Data Types (CDT, e.g. Identifier, Text), and complex object properties could be presented as ASCC. Accordingly, there are also three types of Business Information Entities: Aggregate Business Information Entity (ABIE), Basic Business Information Entity (BBIE) and Association Business Information Entity (ASBBIE). In addition, Business Data Types are types of BBIE, which is based on CDT.

CCTS as a content infrastructure specification now is used as upper standard in developing e-business standards by UN/CEFACT. In our opinion, considering its characteristics, CCTS could be also used in semantic interoperability for financial information.

### III. AN OVERVIEW OF THE APPROACH

In this section, a novel approach about semantic interoperability for financial information is proposed. In order to achieve this goal, we will use metamodel from CCTS to model XBRL GL taxonomies as semantic common understanding for interoperability firstly, and then guidelines for selecting and reusing these component-based XBRL GL taxonomies would be presented to instruct users assemble contextualized financial information.

#### A. Component-based XBRL GL Taxonomies Modeling

Modeling XBRL GL taxonomies under the Core Components Technical Specification is a process of converting XML described taxonomies filings into CCTS component compatible formats. In order to facilitate converting, we would like to design a metamodel-mapping method for aligning concepts. Therefore, the first thing we should do is to find concepts that are similar to each other in these two methodologies, and then construct alignment rules among these similar concepts. Figure 2 depicts the concepts and the concept alignment between XBRL GL and CCTS.

**Aligning element declarations to CCs.** In this concept alignment, modeling elements “Tuple” and “Item” from metamodel of XBRL GL are concepts used for taxonomy element declarations, which is the same as conceptual modeling using core components. Because “Item” is the single element and a “Tuple” is an element that will contain other elements, therefore, we decide to map “Item” in XBRL GL to “Basic CC” in CCTS, and map “Tuple” to “Aggregate CC”. It is worth noting that there isn’t an alignment rule in mapping any concept from XBRL GL to “Association CC” because there is no relationship defined in taxonomies element declaration.

**Aligning content model declarations to BIEs.** It is well known that content models of the same one taxonomy are different declaration in different applications. Taking taxonomy “account” as an example, whose element is declared just once in COR, while its content models are different declared in different subclass of PLT (Figure 3 shows). Therefore, we recognize that a content model of a taxonomy just as a core component used in specific context, so concept “ItemType” could be mapped to “Basic BIE” and “TupleType” could be mapped to “Aggregate BIE”. The relationship between two “TupleType’s” should be mapped to “Association” as well.

Besides above alignment rules, data types used in element declarations should be mapped to Core Data Types, and data types used in content model declarations should be mapped to Business Data Types. If above key concepts from the both sides could be mapped completely, then the alignment rules between XBRL GL and CCTS is established.
B. Contextualized Information Assembling

Two types of assemblers participate in to assemble contextualized information such as documents or messages: business user and system developer.

Business users are domain experts specialize in financial documents generation. Though these various documents could be automatic generated by existing financial information systems, when a format of these documents should be revised, it is also a tedious and time consuming process to provide new requirements to system developers for system reengineering. In our approach, contextualized information is assembled by component-based XBRL GL taxonomies in following three steps, as depicted in Figure 4. In the first step, conceptual core components should be selected, and assembled together as business information entities according to specific needs. And in the second step, this context dependent business information should be arranged in an order following a predesigned document template. In the last step, syntax neutral logical information documents should be translated into physical ones in a specific business environment. Once this specific information has to be changed, information systems needn’t to be reengineered any more, what should business do it to reusing these business information entities into a new document template.

Different from business users, system developers are technique experts who are good at IT infrastructures but do not familiar with finance. Unified Modeling Language (UML) is a standardized general purpose modeling language in the field of software engineering, which could be used to bridge the gap between IT with business. Therefore, A UML profile is supported in our approach to contextualized information assembling. Because XBRL GL taxonomies used in our approach is based on CCTS specification, so this UML profile originates from related UN/CEFACT standard UPCC (UML Profile for Core Components) [9], where a formal UML profile is proposed by mapping of CCTS to UML. Therefore, system developers could use popular UML tools for developing.

IV. IMPLEMENTATION

An initial prototype is designed and implemented in this section. Figure 5 illustrates the architecture of the prototype, which is mainly combined by modules of XBRL GL taxonomy converter, component storage, component assembler and UML tool Adapter.

XBRL GL taxonomy converter is a subsidiary but the most basic tool of our prototype. Prototype administrators would use this tool at first to convert taxonomies in XBRL GL into CCTS component-based counterparts, and these component-based taxonomies act as semantic common understanding in the process of financial information assembling.

Component storage is a database that provides functions for component-based taxonomies registry, repository and semantic selection. Domain experts from financial field enable the reasonable semantic relationships among these component-based taxonomies.

Component assembler and UML adapter are user interfaces for component-based taxonomies management and information generation. Specifically, component assembler is a business user tool implemented by JSP with a graphic user interface. Business users could use this tool to manage (including retrieving, browsing, adding, updating and etc.) their own conceptual component-based (CC) taxonomies, assemble these conceptual component-based taxonomies into context dependent business information entities, and generate specific financial information such as documents and messages. Another user interface, UML adapter, is only a programming interface for prevalent UML tools such as IBM Rational Rose and Enterprise Architect. Component-based taxonomies could be shown in these UML tools using the interface, and UML tools also could be manage, assemble and generate financial information following UML specification.

The prototype is used by our industry partners for financial information system developing and reengineering. Using the prototype, prototype administrators, domain
experts, business users and system developer could fascinate in their works much more convenience than ever.

V. CONCLUSION

This paper proposes a novel approach about semantic interoperability for financial information. In the approach, we define a set of mapping rules for concept alignment that allows taxonomies in XBRL GL could be transformed into CCTS forms. Component-based XBRL GL taxonomies are used as syntax neutral building blocks translate the meanings of information between systems. In addition, guidelines in our approach help both business users of financial information and systems developers to manage these component-based taxonomies and assemble them from conceptual and logical components to physical information. Furthermore, a prototype is designed for implement our approach. Experiences from our industrial partners shows that both our approach and the prototype running well in real.

In the future work, we would like to introduce our approach as a solution for semantic interoperability among various business standards.

ACKNOWLEDGMENT

The research presented in this paper would not have been possible without our colleagues and students at the DAMA CHINA, Shanghai Accounting Association, and Wuhan University.

This research was supported by the National Basic Research Program of China (973) under Grant No.2007CB310801, the National Natural Science Foundation of China under Grant No.60970017, 60903034 and 60803160, the National High Technology Research and Development Program of China (863) under Grant No.2006AA04Z156.

REFERENCES