12-31-2002

Modeling XBRL Using UML: Improving Semantics For Financial Analysis

Vijayan Sugumaran
Oakland University

Joseph Callaghan
Oakland University

Arline Savage
Oakland University

Modeling XBRL Using UML: Improving Semantics for Financial Analysis

Vijayan Sugumaran
Department of Decision and Information Sciences
School of Business Administration
Oakland University
sugumara@oakland.edu

Joseph H. Callaghan
Department of Accounting and Finance
School of Business Administration
Oakland University
callagha@oakland.edu

Arline Savage
Department of Accounting and Finance
School of Business Administration
Oakland University
savage@oakland.edu

Abstract

eXtensible Business Reporting Language (XBRL), a financial accounting application of XML, provides a taxonomy for facilitating the analysis of financial statement information. This taxonomy promises to enhance the financial analyst's ability to process financial information both cross-sectionally and temporally. However, limitations exist in the taxonomy that inhibit the analyst, intelligent agent or application to process information in the most effective way. In particular, financial statement element analysis can only be properly conducted when the additional semantics provided by additional disclosures are incorporated into the process. UML provides a conceptual framework to capture more meaningful semantics, particularly among related, collaborating objects. The purpose of this paper is to develop a framework that will couple the strengths of both technologies in order to develop better systems for financial statement analysis.

Introduction

Extensible Markup Language (XML) is emerging as a fundamental enabling technology for application integration and providing interoperability of data. XML is a set of rules for defining data structures and thus making it possible for key elements in a document to be characterized according to meaning. XML has several advantages. First, it is a descriptive markup language rather than a procedural markup language. Hence, it is possible to understand the semantics of the XML document. Second, it is vendor independent and therefore highly transportable between different platforms and systems while maintaining data integrity. Several standards are emerging in different domains that build on the XML technology to provide a standardized, vender-independent level playing field on which different systems may freely communicate. The eXtensible Business Reporting Language (XBRL) is an emerging XML-based open source specification for financial information, reporting, and analysis that has been jointly developed by more than 80 leading companies and organizations around the world. Its purpose is to facilitate the transfer of richly structured documents between applications over the Web. In a nutshell, it is a financial-based tagging language based on the XML meta-markup standard deemed by the World Wide Web Consortium. XBRL can indicate the values of data contained within a financial document.

There is a growing level of excitement in the financial community around the (XBRL) standard because of the expectation that it will make it easier to collect, aggregate and publish financial results from publicly held companies (Bovee et al., 2002; Coffin, 2001; Zarowin and Harding, 2000). According to XBRL.org, "The goal of XBRL is to provide an XML based framework that the global business information supply chain will use to create, exchange, and analyze financial reporting information including, but not limited to, regulatory filings such as annual and quarterly financial statements, general ledger information, and audit schedules." Thus, using XBRL, investment analysts can easily analyze and compare financial statements from different companies from an investment standpoint (Strand et al., 2001; Watson et al., 2000).

While this new standard facilitates the interoperability of financial data contained in various financial statements, the XBRL taxonomy tags footnotes and other disclosures in a way that does not reflect the natural way in which an analyst, intelligent agent or application works. For instance, before analyzing financial statement elements, an analyst would examine the auditor's opinion and refer to the general accounting policies disclosures, including e.g., inventory valuation and depreciation methods. In this way the quality and meaning of financial statement elements is enhanced. Similarly, when examining a particular financial statement element with a cross-reference to a footnote, the disclosure would add meaning by putting the element in context, elaborating it or disaggregating it. The lack of interaction between element tags and disclosure tags presents processing problems for applications using the XBRL document. UML provides diagrams that permit the modeling of these natural interactions. These diagrams in turn can provide a schema for augmenting XBRL documents that then can be better processed by downstream users, including intelligent agents, and applications.

Framework

Figure 1 depicts the overall framework for the processing of financial information. Financial documents, such as SEC-mandated 10-K filings, provide the raw material for financial analysis. The order in which financial reports are normally presented differs, but trends can be discerned from "Accounting Trends & Techniques," published annually by the AICPA (e.g., AICPA, 2000), which reports the results of a survey of the accounting aspects of the annual reports of 600 corporations. The major components of the financial reporting system, the order in which they typically appear, and the percentage of the occurrence (if available) according to the AICPA 2000 survey, are shown below:

1. Report of the Independent Accountants (or Audit Report) either follows financial statements and notes (53%), or precedes the financial statements and notes (43%).
2. Balance Sheet
3. Income Statement
4. Statement of Stockholders' Equity (which includes Statement of Comprehensive Income in 82% of the sample cases)
5. Statement of Cash Flows (this is the final statement in 48% of the sample cases, and it flows the Income Statement and Balance Sheet in 47% of the cases)
6. Notes to the Financial Statements, the first of which is normally the Accounting Policies disclosure.

Although there is some structure to the SEC filings, the purpose of XBRL is to enhance these documents by developing tag sets for financial statement elements (Debreceny and Gray, 2001). For example, the following snippet provides a self-describing context for a particular financial statement element, *viz.* inventory:

```
<Asset date=12-31-2002>
        <Current Asset>
                <Inventory>10000</Inventory>
        </Current Asset>
</Asset>
```

This structure enhances the downstream processing of financial statement elements. If a set of companies used the same taxonomy at the same point in time, then cross-sectional analysis is facilitated. Similarly, if the same taxonomy is used over time for the same company, time-series analysis is facilitated. Finally, if both techniques are used panel data statistical analysis can be accomplished.

With respect to additional disclosures, the same inventory financial statement element would be referred to under General Accounting policies. For example:

```
<Note 1>… The company uses the FIFO method on a lower-of-cost-or-market basis for
                valuation of inventories….</Note 1>
```

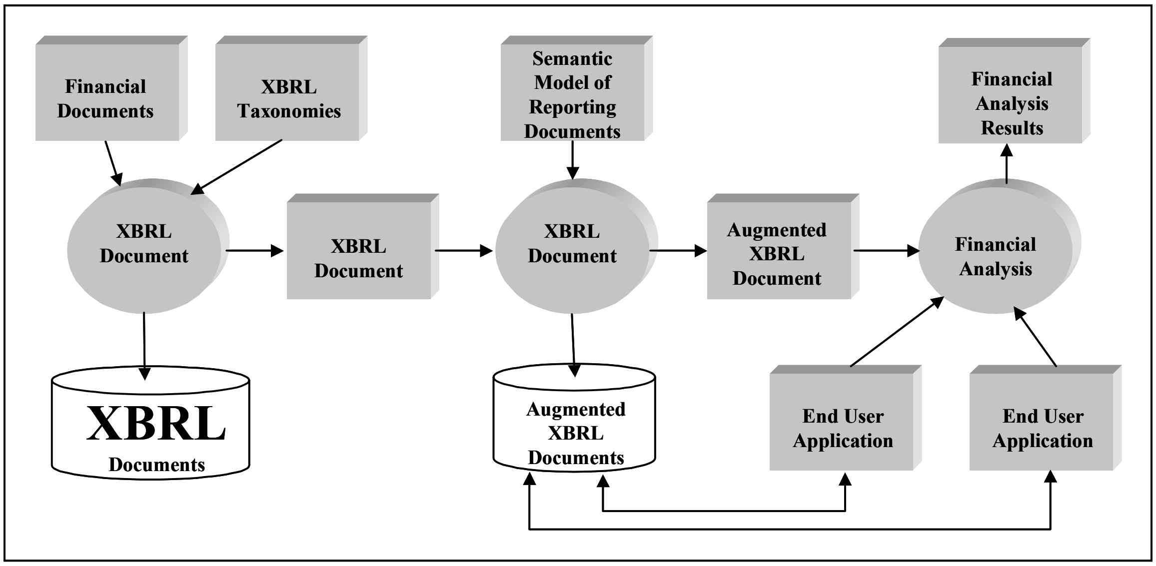

Figure 1. Framework for Financial Processing Using Augmented XBRL Documents

Another disclosure, specifically cross-referenced to the inventory statement element might contain:

<Note 5> The merchandise inventory amount contains $2,000 related to merchandise that had originally cost $3,000.</Note 5>

Thus, the meaning of the statement element is better described by the disclosure, both in terms of context (valuation method) and in terms of disaggregation (the $1,000 markdown to lower-of-cost-or-market**).** Unfortunately, the current XBRL taxonomy does not permit the additional meaning to be systematically garnered. Thus, the value of the XBRL conversion is less than what could be attained with disclosure integration with statement element information.

UML diagrams could better represent the inherent inter-relatedness of these underlying objects. The modeling of these objects and their collaborations would better represent their semantics. A semantic model for reporting documents would enable the augmentation of XBRL documents such that downstream financial processing would be a better value-added proposition. In sum, modeling the more subtle relationships that exist within the financial statement documents and using those models provide a methodology for improving the downstream processing of the document content, whether by analysts, intelligent agents or applications.

UML-Based Modeling of Financial Documents

Unified Modeling Language (UML) is the most commonly used modeling language today. Software engineers routinely use UML in systems development, particularly to specify the structure and behavior of the overall system. The primary benefits of UML are: a) easy transition from design to implementation, and b) UML specification can be executed to check for consistency and accuracy. UML unifies and formalizes many object-oriented approaches advocated by Booch, Rumbaugh, Jacobson and Odell. It provides a rich set of modeling constructs that can be used to capture the intricacies of the requirements of systems as well as the semantics of the application domain.

UML can be used to not only represent the structure and behavior of systems, but also to model various aspects of domain artifacts at a high level (Carlson, 2001). For example, in the financial reporting domain, there are several types of reporting documents that exist with well-defined components and structures. These components are also interrelated, with specific processing implications. It is essential that these relationships are understood and captured to facilitate the automated processing of financial data contained in these documents. UML naturally lends itself in modeling the structural components of financial reporting documents, the relationships between them, and how these components are utilized in financial analysis and how they impact one another in various contexts. Thus, modeling XBRL documents using UML would help enhance the semantic content of the document. This semantic augmentation can be exploited by the applications that use these documents for better financial analysis.

Overall, UML supports the following types of models (Booch et al, 1999): a) use case model – the specification of actions that a system or class can perform by interacting with outside actors, b) static model – describe the structure of components, semantics of data and messages in a conceptual way, c) dynamic model – captures the behavior of components, d) implementation model – describe the component distribution on different platforms, and e) object constraint language – formal language to express additional semantics within UML specification. Each model type supports one more diagrams. Some of these diagrams are listed in Table 1.

Table 1. UML Models and Diagrams

Model Type	Associated Diagrams
Use Case Model	Use case Diagrams
Static Model	Class Diagram Object Diagram Package Diagram
Dynamic Model	Collaboration Diagram Sequence Diagram Activity Diagram Statechart Diagram
Implementation Model	Component Diagram Deployment Diagram
Object Constraint Language	Constraints, Invariants Pre and Post Conditions Navigation Paths

We use a subset of the diagrams shown above in modeling an XBRL financial document. In particular, we use the class diagram and the package diagram to capture the structure of the financial document. The collaboration diagram is used to capture how the individual components of the document interact with each other, thus facilitating the integration of these components. Some of the elements of the financial document are used as input to certain types of financial processing activities, which could be expressed using the activity diagram. Figure 2 shows a high level overview of UML components for SEC filing documents. The interaction between the reporting elements and the disclosure elements is the key to understanding and using these filing documents.

Processing XBRL Documents

Once the SEC filings are converted into XBRL documents, they can be utilized by various external entities that perform different types of financial analyses. The semantics of the XBRL documents improved through UML constructs, results in the effective use of the contents of these documents. In addition to human analysts, intelligent agents can also be employed in the analysis process. In particular, agent-based applications can be developed that can use the enhanced XBRL documents and automate the analysis process for a variety of financial computations. For example, standard ratio analysis may be performed, but on items from companies meeting certain disclosure requirements, e.g. only companies with clean opinions and LIFO inventories. Figure 3 shows a use case diagram that would serve as a starting point for developing software applications that can assist the analyst in processing financial statements.

Summary

This paper has discussed the need for improving the semantics of SEC filing documents and has presented a framework for augmenting the XBRL documents through the use of UML constructs. Such an enhanced document can be processed by not only human agents, but also software applications and intelligent agents. This results in the increased utilization of XBRL documents in financial analysis and decision-making. Future work includes, clearly articulating and modeling the various elements of disclosures and how they relate to financial statement elements, and developing a proof-of-concept prototype application. This framework for using UML to augment XBRL financial documents will provide richer semantics in the financial statement analysis domain. The UML schema could be used to generate code through forward engineering that would further process the XBRL document, making the augmented documents more amenable to further processing by financial analysts, intelligent agents and downstream software applications. Financial analysts could develop screening mechanisms and standard ratio analysis based upon the augmented documents. Intelligent agents could search for companies meeting conditions contained in the augmented documents. Software applications could further process the augmented documents performing standard analyses on a broader set of financial constructs. In general, the integration of the two technologies affords a synergistic method for enhanced financial statement analysis.

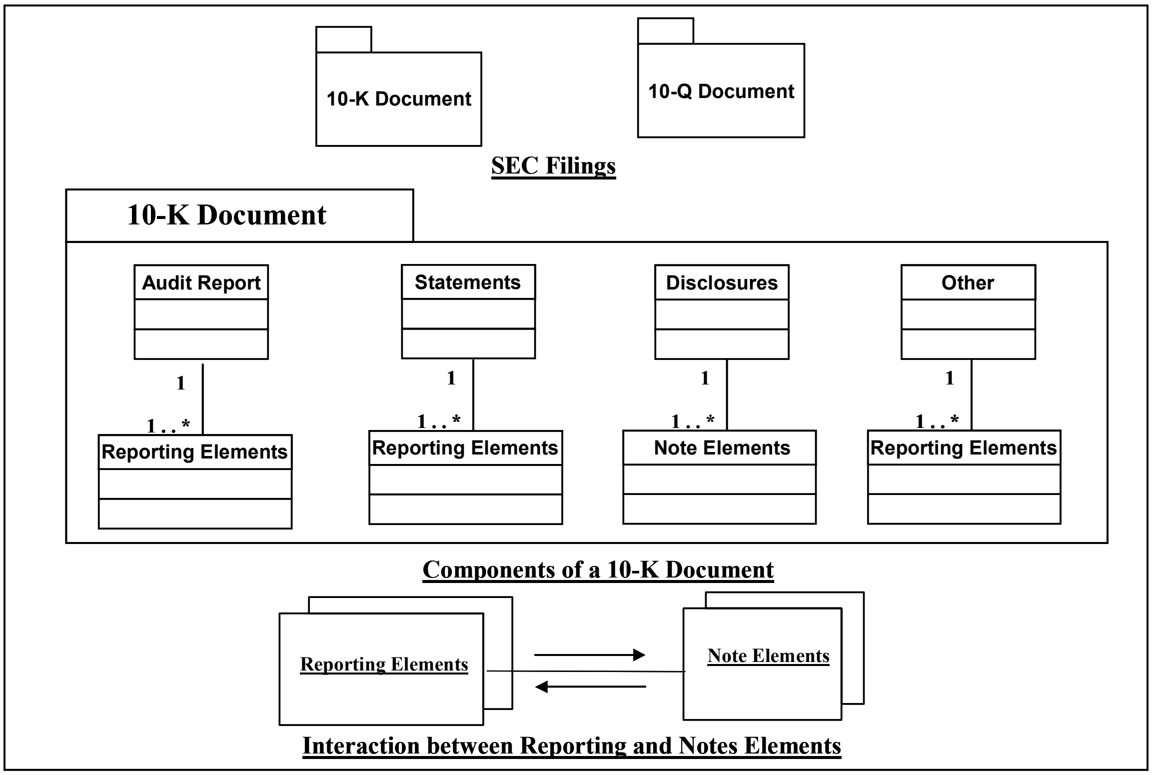

Figure 2. UML Components Representing SEC Filing Documents

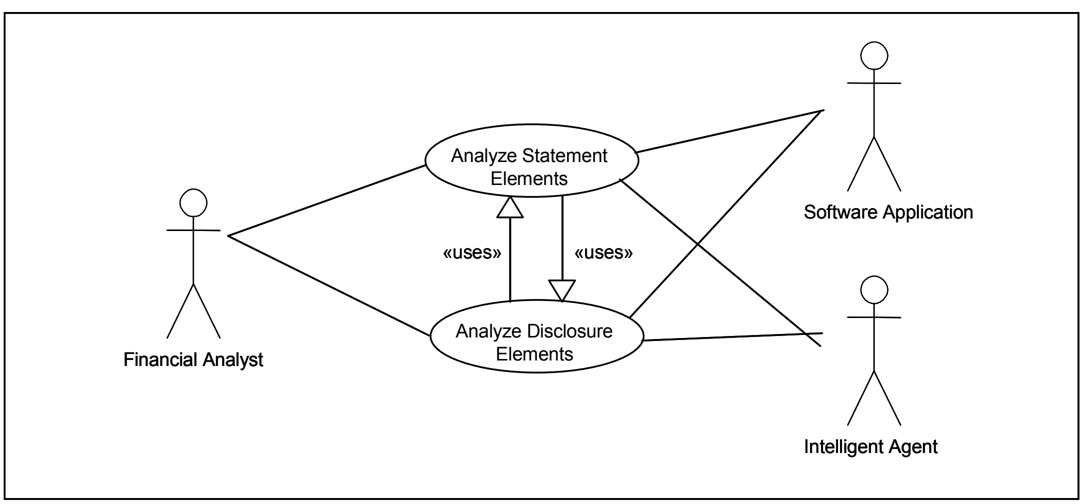

Figure 3. Use Case Diagram for Analyzing SEC Filing Documents

References

AICPA. "Accounting Trends & Techniques" (54^{th} Ed.). New York: AICPA, 2000.

Booch, G., Rumbaugh, J., Jacobson, I. The Unified Modeling Language User Guide, Addison Wesley, New York, 1999.

Bovee, M., Ettredge, M., Srivastava, R.J. & Vasarhelyi, M. "Enabling Semantic Consistency for Electronic Business Reporting: Does the Year 2000 Taxonomy Accommodate Current Business Financial Reporting Practice?" *American Accounting Association Information Systems Mid-Year Conference*, Orlando, FL, 2002.

Carlson, D. Modeling XML Applications with UML, Addison Wesley, New York, 2001.

Coffin, Z. "The Top Ten Effects of XBRL," *Strategic Finance*, June, 2001, pp. 64-67.

Debreceny, R. & Gray, G.L. "The Production and Use of Semantically Rich Accounting Reports on the Internet: XML and XBRL," *International Journal of Accounting Information Systems*, Vol. 2, 2001, pp. 47-74.

Hamscher, W. & Vun Kannon, (Eds.) "Extensible Business Reporting Language (XBRL) Specification," July 31, 2000. http://www.xbrl.org

Strand, C.A., McGuire, B.L., Watson, L.A. & Hoffman, C. "The XBRL Potential," *Strategic Finance*, June, 2001, pp. 58-63.

Watson, L.A., McGuire, B.L. & Cohen, E.E. "Looking At Business Reports Through XBRL-Tinted Glasses." *Strategic Finance*, September, 2000, pp. 40-45.

Zarowin, S. & Harding, W.E. "Finally, Business Talks the Same Language," *Journal of Accountancy*, August, 2000, pp. 24-30.