ASSURING DATA INTEROPERABILITY THROUGH THE USE OF FORMAL MODELS OF VISA PAYMENT MESSAGES

(Category: Practice-Oriented Paper)

Joseph Bugajski
Visa International
JBugajsk@visa.com

Philippe De Smedt
Visa International
pdesmedt@visa.com

Abstract: Visa desires a high-quality payment experience for its cardholders, merchants, and Member banks. The application of formal techniques to model payment messages exchanged among the participants in a transaction has led to a substantial reduction in the incidence of interoperability problems. These occur when message fields that capture information about a transaction are incorrectly populated. Models allow the precise specification and shared understanding of the structure, semantics, and lifecycle of the data, and thereby enhance the overall payment experience. A monitoring environment is in place to analyze all transactions in the global VisaNet system against expected values and combinations of values specified in the models, and to take corrective action when interoperability issues are identified. Various standards activities are underway where models are applied to the creation of new financial messages and the evolution of existing ones.

Key Words: Payment Messages, Interoperability, Data Modeling, Process Modeling, UML Models, International Payment Standards, baseline analysis, statistical analysis

INTRODUCTION

The payment experience is directly related to the quality of the data in the payment messages exchanged over the course of a transaction. As the number of payment products and services offered by Visa increases in response to market opportunities, the added complexity necessitates a novel approach to specifying the structure and meaning of the information associated with the various types of transactions.

The framework that underlies Visa's data quality program [1] has as one of its major components the ability to specify formal and machine-readable models of the messages and the lifecycle of the messages. The models serve a dual purpose: first, they provide a precise and unambiguous mechanism for specifying what the data is supposed to look like, removing the possibility of differences in interpretation by business analysts and software developers of traditional written documents; second, they provide a mechanism for analyzing actual transaction data against what the models prescribe.

Interoperability is defined as the capability of all the participants in a transaction to uniformly apply the complicated rules for populating and reading data elements in messages along the payments chain. Interoperability problems generally occur when population of message fields is done incorrectly anywhere in the transaction. Most often, this is due to misinterpretation of the textual description of the structure, semantics, and use of the impacted data fields. Lack of interoperability manifests itself as data quality problems. These problems may cause transactions to fail, impacting authorization rates, causing difficulty in clearing and settling transactions, and resulting in the expense of researching failed transactions.

To put the challenge in context, the Visa payments chain involves many participants: 20,000 Member banks issue Visa cards or acquire Visa card business; there are 1.59 billion cards in circulation; 59 billion transactions are conducted annually, with peaks of over 6,800 transactions per second; and Visa cards are accepted in 170 countries. The resulting annual transaction volume is USD 4.8 trillion [3].

A large variety of payment products and services is offered by Visa: debit, credit, ATM, mobile, proximity, web-based, 'traditional' card-swiped payments, bill and recurring payments, and others. Each of these results in a slightly different population of the fields that compose the payment messages, i.e., the authorization message and the clearing and settlement message. Furthermore, as the message moves from one participant to the next, its field values may change.

With a multi-billion dollar global investment in payments infrastructure in place at merchants, Member banks, and Visa itself, it is only natural to leverage that investment by reusing or overloading existing message formats rather than to impose costly changes caused by new message formats every time a new product or service is introduced.

To avoid any ambiguity when overloading existing formats, formal message modeling technology is introduced. The Unified Modeling Language (UML) [4] specified by the Object Management Group (OMG) is the most widely used and implemented modeling standard. Models are precise and unambiguous. They can be shared by all participants along the payments chain and can be processed by any tool that is compliant with the specification. In addition, the OMG's Model-Driven Architecture (MDA) is the cornerstone for specifying model-based software systems and, eventually, for generating software directly from the models (also known as Model-Driven Development (MDD)). With these properties, UML models are the ideal vehicle for formally specifying payment messages, rendering the payment software development process much more efficient, eliminating the possibility of misinterpretation of written format specifications, and facilitating the introduction of new payments products and services.

The financial services industry as a whole has recognized the need to approach the evolution of financial messages in support of new products and services through the use of models. Global, industry-wide standardization efforts underway include those being conducted under the International Standards Organization's (ISO) Technical Committee (TC) 68 (Financial Services), which is developing the ISO 20022 (UNIFI) standard [5], and the OMG's Finance Domain Task Force (FDTF). These activities are further elaborated upon in this paper.

Modeling for the purpose of addressing information quality is a novel approach. It allows for both the specification of expected results in an objective, unambiguous, and machine-readable fashion, and for the validation of actual results against expected results, in a unified manner. It thus supports the end-to-end capability of the Visa data quality analysis framework, capable of handling very complex interoperability problems. The work reported in this paper has been conducted by the authors and their colleagues at Visa.

Furthermore, the modeling approach dovetails very nicely with the multi-dimensional model of data quality described in [2], in that it fundamentally supports notions such as completeness (models can capture any data elements desired), interpretability (definitions are clear), correctness (result values can be validated against expected values), understandability (models are precise, unambiguous, and easily comprehended), relevancy (only data fields applicable to specific transaction types need to be modeled), etc. The application of models is expected to enhance the quality of financial messaging in these areas.

THE PAYMENT MESSAGES MODELING APPROACH

Application of the UML to solving the interoperability challenge enables the capturing of the structure and semantics of the data, the specification of process and lifecycle models, the ability to trace data to the process steps in which it is manipulated, the exchange of message formats with transaction partners, the definition of next-generation financial messages, and the gradual transition of existing message formats to next-generation ones. Each of these points is discussed in more detail below. Before doing so, a graphic showing the flow of a Visa authorization message is introduced.

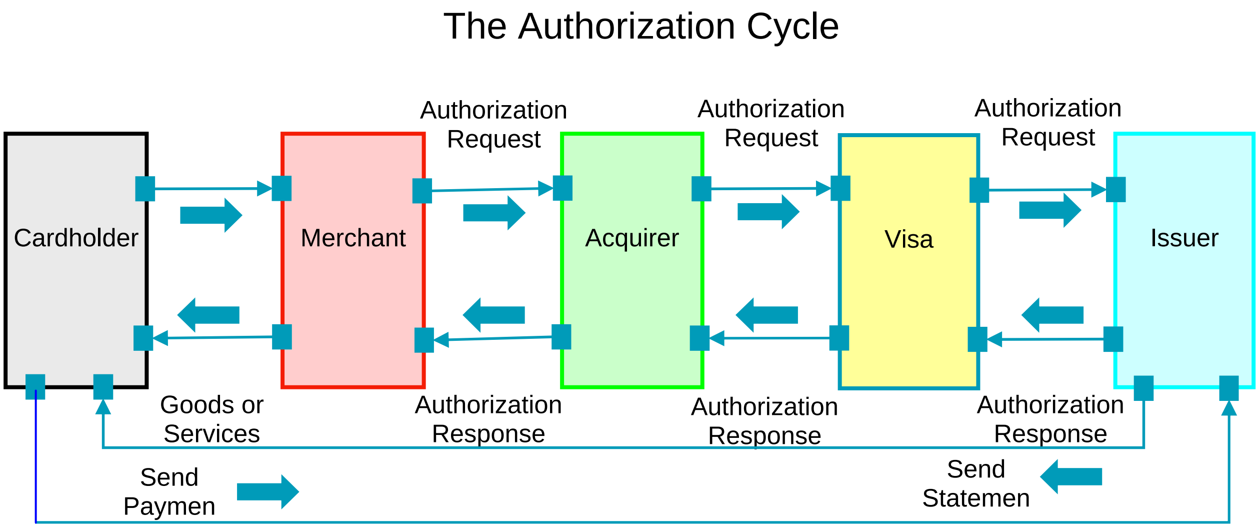

Figure 1. The Authorization Cycle

Figure 1 shows the key participants in a typical card payment scenario: the cardholder, the merchant, the acquirer (i.e., the merchant's bank), Visa, and the issuer (i.e., the cardholder's bank). It also shows the exchanges of goods and services, and of information about the transaction. Message models show precisely the structure and semantics of the data, as well as how data is generated, manipulated and consumed at each step in the authorization process. Similar models for clearing and settlement messages can be constructed.

A Visa authorization message is based on the ISO 8583 standard [6], and is structured as shown in figure 2.

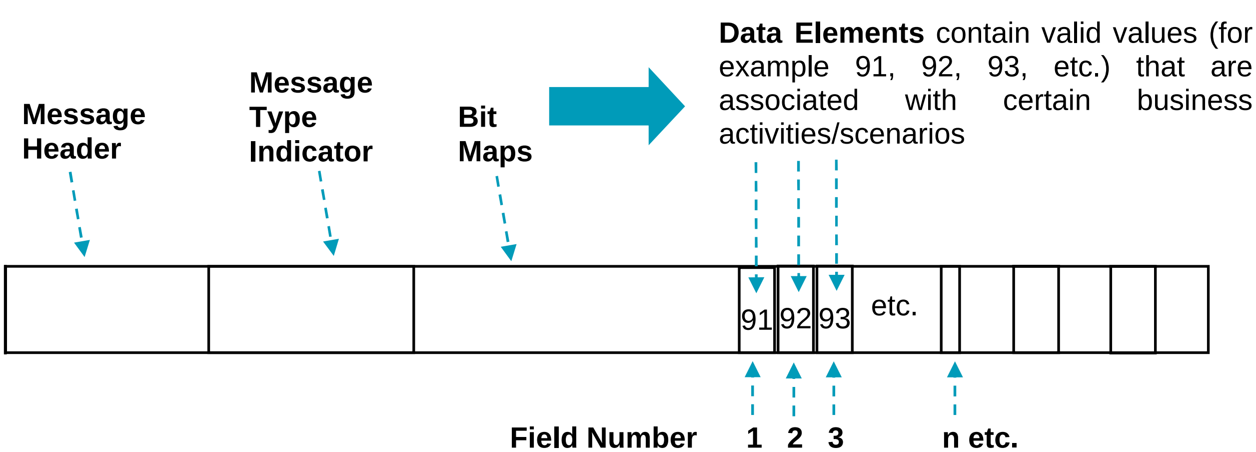

Figure 2. Structure of an authorization message

The authorization message consists of the following parts: message header, message type indicator, and bitmaps indicating which data elements are present. The data elements are the fields that carry the information about the transaction itself, e.g. account number, merchant identification, amount, etc.

Capturing the structure and semantics of the data in payment messages

While ISO 8583, designed originally in support of 'traditional' credit card transactions, provides a powerful means for capturing payment transaction data, it is becoming increasingly difficult to adapt to new payment products and services which require more complex information to be captured about the transaction's characteristics. ISO 8583 is a complex document susceptible to ambiguity and with the potential for misinterpretation. That said, over the 20 years of its existence, it has served the payments industry very well. To continue this trend and take full advantage of the capabilities and bandwidth that the format provides, a UML-based modeling approach is applied.

For each type of transaction (e.g. card present, recurring transaction, etc.) and possible scenario for such a transaction (e.g. card swiped vs. manually entered), a UML class diagram is constructed that maps the scenario to the set of relevant fields and their values (or combination of values). A simplified, generic example is shown in figure 3, for a transaction type 'TransactionTypeABC' and scenario 'Scenario9'.

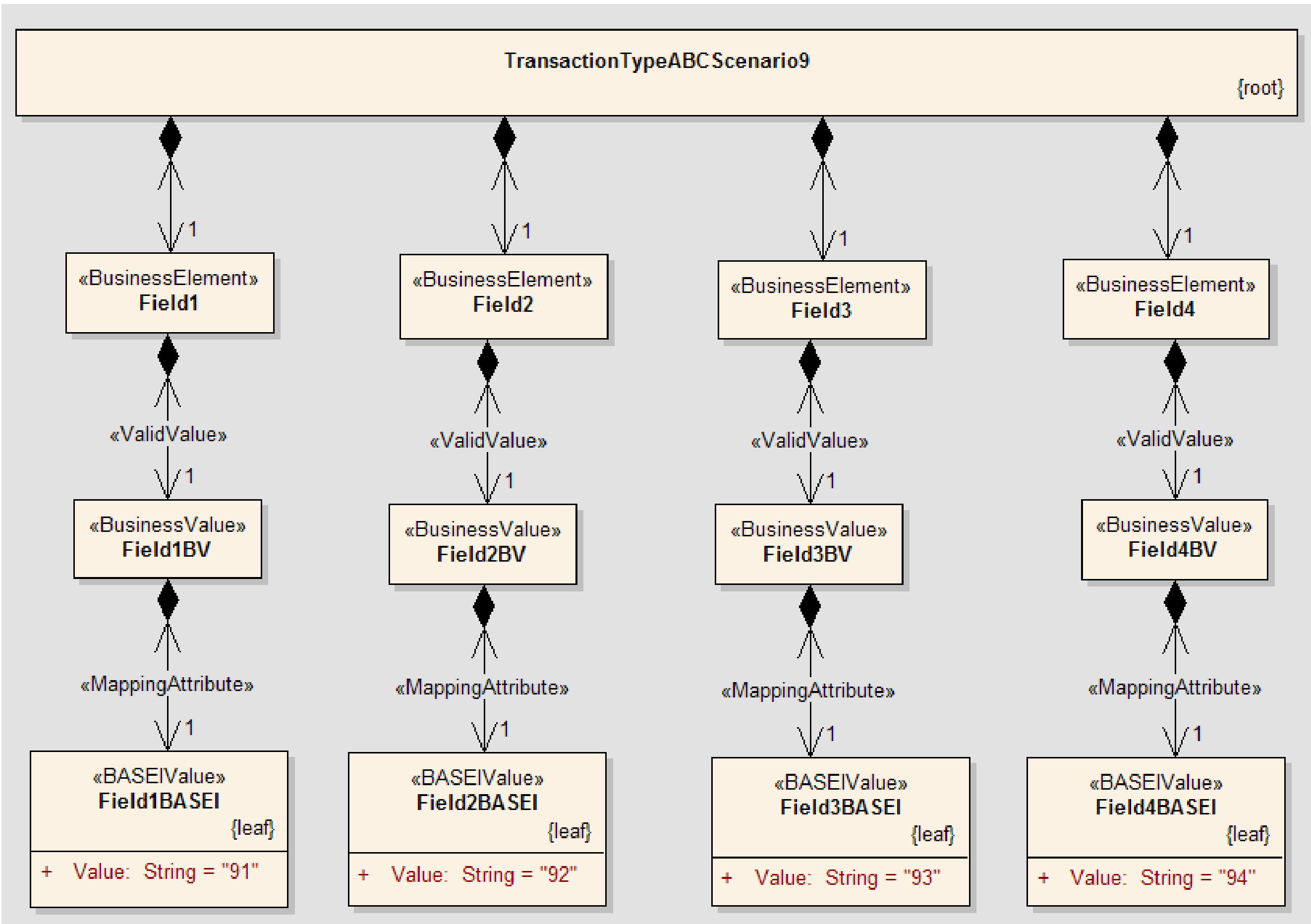

Figure 3. A UML class diagram showing the data fields associated with a specific payments scenario

In this diagram, leaf nodes represent physical information about each field in the message: its valid value(s), including combinations of valid values for multiple fields, and structural information (location, length, data type, etc.). Non-leaf nodes represent how the leaf nodes map to business concepts, allowing modeling at the business level. The root node represents how the business concepts are mapped to the particular use of the message (i.e., the combination of transaction type and scenario).

Specifying process and lifecycle of the data in payment messages

It is absolutely essential to also understand how a payment message moves along the payments chain, and the corresponding manipulation of the fields in the message from generation to consumption, and all steps in between. An example is the process of passing the authorization request to Visa by the acquirer, passing it on to the issuer by Visa, and the return of the authorization response from the issuer to acquirer (and then to the merchant) via Visa. Figure 4 shows a UML activity diagram for this process.

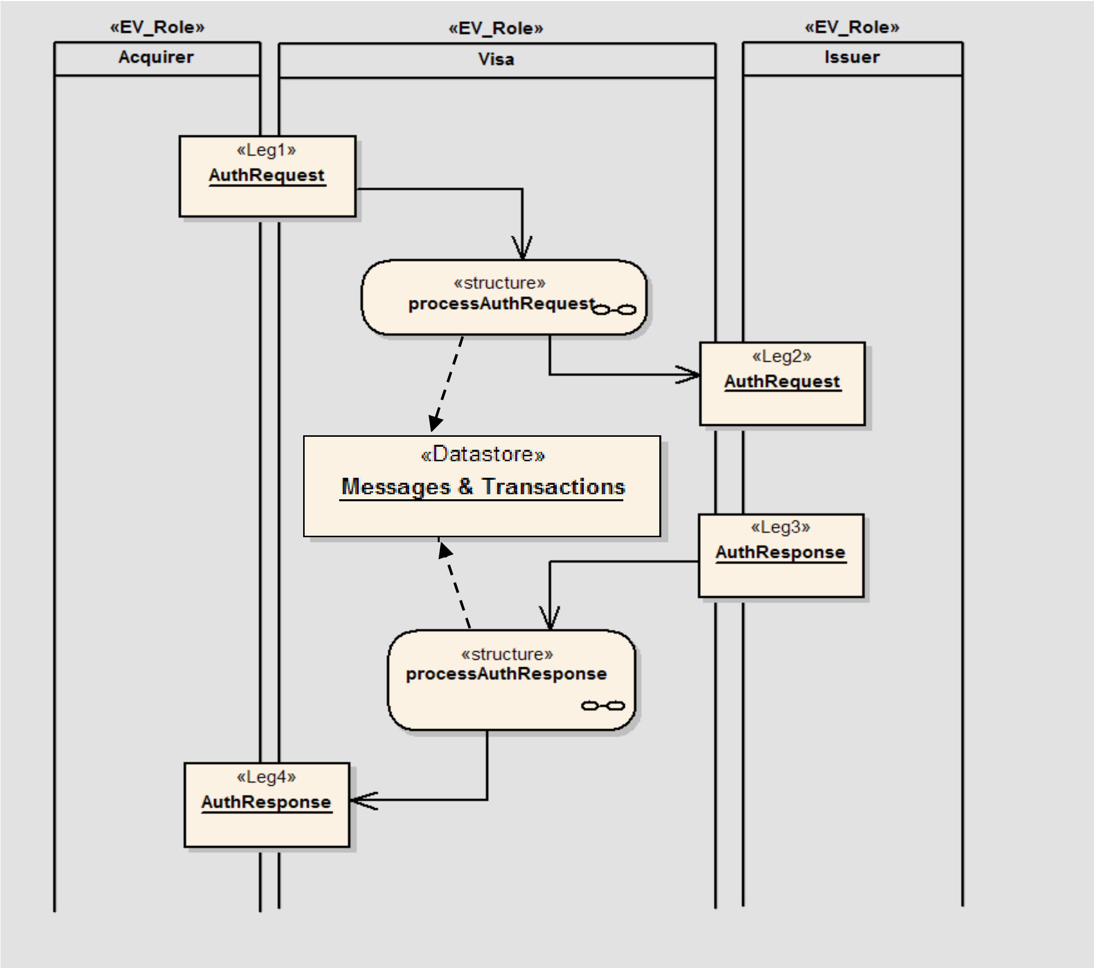

Figure 4. A UML activity diagram of the authorization request and response flow

To further understand where interoperability problems may arise, figure 5 shows a UML activity diagram of the population of an authorization request for a recurring transaction at the merchant (e.g. a utility). In this diagram, a process is depicted for checking whether an account is due for payment and composing the authorization request from a number of data items. Each of these can be traced back to the non-leaf nodes in the class diagram for the recurring payments scenario.

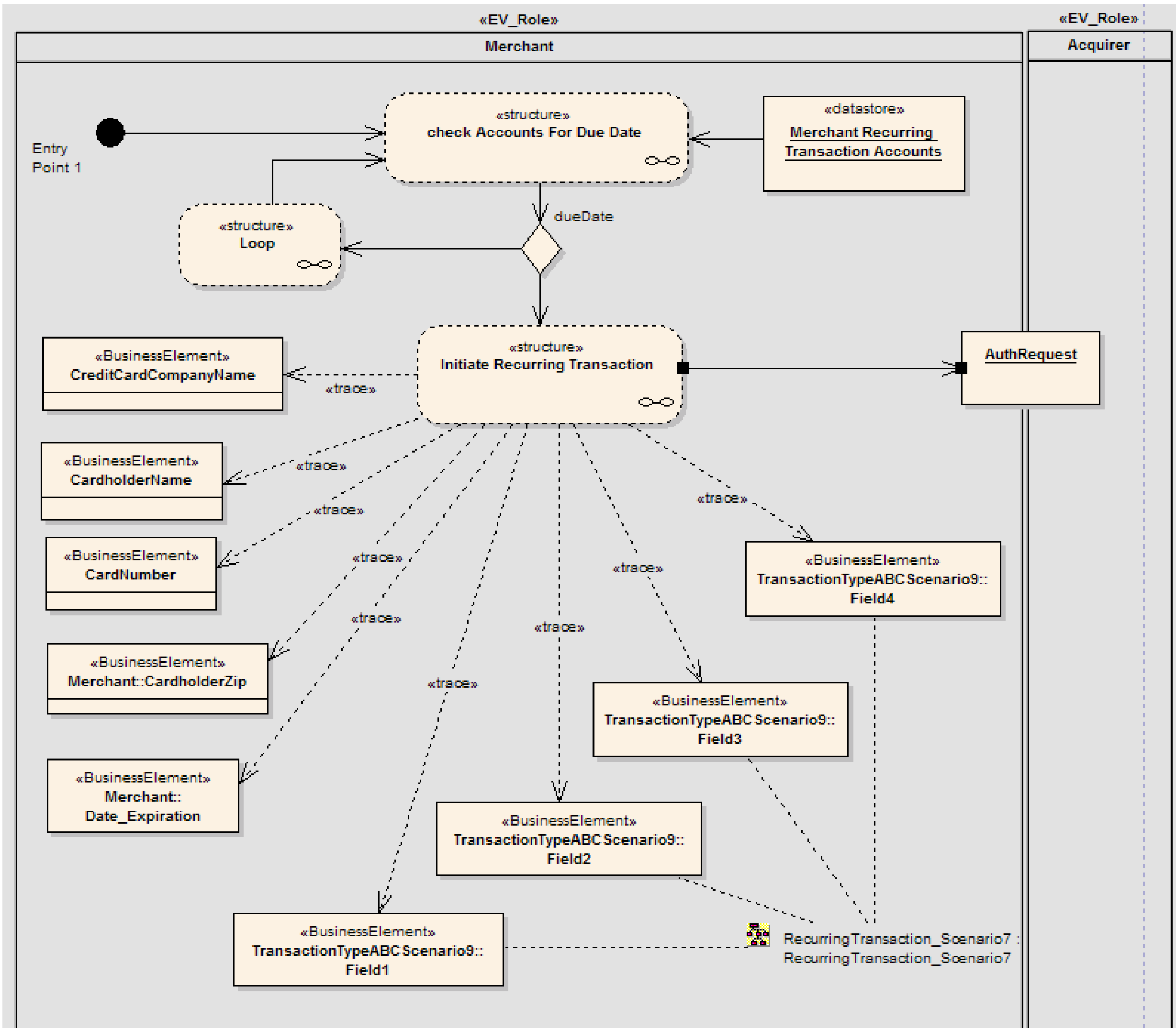

Figure 5. Population of an authorization message in a recurring payments scenario

Tracing data to process steps

Traceability of data to process steps facilitates the understanding of how data is manipulated along the payments chain and assists in pinpointing the source of interoperability problems.

In figure 5, for instance, it is clearly shown how the values submitted in the authorization request are derived from those specified in the class diagram in figure 3.

Exchanging models with other participants in the payments chain

A powerful feature of UML class models is that they can be serialized into a format that is standardized across all tools that faithfully implement the UML specification. The serialization process is known as XMI (XML Metadata Interchange) [7], an OMG specification, and converts class diagrams into XML documents. These can then be exchanged with other participants in the payment process. The reason for a tool-neutral exchange format is that it cannot be assumed that all participants use the same UML tool in their IT organization.

As indicated earlier, once the models have been communicated by Visa, they can be used in a variety of ways, for instance as a documentation aid augmenting the written specifications or, ideally, as computationally consumable artifacts integrated into the development process. The extent to which models are applied is dependent on the maturity level of the organization.

Defining next-generation financial messages

The global financial services industry has identified a need to create a standard approach, a 'recipe', for the creation of new messages and the evolution of existing ones. As the industry's pre-eminent standards body, the ISO, through its Technical Committee (TC) 68 defines the ISO 20022 (UNIFI) standard. The intent of this standard is to eliminate or vastly reduce the potential ambiguity that is inherent in standards defined in 'written' form, rather than supported by a formal modeling methodology, such as UML. It is expected that the quality of financial messages created in compliance with UNIFI will be an improvement over existing ones.

The standard, a new version of which is currently under development, provides the following: a UML-based meta-model for the definition of financial messages and processes, a mechanism for precisely defining the content and structure of a message, a repository structure to which UNIFI-compliant message components can be contributed, and a mapping from the models to an XML format (at present, a proprietary mapping is specified, but application of XMI is under consideration; in addition, wire formats other than XML can be specified as well, but they are not currently part of the standard).

By applying a modeling formalism to the creation of new messages and by defining an infrastructure to share the message definitions, UNIFI will go a long way toward fostering financial message standards that are precise and can be evolved to support new financial products and services. UNIFI is not restricted to the payments space, but also covers securities, foreign exchange, trade services, and others.

Transitioning from existing formats to next-generation message formats

It was mentioned earlier that very substantial investments were made in support of the global financial services, including payments, industry. While new approaches and standards are very appealing, one must be sensitive to the fact that existing formats will continue to exist for some time at some organizations, while others will want to adopt the new formats sooner, because of all the benefits that they provide.

Two approaches address this issue. The first approach is the definition of two compliance levels within UNIFI: the first one is compliance at the model level; the second one is compliance at the wire level (currently XML). By not making wire level compliance a requirement, organizations with existing or legacy messages can continue to use them in their operations, but be in compliance with a formal of the messages. This is an approach that is very suitable to an existing and widely used format such as ISO 8583. By adopting this approach, organizations can gradually move toward the adoption of new message formats.

The second approach involves the application of conversion mapping technology. The OMG's Finance Domain Task Force is completing a submission to an RFP that calls for model-based conversion maps between message formats [8]. By following a formal methodology for dissecting an incoming message into its elemental semantic components by abstracting its syntax, then mapping these components to common semantic elements maintained in a repository (such as a UNIFI-compliant one), and then mapping from these common elements to the semantic elements of the outgoing message, applying the syntax of the outgoing message, one can be assured that proper, lossless mappings between messages in differing syntaxes can be achieved.

APPLYING MODELS TO THE DETECTION OF INTEROPERABILITY PROBLEMS

The previous sections have explained the usefulness of UML models in gaining a precise and unambiguous understanding of the data in payment messages, and have shown the mechanics of modeling.

Visa has taken the additional step of applying models to the detection of interoperability problems. A sophisticated monitoring system analyzes all transactions in the VisaNet system for the existence of such problems. Baseline analyses provide the insights into which transactions should be further analyzed. Powerful analysis and statistical tools are available to Visa analysts [9] to investigate such transactions. This set of tools has been augmented by a model-driven filtering system that checks field values for each transaction against those specified in the message models. Figure 6 depicts the mechanism (note that the left side of the picture contains the diagram of Figure 3).

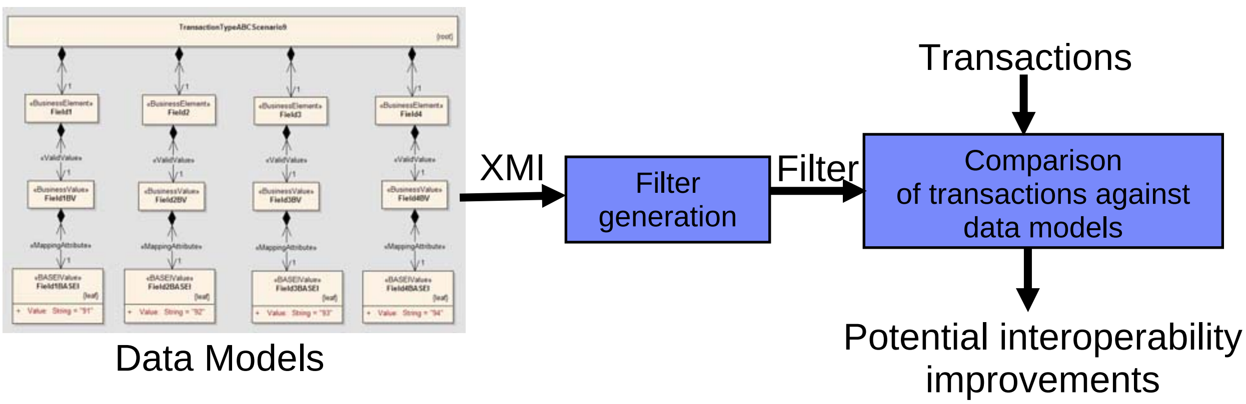

Figure 6. Applying UML models to the detection of interoperability problems

The data models, captured as UML class diagrams, are converted to XML documents using the XMI transformation mechanism. Once they are in XML format, and thus machine readable, filters are programmatically constructed by extracting the information in the leaf nodes of the class diagrams. These filters are then applied to all incoming transactions and used to check compliance of the field values in a transaction against those specified in the models. The outcome of this activity is a clear picture of how well transactions stack up against their corresponding models. Visa is using this approach to identifying potential interoperability improvements to great benefit.

CONCLUSION

UML modeling has proven to be a very successful approach at Visa for identifying opportunities for interoperability improvements. By formally specifying what payment messages should look like and how the information in payment messages is created, manipulated, and consumed along the payments chain, the quality of the information in the messages is greatly enhanced, resulting in an enhanced payment experience for Visa's customers. A model-based monitoring environment, implementing an end-to-end framework for data quality analysis, allows Visa to check for the occurrence of interoperability problems and take remedial action once such a problem is identified. The global financial services community is working hard to drive the adoption of modeling standards for financial messages in various domains.

REFERENCES

[1] Joseph Bugajski and Robert L. Grossman, An Alert Management Approach to Data Quality: Lessons Learned from the Visa Data Authority Program, submitted for publication

[2] Leo L. Pipino, Yang W. Lee and Richard Y. Wang, Data Quality Assessment, *Communications of the ACM*, *Volume 45*, 2002, pages 211-218

[3] Visa Operating Certificates statistics, as reported by member financial institutions globally (31 March 2007), and therefore subject to change

[4] The Unified Modeling Language (UML), Version 2.1.1, Object Management Group (OMG)

[5] ISO 20022 UNIversal Financial Industry message scheme (UNIFI), www.iso20022.org

[6] ISO 8583 Standard for Financial Transaction Card Originated Messages, 1987 version, 1993 version, 2003 version

[7] The XML Metadata Interchange Standard, Version 2.1, Object Management Group (OMG)

[8] Conversion Models for Payment Message Standards, OMG RFP, doc.omg.org/dtc/2005-9-13

[9] Joseph Bugajski, Robert L. Grossman and Eric Summer, Steve Vejcik, Monitoring Data Quality for Very High Volume Transaction Systems, *Proceedings of the 11th International Conference on Information Quality (ICIQ)*, 2006