

Context Mediation on Wall Street

Allen Moulton

Stuart E. Madnick

Michael D. Siegel

*MIT Sloan School of Management
Cambridge, Massachusetts 02139
{amoulton, smadnick, msiegel}@mit.edu*

Abstract

This paper reports on a practical implementation of a context mediator for the fixed income securities industry. We describe industry circumstances and the data and calculation services (DCS) mediator developed and deployed in the early 1990s. The mediator was designed as an interpretive engine controlled by a static declarative knowledge structure and client preference data. In addition to heterogeneous, autonomous data sources, the mediator integrated autonomously developed local and remote procedural components. Client access to both data and computational resources were provided through an active conceptual model. Structural and semantic context conversions were used to integrate disparate components and to support varying client needs. Lessons learned from the implementation and usage of this mediator provide insight into the requirements for a successful context mediator.

1. Introduction.

The fixed income securities industry is an environment with potential for major benefits from rapid and effective interchange of semantically heterogeneous information among autonomous entities. Many firms in the industry develop specialized, and often proprietary, knowledge. That knowledge is more valuable when used in combination with knowledge from other specialists. Although standardization can facilitate knowledge interchange, the rapid evolution of financial products, transactions, and markets characteristic of the fixed income securities industry suggests that a dynamic context mediator approach may have significant value.

The challenge of the Data and Calculation Services (DCS) mediator product described here was to bring together six highly competitive firms to cooperate in providing decision support information to their customers while retaining autonomous control of proprietary data and analytic resources. Wiederhold [1][2] described mediators as an extra software layer to provide client

applications with access to the integrated knowledge available from heterogeneous, autonomous data servers. The DCS mediator extended the definition to include computational resources as well as data sources.

Context refers to the implicit assumptions that add meaning to symbols. Autonomous organizations often develop different contexts, resulting in semantically heterogeneous representations of the same knowledge[3]. In making investment decisions, it is vital the information conveyed among parties be understood by all. In the fixed income industry, regulators, industry associations, and publications, such as Fabozzi's series of books (e.g., [4]) disseminate common theories and practices. Context mediation can fill the gap where industry association and regulatory standards leave off.

In COIN [5][6][7], source and receiver contexts are independently described to the mediator, which uses metadata to detect conflicts in data representation and interject conversion operations in the query plan. The DCS mediator supported context definition and automatic conversions for parameters and values of computational functions as well as source and receiver data. In addition, the DCS mediator experimented with context conversions among different representations of security valuation related through abstract models, often implemented in proprietary analytic software.

The DCS mediator employed a domain model of information applicable to fixed income decision-making. SIMS [8], Information Manifold [9], Infomaster [10], and Garlic [11] also employ a domain-specific ontology. In DCS and these other projects, data sources are integrated through a shared model of knowledge specific to the problem domain. Users access information through the mediator by working directly with the domain model. DCS differs from the approach of TSIMMIS [12], which uses an ontology-free approach where data is integrated and delivered in its own terms, from Carnot [13][14] which draws on an ontology of general common-sense knowledge, and from Benaroch's [15] use of macro-level ontological knowledge to integrate disparate knowledge-based systems for financial risk management.

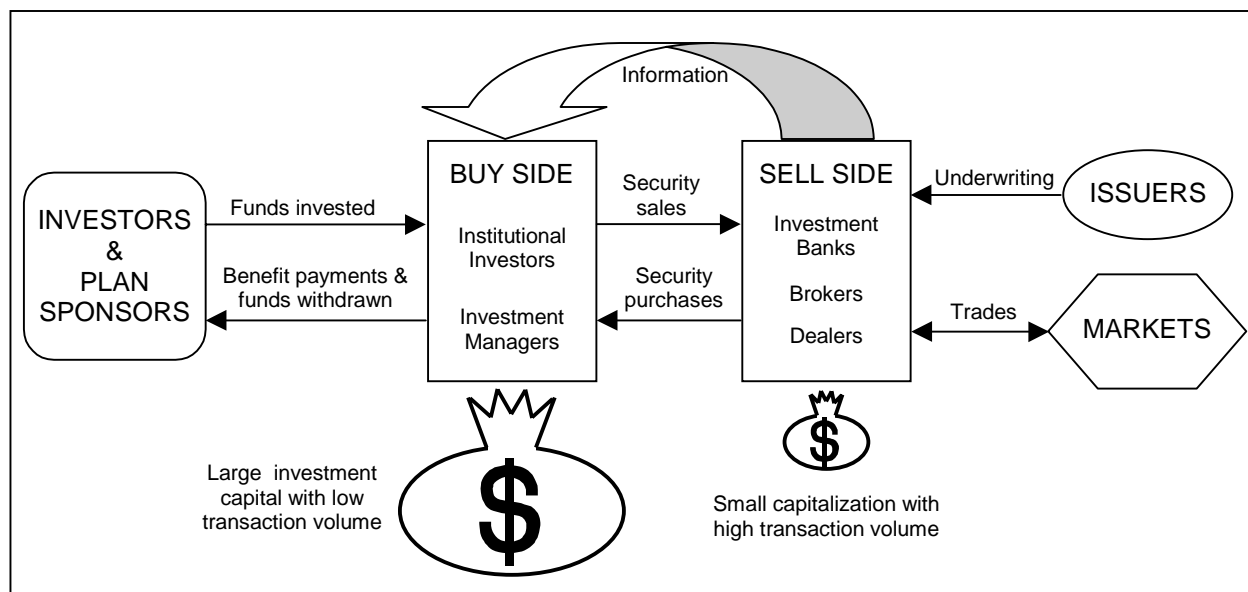


Figure 1. The fixed income securities industry

2. The fixed income securities industry.

Information is the critical resource in the fixed income securities industry – information about securities and their issuers, information about markets, information about economic conditions and events, and information about methodologies and models (see Figure 1). Billions of dollars of debt instruments trade every week. Firms on the “buy side” (institutional investors and investment managers) manage capital on behalf of investors and benefit plan sponsors. Firms on the “sell side” (investment banks, brokers, and dealers, often known as “Wall Street”) bring new securities to market and interact to create capital markets. While the popular connotation of sell-side firms is of great wealth, the capitalization of these firms is actually relatively low. The big money is on the buy-side. The sell-side makes money from a small commission or price spread on large volumes of transactions. The sell-side offers buy-side managers access to its skill, expertise, and knowledge in anticipation of purchase and sale orders.

Bonds and other fixed income securities are obligations to pay sums of money at points in time over the life of the security. Unlike equity securities, an investor has no stake in the financial entity that issued the security. To an investor, a fixed income security represents a stream of future cash flows. A purchase decision trades present money for future payments. There may be optional events that change the cash flow stream and there may be risk of default. In essence, however, all fixed income securities are interchangeable commodities from the point of view of an investor. The cash flows

from one or more obligations may even be repackaged by selling off rights to payments or by combining rights to payments into new composite securities. This repackaging, or “financial engineering” can produce securities known as “derivatives.” Faced with a vast array of combinations of cash flows, risks, and optional events, every industry participant needs timely information and effective methods for determining investment value from raw data.

In 1990, six major Wall Street investment banking firms formed a partnership establishing an “Electronic Joint Venture” aimed at using advanced technology to improve communication with their buy-side clients. During the 1970s and 1980s, fixed income investment management had become substantially more complex. Interest rates had fluctuated over a wide range. Huge quantities of new debt instruments had been issued. New securities products had been regularly introduced. Analytic methodologies had advanced to cope with the new complexities. Markets had become more globally integrated and moved at a faster pace.

To cope with the new complexity, sell-side firms hired bright young “rocket scientists” with mathematical and engineering backgrounds into “fixed income research” departments. They developed databases with detailed descriptive information about securities as well as historical prices, interest rates, and economic statistics. Proprietary real-time broadcast networks brought current market information and news to trader and investment analyst desktops. Application software was built to display securities data, perform securities and portfolio analytics, and apply the models developed by financial engineers. Information technology had also been

advancing rapidly. Business mainframes stored databases, supercomputers performed the most complex calculations on an overnight basis, while a combination of dumb terminals, PCs, and workstations were piled high on trading and sales desks.

In the midst of this jumble of technology, buy-side firms began to demand direct access to data, analytics, and models. Previously, the sell-side would offer ideas and supporting analysis on the phone and by FAX. The buy-side wanted more control over the analysis used to make investment decisions. Investors recognized that the sell-side had the knowledge that was needed. Investors wanted to be able to integrate proprietary sell-side data, analytics, and models with their own internal data, analytics, and models so that they could make better decisions. The sell-side firms were willing to provide information resources to the buy-side, but wanted to protect proprietary information from being passed on to competitors. The sell-side also hoped to receive information from the buy-side about positions held in their portfolios in order to be able to make better trading recommendations.

Investment management involves buying, selling, and swapping securities in order to improve the performance of an investment portfolio while meeting the sponsor's objectives. In fixed income, quantities of securities are generally measured in face value. Prices, interest, yields and many other values are usually expressed as a percentage of face value. The price (or valuation) of a security depends on the general interest rate market and the security's predicted cash flow compared to other investments on the market. A variety of methodologies, such as internal rate of return (IRR), horizon rate of return (ROR), and option-adjusted spread (OAS), are regularly used to assist in investment decision-making. An investment manager needs to be able to apply the same methodology to all securities held in a portfolio and all potential investments. A requirement for an effective decision support tool is the ability to apply a common methodology across diverse securities. An effective tool must also be extensible to new kinds of securities invented by financial engineers and give investment decision-makers access to and control over assumptions and models used in valuation methodologies.

The mission of the Venture was to provide the infrastructure to facilitate the interchange of information between buy-side and sell-side firms. The technology plan involved six major areas: physical connectivity via a proprietary network, a desktop workstation platform, a suite of standard applications, standard databases, standard analytic calculations, and the data and calculation services (DCS) mediator. By offering "generic" or industry standard applications, databases, and

analytic calculation software, the venture would be able to replace duplicative efforts on the part of clients and its partner firms. All Venture-supplied applications, databases, and calculations were required to be designed on an open architecture that would allow partner firms and clients to develop their own proprietary extensions and substitute parts. For example, a partner firm might provide a proprietary application that would use and extend the Venture's standard databases and calculation software. A partner firm might alternatively supply a proprietary algorithm for use within a standard application.

3. Architecture of the DCS mediator.

The DCS mediator (see Figure 2) was designed as a demand-driven general interpretive engine controlled by a static declarative knowledge. The declarative knowledge structure included a domain model for fixed income securities and analytics, database retrieval and calculation function wrapper specifications, and a dependency graph. The declarative knowledge was prepared in a definition language script compiled into a binary structure for run-time interpretation. Application software accessed the mediator through a C-language application-programming interface (API). Ad-hoc user access was provided through a spreadsheet interface. Generic wrapper interfaces were developed for the Sybase relational DBMS, C function libraries, an object technology contributed by one of the partner firms, and remote procedure calls (RPC).

Client preferences and assumptions were contained in global data objects that could be viewed and manipulated through the Control Panel application. By changing values in these global objects, the user could set the financial assumptions, data sources, and analytic models used by the mediator in responding to other application queries. In essence, the Control Panel allowed the user (and the client firm's management) to define the context, the implicit assumptions, of information to be received.

3.1. A fixed income securities domain model

The core of the domain model was a DCS conceptual model, which defined an abstract view of the entities, attributes and relationships available through the mediator and served as a shared ontology for integrating heterogeneous, autonomous data and calculation resources. Publication of the conceptual model served as the principal documentation of the data and analytic functionality available from the mediator. Each entity class defined a collection of attributes accessible through "get" and "put" operations provided in the API. No object methods or parameterized procedures were allowed.

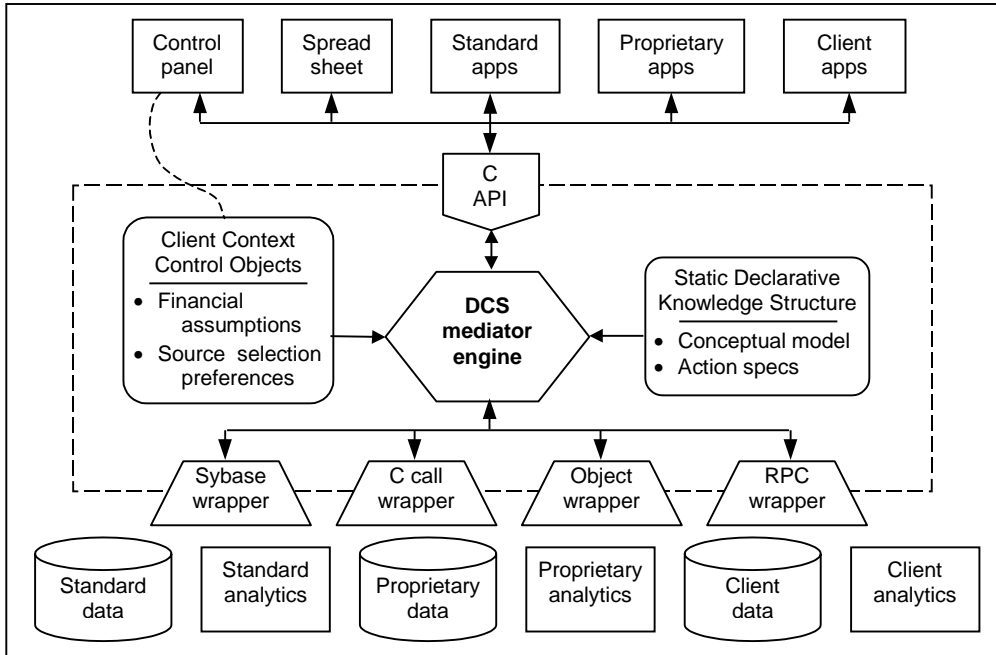


Figure 2. DCS mediator architecture

Application and spreadsheet developers worked directly with conceptual model objects to access DCS resources. The user asked for information from the DCS mediator. The mediator decided whether that information would be calculated or obtained from a data source.

Figure 3 shows some of the abstract classes for in the conceptual model for fixed income securities and analytic methodologies. In the center is the *Security* class, representing the generic fixed income security with a many-to-one relationship to the *Issuer* class with information about the issuing financial entity. There is also a one-to-many relationship to *Valuation*, which contains information about end of day prices from various sources. The *PyCalc* (price-yield calculator) class represents information about an actual or hypothetical trade in a security and analytic measures describing the trade. The *Assumptions* class holds user and client financial assumptions, and control information about data sources, and analytic models for use by the mediator in responding to application queries.

The DCS domain model supported single inheritance and dynamic polymorphic subclassing. Figure 4 shows a

sample of the polymorphic subclasses of the *Security* and *PyCalc* abstract classes. In each case, the subclass is determined by the value of a meta-attribute in the parent class. *Security* subclasses correspond to different kinds of descriptive data relevant to each type of security: *SecBond* for bonds, *SecPool* for mortgage pools, *SecMGen* for a generic mortgage model, and *SecTranche* for a tranche of a collateralized mortgage obligation (CMO). For the *SecTranche* subclass, the security is related to a *CMODeal* entity and in turn to a set of *Collateral*, each of which may be a position in a pool or a generic mortgage.

Subclasses of *PyCalc* correspond to alternative procedures for performing the functionality of the abstract class depending on the type of security involved. In some cases, a *PyCalc* subclass might also provide extended functionality specific to that type of security. Figure 4 shows subclasses *PyBond* for bonds, *PyMort* for mortgage generics and pools, and *PyTranche* for CMO tranches. The subclass of *PyCalc* is derived from the type of *Security* entity referenced by the "sec" attribute in the calculator. Dotted lines indicate the association of *PyCalc* subclasses with *Security* subclasses.

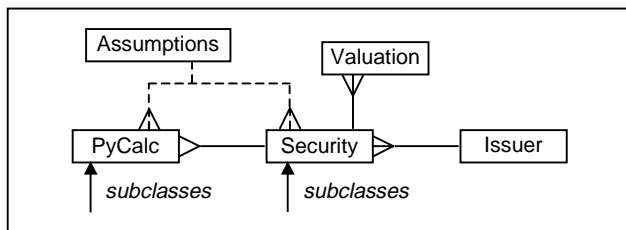


Figure 3. Abstract classes and relationships

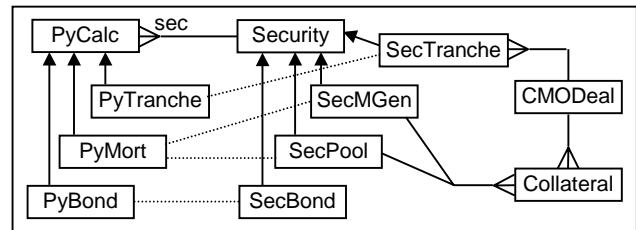


Figure 4. Polymorphic class relationships

Enumerated metavalues provided a standard terminology for mediating among different representations of security and calculation features. For example, the description of coupon interest payment frequency and interest accrual convention for a security used metavalues. Similarly, the yield calculation conventions in the *Security* and *PyCalc* classes used metavalues. The context definition for a data source or computational function library included relations to map context-specific representations to the DCS domain model standard. The mediator used these mapping relations to convert representation across contexts.

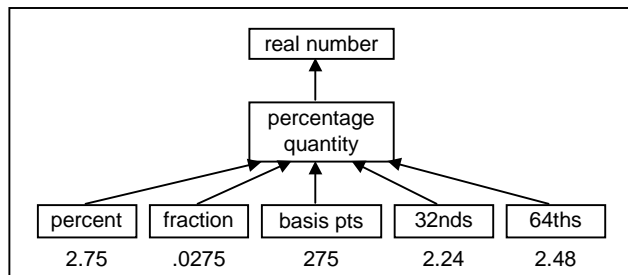


Figure 5. Representations of percentages

DCS context conversion also helped alleviate one of the pesky problems of fixed income software development by providing automatic conversion among different representations of percentage quantities. Securities are priced at a percent of face value, interest rates and yields are percentages, as are spreads between two prices or yields, the interest accrued on a unit investment as of some point in time, and many other quantities. Figure 5 shows five alternative representations of the same 2.75%. The "fraction" form is favored by spreadsheets and many programmers. Basis points (hundredths of a percent) are handy for talking about small spreads. The "32nds" and "64ths" forms are used for bond markets that trade in those increments. The DCS domain model provided types for alternative representations of percentages and automatic conversion among them.

Figure 6 shows a sample of the "placemat" depiction of the generic *Security* class, showing some attributes and data types. The *type* attribute in the lower right provides an explicit representation of the type of the security. Type may be implied (because a source is limited to a single type) or derived from stored information. The *maturity* and *coupon* boxes inside the diagram are internal structures describing the maturity and the coupon interest payments of the security. Attributes declared as *percent* could be filled with any form of percentage quantity. Many-to-one relationships are shown as role attributes, such as *issuer*, referring to a single object. One-to-many relationships appear as role attributes, such as *valuation*, referring to a set of objects.

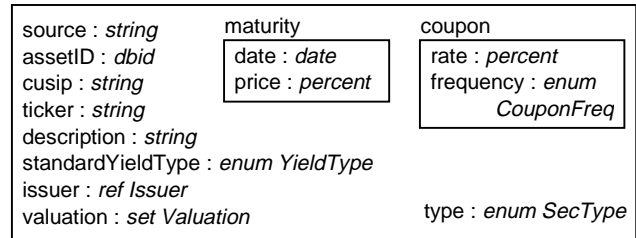


Figure 6. Generic Security class placemat

An object of the *Security* class could be used as a hypothetical security, with attribute values filled in by the application program. More often, a *Security* object would be associated with stored data using key attributes such as *cusip*, *ticker*, or *assetID*. CUSIP is an industry standard identifier used in the United States and Canada [16]. An *assetID* was an artificially constructed globally unique key for all security data objects accessible through DCS. Venture-supplied databases were designed with *assetID*. Autonomous databases are made accessible by extracting the primary keys of each security record into a lookup table and assigning a unique *assetID* to each. The *assetID* also contained codes indicating the source for the data and the security type.

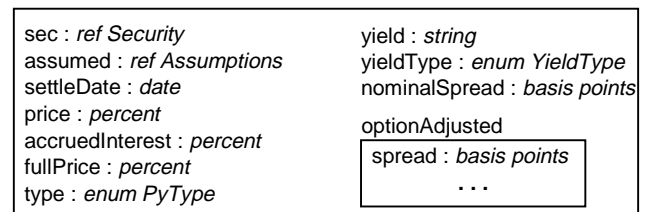


Figure 7. Generic PyCalc class placemat

The *PyCalc* class (Figure 7) provided both IRR and OAS analytics for a real or hypothetical trade in a security. The *sec* attribute provides a reference to a security object. The *assumed* attribute connects to an *Assumptions* object with data about user preferences and financial assumptions. The *type* attribute is derived from the security's type and controls subclassing. Subclasses of *PyCalc* primarily represent different methods of computing standard analytic results based on the class of security involved.

One of the requirements of the DCS mediator was delivery of a common analytical framework across all kinds of securities. Many of these calculations are specified in detail by industry conventions or regulatory agencies. Since the characteristics of securities vary widely, the parameters and procedures for calculations also vary widely. In some cases, security cash flow must be projected using a predictive model and financial assumptions. For example, an attribute of the *PyMort* subclass controlled the choice of prepayment model used in projecting cash flow for *SecMGen* and *SecPool* securities. If no value was set by the application, the value

for that attribute was obtained from the user's setting in the *Assumptions* object, allowing the prepayment model to be varied without the application's knowledge.

The calculation of price from yield, and yield from price, are so common and important that regulators and industry standards bodies have set down precise rules for computations and precision of results. For bonds, bills, strips, and other straightforward securities, there are closed form analytic formulas for computing present value. Mortgages and mortgage-backed securities require prediction of the cash flow stream using a prepayment model followed by a general discounted cash flow method. Calculating yield from price requires a root-finding algorithm for most securities. Industry standards specify the root finding algorithm and the initial and terminating conditions. As some programmers have learned the hard way, a "more accurate" result is worthless if it differs from what everyone else in the industry expects.

3.2. Integrating databases.

DCS did not include a full query-writing capability for databases. Instead, it relied on pre-coded SQL statements provided in database "retrieval packages." Each package was associated with a class in the conceptual model. Attribute references were specified using path names relative to that base class (e.g., "coupon.rate" based on Security). Each package included a list of attribute paths for required parameters, which could be associated with parameter slots in the SQL statement. Each element in the SELECT clause of the SQL statement, mapped to a corresponding attribute path where retrieved data would be stored. The mediator would not execute the SQL in a package until all required parameter values were present.

3.3. Integrating computational functions

Computational functions and simple arithmetic expressions could be associated with attributes of classes in the domain model. When the value of an attribute was requested, the mediator would first check to see if a value had previously been set or derived. If not, the expression or function would be evaluated and the result value assigned to the requested attribute. Inputs to the calculation were specified as attribute paths based on the containing object. Intermediary results were held in "value packages" along with metadata describing the DCS domain model type. In preparing a computational function call, the mediator consulted the function declaration and performed necessary data conversions. Independently supplied functions did not need to be changed when

integrated into DCS, since the mediator would automatically handle conversion of input data from its original context to the context of the called function. The value returned was held in a value package until the context of its destination was determined.

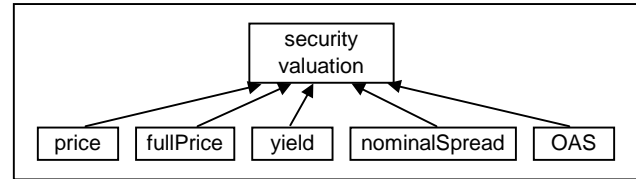


Figure 8. Representations of valuation

DCS required the domain model designer or the function integrator to specify each calculation alternative for mutually dependent attributes. The five attributes of *PyCalc* shown in Figure 8 are financially equivalent representations of security valuation. For example, *fullPrice* is equal to the sum of *price* and *accruedInterest*. Given any two values, the third can be calculated. To represent this situation, all three expressions are declared. Given a request for *price*, for example, the mediator would select the expression subtracting *accruedInterest* from *fullPrice*, proceed to find either constraints or other expressions to determine values for these two attributes, then take the difference. The mediator tracked attributes visited to detect cycles and return an error.

3.4. The DCS C and spreadsheet APIs

The DCS API supplied four basic operations: create or delete object and get or put attribute value. The create operation (*dcsNew*) returned a handle to a runtime object of an entity class in the Conceptual Model. Get and put operations (*dcsGet* and *dcsPut*) accessed the value of an attribute referenced by a dotted path name relative to an object handle. The C API provided variants of the get and put operations for common C data types (e.g., *dcsGetInt*, *dcsGetString*), as well as specially defined types, such as "percent." Where possible and necessary, data values were automatically converted between the C data type used in the application and the representation in a source.

Figure 9 shows a fragment of code using the C API. Programs would perform a series of put operations to initialize the runtime object with parameters or search criteria. Get operations would then be used to obtain the results of calculations or data retrieval. For the spreadsheet interface, the sequence of put operations was replaced by a range of name-value pairs passed to an object declaration function which returned a spreadsheet handle to a DCS object. Individual get functions were used in each cell where an output value was required.

```

hCalc = dcsNew( "PyCalc" );
dcsPutString( "sec.cusip", hCalc, "887315AT" );
pDescr = dcsGetString( "sec.description", hCalc );
dcsPutBasisPoints( "nominalSpread", hCalc, 275 );
yield = dcsGetPercent( "yield", hCalc );
price = dcsGetFraction( "price", hCalc );
dcsRelease( hCalc );

```

Figure 9. API example

4. The DCS mediator in use.

After about a year of design and development, a beta version of the DCS mediator was made available to application developers in 1991. Product release was in mid-1992. After the initial development of the mediator engine, the efforts of the DCS development team focused primarily on extending the domain model, verifying the correctness of calculations and data integrated through the mediator, and documentation. Documentation included "placemat" diagrams of object classes similar to the small samples in Figures 6 and 7, along with explanations of the semantics of each attribute. A glossary of fixed income terminology (still available on the web at [17]) clarified definitions and semantics as used in the domain model.

The DCS mediator eliminated the need for application developers to understand the individual semantics of each database and calculation program library. The procedure that would have been required to integrate the separate components was part of the mediator declarative knowledge rather than each application program. Application developers were freed to focus on the business problem, the user's needs, and user interface design relying on the mediator to find and integrate the necessary data and calculation components.

As new types of securities, new calculation libraries, new analytic methodologies, or new data sources became available, additional classes and subclasses could be added to extend the DCS domain model without requiring changes in application software. Using the Control Panel application, the end user could swap calculation methods and data sources used by other running applications.

The DCS mediator offered the appearance of a large database capable of providing integrated answers to a broad range of application queries. A few small examples may be helpful in understanding how DCS worked. The code in Figure 9 represents a small fragment of the sort of query that might be used in a "portfolio review" type of application, which scans through a list of securities displaying some descriptive information along with several valuation measures for each. A portfolio review application must be able to retrieve indicative data from an appropriate database and apply appropriate calculation

procedures for each kind of security found in the portfolio. Restated in an SQL-like query language, the DCS request in Figure 9 might look like:

```

SELECT sec.description, yield, price
FROM PyCalc
WHERE sec.cusip='887315AT'
AND nominalSpread=275

```

4.1. Retrieving security indicative data

Much of the information about a security is stored data that can be obtained from a database directly. For example, the first part of the sample query above asks for a text string (sec.description) describing the security whose CUSIP number is given (sec.cusip). For CUSIP number 887315AT, the string might be:

TIME WARNER 7.48% DEB 01/15/2008

Some securities databases may store a description string; others may not, since the description is redundant with other information stored for the security. The description begins with the name of the issuer (Time Warner), which in a normalized database would appear in a related Issuer table. The other elements of the string are the coupon interest rate (7.48%), the type of security (debenture), and the maturity date (01/15/2008).

Without DCS, the application developer would have to determine whether the database contained a descriptive text field. If the field were missing, the developer would have to retrieve the other four attributes and write a program to combine them into a descriptive string.

Using DCS, the application developer can request the security's description and leave the details to the mediator. If the database stores the description, DCS acts as name-mapping view. If the database does not contain a suitable description attribute, the mediator uses the declarative knowledge structure to find a "calculation" function to produce a description from other available attributes.

4.2. Working with valuation measures

The five measures of security valuation shown in Figure 8 are alternative representations used to express how much a security is worth as of a given settlement date. For analytic purposes, the settlement date is the date when a hypothetical sale of the security would take place. Price is the percentage of face value to be paid for the principal of the security. Full price adds the amount of interest accrued from the last payment through the settlement date. For analytic models, full price can also be considered the present value of a future cash flow stream. Yield is the parameter used for discounting future cash flows in the internal rate of return valuation model.

Nominal spread measures the difference between the yield on the security and the yield for a benchmark security of similar maturity. The benchmark yield may be taken from an actual security or from a point on a yield curve. The notion of spread is that one security may be valued relative to another or relative to an approximation of the market captured in a yield curve. The most common benchmark is the U.S. Treasury yield curve, which is taken as a proxy for the yield curve for risk free investments.

The final measure, option adjusted spread (OAS), measures the expected value based on the probability distribution of future interest rates. OAS is also a relative value measure based on a time-varying risk free investment return rather than the single yield used in nominal spread. For complex securities with cash flows that depend on future interest rates, the calculation of OAS can be complicated and compute-intensive.

The sample query illustrates how the DCS mediator assists the application developer for valuation measures. The query asks for yield and price given nominal spread. To calculate yield, a benchmark yield is needed. Since the benchmark is not provided in the query, the mediator consults the financial context to find which yield curve or security to use. The value returned to the application will depend on the context setting controlled by the user.

Once the yield has been found, the mediator must then apply the appropriate yield-to-price calculation algorithm for the security involved. In addition, the mediator retrieves the additional security descriptive data needed for the calculation. Both the algorithm and the data required vary with the security. The final ingredient for calculating price is a settlement date, which was not specified in the query. The financial context can provide a default settlement date for a series of queries.

While valuation measures are generally thought of as calculations, many systems pre-calculate and store values that are frequently needed. For example, an overnight pricing program may calculate yield, nominal spread, OAS, and other measures for all securities in a database based on the closing price for the day. When a portfolio manager asks for a morning review, the data can be retrieved from the database instead of calculated. By using the mediator, the decision to calculate or retrieve is transparent to the application.

The security valuation measures also illustrate the contextual basis of representational equivalence. Price, nominal spread, and OAS are equivalent measure of value given a nominal or option-adjusted abstract model of relative security value based on another abstract model of return on investment. Nevertheless, specifying one or another valuation measure may have different operational implications. For example, specifying the price may mean

that the expected selling price of a particular security is known. Alternatively, by specifying nominal spread, the price is left to fluctuate with changes in the market or changes in user assumptions, both of which are represented by the benchmark yield curve. The DCS domain model allowed the application developer and the user to control and work with valuation models as circumstances required.

5. Conclusions.

DCS proved the concept of context mediation in the pragmatic world of an information-intensive and highly competitive industry. When the original design was presented to clients, response was enthusiastic. Developers liked the idea of integrating data and calculations, allowing them to avoid the work of connecting components together in software. They found the notion of treating computational results as data to be appealing.

Producing a mediator for production business use required a major focus on quality assurance and training. The major quality assurance effort was not directed at the mediator itself. Clients saw the results delivered through the mediator as the responsibility of the mediator. In consequence, the mediator development team spent most of a year testing and fixing calculation software as it was integrated. Considerable effort was also devoted to documentation and training, including a manual on domain model design using a tennis shoe sales example.

The DCS context mediator was similar in some ways to an object broker approach. The difference was a design based on declarative knowledge used to dynamically decide on methods to materialize elements of the conceptual model.

For some users, the structure imposed by the standard conceptual model was helpful. More advanced developers and client shops, however, wanted to be able to specify their needs as user views or in their own conceptual models. The mediator would need to go beyond context conversion for individual pieces of data, to be able to map the central domain model into the structure of client views and models. Clients also wanted to be able to treat the components of the domain model as building blocks that could be arranged to suit their purposes. For example, clients would have liked to pose queries that combined entities as needed rather than using pre-planned relationships. A query language like SQL or OQL would have been helpful in place of the cumbersome API.

The DCS mediator offered a new paradigm for financial application development. Traditionally, the application programmer retrieves data and sets up calls to calculation libraries. With DCS, the application

programmer works at a higher level of abstraction, in terms of the results of applying analytic methodologies. The mediator takes responsibility for finding and meshing the right programs and data to fulfill the application's needs.

By decoupling applications from the methods and data used in computing analytic information, DCS allows plug 'n play dynamic substitution of functionally equivalent components. For example, mortgage prepayment models are critical to projecting performance on portfolios holding derivative securities. When a Wall Street firm refines its prepayment model, it wants to make that new knowledge available to its customers quickly. Buy-side portfolio managers may have preferred portfolio analysis application tools that can gain immediate benefit from the new model delivered through the mediator. The mediator can also facilitate the sell-side analyst in developing proposed trades by bringing buy-side portfolio holdings data into the analyst's applications. Improving information flow between buy-side and sell-side can also improve capital market efficiency.

Experience with DCS suggests the potential value of further research on query processing reasoning algorithms for mediating among application developers' conceptual models and component supplier models using industry-specific abstract knowledge.

References

- [1] G. Wiederhold, "Mediators in the architecture of future information systems," *IEEE Computer*, March 1992, pp. 38-49.
- [2] G. Wiederhold, "Mediation in information systems," *ACM Comp. Surveys* vol. 27, no.2, June 1995, pp. 265-267.
- [3] S. Madnick, "Are we moving toward an information superhighway or a Tower of Babel? The challenge of large-scale semantic heterogeneity," *Proc. IEEE International Conference on Data Engineering*, 1996, pp. 2-8.
- [4] F. Fabozzi, ed., *The Handbook of Fixed Income Securities*, 3rd ed., Irwin, Homewood, Illinois, 1991.
- [5] C. Goh, "Representing and reasoning about semantic conflicts in heterogeneous information systems," Ph.D. Thesis, MIT Sloan School of Management, 1996.
- [6] C. Goh, S. Bressan, S. Madnick, and M. Siegel, "Context interchange: representing and reasoning about data semantics in heterogeneous systems," MIT Sloan School of Management Working Paper #3928, Dec. 1996.
- [7] M. Siegel and S. Madnick, "A metadata approach to resolving semantic conflicts," *Proc. 17th International Conference on Very Large Data Bases*, 1991, pp. 133-145.
- [8] Y. Arens and C. Knobloch "Planning and reformulating queries for semantically-modeled multidatabase," *Proc. of the Intl. Conf. on Information and Knowledge Management*, 1992.
- [9] A. Levy, A. Rajaraman, and J. Ordille, "Query answering algorithms for information agents," *Proc. 13th National Conf. on Artificial Intelligence*, Aug. 1996, pp. 40-47.
- [10] M. Genesereth, A. Keller, O. Duschka, "Infomaster: an information integration system," *Proc. 1997 ACM SIGMOD Conference*, May 1997, pp. 539-542.
- [11] Y. Papakonstantinou, A. Gupta, and L. Haas "Capabilities-based query rewriting in mediator systems," *Proc. 4th Intl. Conf. on Paralled and Distributed Information Systems*, 1996.
- [12] H. Garcia-Molina "The TSIMMIS approach to mediation: data models and languages," *Proc. Conf. on Next Generation Information Technologies and Systems*, 1995.
- [13] D. Woelk, P. Cannata, M. Huhns, W. Shen, and C. Tomlinson, "Using Carnot for enterprise information integration", *Second International Conf. on Parallel and Distributed Information Systems*, Jan. 1993, pp. 133-136.
- [14] D. Woelk, W. Shen, M. Huhns and P. Cannata, "Model driven enterprise information management in Carnot", in Charles J. Petrie Jr., ed., *Enterprise Integration Modeling: Proc. of the First International Conference*, MIT Press, Cambridge, MA, 1992.
- [15] M. Benaroch, "Toward the notion of a knowledge repository for financial risk management," *IEEE Trans. on Knowledge and Data Engineering*, vol. 9, no. 1, Jan. 1997, pp. 161-167.
- [16] Standard & Poors CUSIP Service Bureau <http://www.cusip.com/>
- [17] Bridge Info. Systems, "EJV Partners Financial Glossary" <http://bondchannel.bridge.com/html/EJVGlossary.html>