DATABASE RESEARCH GROUP, INC.

DATA BASE newsletter

A FORUM FOR INFORMATION RESOURCE PLANNING, DEVELOPMENT AND ADMINISTRATION

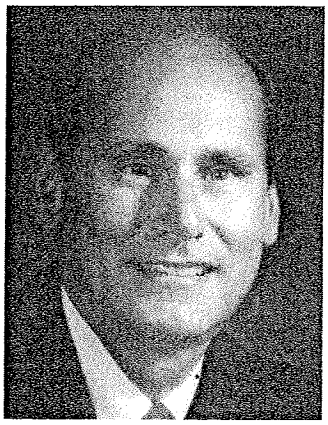
RONALD G. ROSS
EDITOR/PUBLISHER

VOLUME 18, NUMBER 2
MARCH/APRIL 1990

Interview with

George Schussel

President
Digital Consulting, Inc.



George Schussel

The information systems industry has recently seen significant changes in the areas of distributed processing, database, and CASE technology. Cooperative processing and client/server architectures are presenting new opportunities and challenges for system development and information planning. The CASE market has also experienced a major shift as a result of IBM's announcement of AD/Cycle, a framework for integrated CASE technology.

In the following interview, George Schussel, President of Digital Consulting, Inc., discusses the relationship between cooperative processing and client/server architecture, the impact client/server architecture may have on the distributed database market, and changes in the CASE market resulting from the AD/Cycle announcement.

Newsletter: Interest in cooperative processing and client/server architecture has increased significantly over the past few years. What role does distributed database play in this arena?

Continued on page 14

"Three years ago when the move toward distributed database began, it was not really clear how useful the technology would be. Now many companies are discovering that distributed database is essential to implementing certain of their applications."

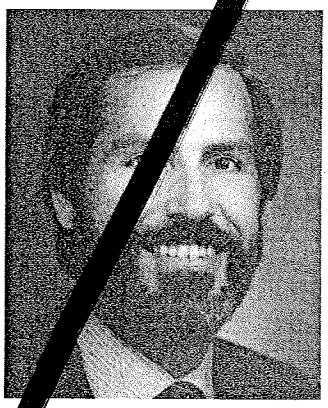
INSIDE THIS ISSUE . . .

The IBM Repository

Pages 2 and 3

Object Orientation and its Software Implementation

By James J. Odell



James J. Odell

Object orientation (OO) provides a unique approach to constructing data processing systems. The OO "paradigm" suggests that first the things (or objects) in the system should be examined—and then their behavior. For many in DP this requires a 180° shift. Most professionals have been taught to define a "problem domain" in terms of processing requirements and of those data items needed-and-yielded "in the process." Because of this, the world is viewed as consisting of processes that manipulate things. OO is different. In OO objects are viewed as being in charge of their own processing destiny.

Process orientation has been eloquently described by Dan Ingalls of Smalltalk fame as responsible for "bit-grinding processors raping and pillaging data structures." For him, object orientation provides a solution that leads to "a universe

Continued on page 5

Object Database

By Mary E.S. Loomis

Hewlett Packard

Object technology is in its infancy, even though its basic concepts have been around for many years. The technology has received much attention from computer vendors, academicians, industry press, and potential users. Yet there remains much confusion about exactly what object technology is and how it can best be applied. This article attempts to eliminate that confusion.

What is an object? An object is like a software building block. It is a data structure that is encapsulated with code that understands the data structure and provides desired ser-

Continued on page 10

ucts. Formal and informal groups are now actively addressing these issues. A reasonable balance must be reached carefully between standardization that enables interoperability and communication, and standardization that stifles innovation early in the lifespan of this relatively new technology.

Support. In general, object database is not a technology that can be thrown successfully over the wall to customers. A few elite groups will be able to evaluate the products and perhaps prototype applications, but the user community at large will not know how to reap the benefits of object technology just from reading an object DBMS reference or concepts manual. It will be very important that the vendors be able to provide their customers with sufficient support and project-related services to ensure successful application of the object DBMS products.

Preparing for Object Technology

Potential users of object database can do several things now to prepare for availability of object technology.

Awareness and education. The availability of information about this area is increasing rapidly. Several publicly offered seminars on object technology are now being offered. Most of these emphasize a particular object programming language. A few focus on a particular object DBMS product or object analysis methodology. Tutorial information is also available in the growing number of books on the subject. The introductory chapters of various object programming language texts can be quite helpful in explaining the fundamental concepts of objects even if the professional does not want to tackle the details of a particular language. Expect numerous new books on object methodologies to appear on bookstore shelves in the next year. Articles about object technology are also beginning to appear in leading-edge database and software periodicals.

Evaluation and hands-on experience. After learning about basic object concepts, getting real hands-on experience with an object product may be beneficial. This might mean prototyping an application using Smalltalk, or building a prototype with an evaluation copy of an object DBMS product.

Methodology. Perhaps the biggest step toward successful use of object technology is to move toward adopting an object analysis and design methodology. This tends to be easier if approached from a data-centered rather than process-centered philosophy. Data modeling should therefore be incorporated into analysis and design work. Use data dictionaries to inventory and control data recourses. Plan for evolution of legacy systems. Identify those parts that are candidates for re-implementation in the next several years. Investigate applicability of the available object analysis and design professional services of a given environment.

Conclusion

The major benefit of object technology is generally expected to be greater productivity over the lifetime of information systems. This productivity arises from the following.

- Models that better capture the characteristics of real-world systems.
- Software that is more adaptable and flexible to accommodate changing business requirements.
- Software modules that are more easily re-used across applications.
- Modes of interaction that are more natural for people.

Object database is just one of the technologies that will be instrumental in achieving the potential benefits of object orientation. As far as the end-user is concerned, object database will be an invisible part of the computing environment. One of the benefits to the business is that database will become more invisible to the application developer as well. ■

SCHUSSEL from page 1

Schussel: Distributed database provides the highest level of system services required to support cooperative processing. In a cooperative processing environment, program functions that traditionally ran exclusively on the mainframe or minicomputer are executed across diverse processors, each program function running on the processor to which it is best suited. For example, editing and screen handling may be performed on local PCs because of their graphics capabilities, response times, and relative economy when compared with the mainframe.

With distributed DBMS, placing applications on diverse processors becomes quite simple. Applications are developed using a common logical data model and database interface (e.g., SQL), and then distributed to different processors using database administration utilities. The distribution of the underlying physical databases is transparent to both the programmer and the end-user.

Distributed database is the technology that simultaneously enables the user to gain the advantages of dedicated processing, data sharing, and transparent communications with databases on diverse processors.

Newsletter: Is it necessary that the database management system and data dictionary also be distributed?

Schussel: Yes.

Newsletter: What distinction is there between cooperative processing and distributed database?

Schussel: Cooperative processing involves executing transactions in a network of interconnected processors. The distinguishing characteristic of cooperative processing is that transaction processing functions are executed on the processors to which they are best suited. Cooperative processing systems usually consist of a single minicomputer or mainframe networked with various PC workstations. Some of the transaction processing functions are performed on the PC and some on the mainframe or minicomputer.

Distributed database refers to the physical distribution of database files to different processors.

Newsletter: Must a "true" distributed database support transparent joins across processors?

Schussel: Yes. For the definition of "true" distributed database, I always refer to the definition developed by Mike Stonebraker et al. at Relational Technology, Inc. (now Ingres). The ability to perform transparent joins across processors is one of the requirements in their definition.

Newsletter: Must a distributed database also support referential integrity across processors?

Schussel: Distributed referential integrity support is desirable, but it is not available among current products.

One of the major problems with implementing distributed referential integrity is performance. Implementation involves a multi-phased commit process that can result in prohibitive response-time delays. Even if distributed referential integrity can be made to work properly, the costs may outweigh the benefits of this capability for some users.

Newsletter: Semantic DBMSs would involve even more integrity constraints and rules than do traditional DBMSs. If distributed referential integrity support is a performance problem in current distributed DBMSs, performance problems may be even more severe for distributed semantic databases. Are the trends we see in the database market today—toward more semantic content and distributed support—inevitably incompatible?

Schussel: Not necessarily. In many cases, the rules enforced by the database will be executable on the local processor and not require enforcement across processors.

The performance problem itself may not be quite so severe. In most applications, the majority of database accesses will be performed on the local processor. Those accesses that require external communications will usually involve a small volume of data. If large volumes of data are accessed or external access is quite frequent, database replication may be necessary.

Newsletter: Is the ability to replicate databases included in Stonebraker's distributed database definition?

Schussel: Yes.

Newsletter: How useful is distributed database proving to be?

Schussel: Three years ago when the move toward distributed database began, it was not really clear how useful the technology would be. Now many companies are discovering that distributed database is essential to implementing certain of their applications.

"One essential prerequisite for successful implementation of the client/server architecture was, of course, the establishment of a standard database interface, i.e., SQL."

Newsletter: One potential application of distributed database is supporting the distribution of repository information to multiple processors. This type of application, however, requires both semantic integrity enforcement and reasonable response times. Is this an example of an application that will be hindered by the performance constraints involved in distribution?

Schussel: Distributed database and repository are both very new technologies. As a result, there is no base of experience to draw upon to determine how significant any particular problem will be.

Newsletter: Repository, a strategic product for IBM, requires very complex, distributed management in order to be effective. Is repository management a distributed database application?

Schussel: Yes and no. IBM's approach is to require distributed processors to check objects out of the central repository in order to update them. While they are checked out they are not available to any other user. This is less complicated than true distributed database. On the other hand, DEC allows distribution of objects and updates them with a two-phased commit protocol.

Newsletter: Isn't CASE in the abstract sense, however, a distributed database application?

Schussel: Yes.

About George Schussel

Dr. George Schussel is a leading lecturer on DBMS and software trends. He is President and founder of Digital Consulting, Inc., a high-technology education and management consulting firm that specializes in advanced software tools and technologies, including computer-aided software engineering (CASE), database management systems, fourth generation languages, corporate connectivity, and artificial intelligence and expert systems. Dr. Schussel is Chairman of the Database World and Software Futures Conferences.

As a consultant and educator, Dr. Schussel has taught over 25,000 seminar attendees about state-of-the-art data management concepts and technologies. He is a frequent speaker at industry conferences worldwide. In his consulting practice, he has conducted top-level management studies for AT&T, NCR, New York Blue Cross/Blue Shield, General Electric, Software AG, Hewlett Packard, Computer Associates, and many other firms.

Dr. Schussel received his doctoral degree from the Harvard Business School. He is a Fellow of the American Association for the Advancement of Science and holds CDP certification from the Data Processing Management Association. Dr. Schussel has published over 70 articles in professional journals. ■

Client/Server Architecture

Newsletter: How does client/server architecture differ from distributed database?

Schussel: Client/server architecture is a subset of "true" distributed database. The current client/server approach originated at Sybase. Sybase's developers were familiar with the database machine concept in which the DBMS and its databases reside on a special-purpose back-end processor. In 1987, Sybase decided to apply the concept of a dedicated database processor to general-purpose hardware. Rather than building a specialized processor for database support, they wrote system software that let a general-purpose processor function as a database machine.

Early in its evolution, the benefits of this new database machine, i.e., the client/server architecture, were unclear. As cooperative processing became more popular, however, it became obvious that client/server architecture was very well suited to cooperative processing support. Given the proper software, a very inexpensive processor could function as a database machine (server) providing database support to other workstations (clients) in the network.

One essential prerequisite for successful implementation of the client/server architecture was, of course, the establishment of a standard database interface, i.e., SQL.

Newsletter: What is the market today for client/server architecture?

Schussel: Sybase, Oracle, Novell, XDB, and Gupta are the major vendors currently developing client/server software. The Oracle and Novell products have just begun delivery. IBM's OS/2 Extended Edition isn't yet available as a client/server.

Very few client/server networks are running applications today. The first applications implemented are likely to be

decision support and gateway types. Many companies are interested in client/server architecture with PC servers as replacements for mini-hosted applications.

Newsletter: What is the potential market share of client/server architecture?

Schussel: Client/server computing will eventually comprise 90% of distributed database processing, whereas "true" distributed database will comprise only 10%. Client/server computing is like the Chevrolet of the distributed database market. It is more predictable and less complicated than true distributed database and is volume-priced.

Newsletter: Given how new this technology is, is it too early to predict limits to the size of applications that may be implemented using the client/server architecture?

Schussel: No. Evidence indicates that a server on a 25 mHz 386 processor can process 10-15 TP1 (debit/credit) transactions per second on an Ethernet LAN. The comparable processing rate for an IBM 4381 is 20 transactions per second. Processing 10-20 transactions per second is equivalent to supporting 250 automated teller machines (ATMs).

Applications of this type formerly required at least a minicomputer at a cost of about $250,000. With client/server computing, PC technology can provide the necessary database support at a fraction of the cost.

"Considering current trends in hardware/software capabilities, PC-based architectures will soon be more powerful than many current minicomputers and mainframes."

Newsletter: Granted that the client/server architecture in a LAN environment can provide adequate processing rates and considerable economy, aren't up-time reliability and vendor support less acceptable in the LAN environment than in a mainframe or minicomputer environment?

Schussel: Yes. Because these products are so new, vendors lack the distribution channels to provide timely support. Companies that adopt this technology must therefore develop their own internal support staff. With respect to up-time reliability, PC LANs still don't have a track record comparable to that of minicomputers.

Given the weaknesses of PC LANs and client/server software, vendor support and up-time reliability requirements must be seriously considered for each potential client/server application.

Newsletter: How powerful will PC LAN-client/server networks become in the next few years?

Schussel: Considering current trends in hardware/software capabilities, PC-based architectures will soon be more powerful than many current minicomputers and mainframes.

Newsletter: Are there likely to be communication bottlenecks in these PC networks that will limit their capacity?

Schussel: No. With the introduction of fiber optic networks, communication rates should not be a problem.

Newsletter: How should IBM respond to the introduction of client/server architecture into the cooperative processing environment?

Schussel: IBM should respond by introducing their own version of this technology, which is what they are doing with OS/2 Extended Edition.

Newsletter: IBM has historically positioned itself more as a hardware vendor than as a software vendor. Isn't the introduction of client/server software and the resulting move away from large mainframes contradictory to IBM's historical direction?

Schussel: Client/server architecture isn't necessarily antithetical to mainframe computers. IBM has already begun to talk about mainframes as "servers" in the SAA arena. SAA provides the "glue" to connect PCs, minicomputers, and mainframes into a single network, each using common database access methods. Mainframes will be available for client/server users as upgrades when they need more power than PC-technology servers can provide.

If AD/Cycle is well-accepted by large IS organizations, it will provide further impetus to purchase mainframes.

Newsletter: Wouldn't a mainframe functioning as a database server be quite expensive compared to other competitive offerings?

Schussel: Yes, but mainframes still offer unique advantages. Even with the increased power of PCs in the future, mainframes will still be able to handle larger volumes of data.

Newsletter: IBM has traditionally avoided marketing specialized nodes—e.g., database machines—because of the potential for competition by other vendors. Does IBM's entry into the client/server market represent a significant change in this marketing strategy?

Schussel: No. To be able to use a mainframe as a server in an SAA environment, IBM software products like MVS, DB2, APPC, and AD/Cycle will still be necessary. These products will remain proprietary to IBM.

Within the AD/Cycle environment, substitution of other vendors' modeling tools and code generation tools may be possible, but the central portion—i.e., the repository manager—will be proprietary to IBM and not open to competition by other vendors.

Even though there will be considerable competition in the CASE market, IBM will capture the lion's share of software revenues from AD/Cycle users.

Newsletter: Will IBM have the largest market share even in the short term?

Schussel: Yes. When IBM begins to deliver its products, license fees will generate a substantial revenue stream.

Newsletter: Given how expensive IBM's products are and how powerful PC LANs are becoming, should we expect to see renewed competition in the computer systems market?

Schussel: The hardware market will probably become two-tiered. The top tier will consist of very large, powerful, and expensive computer systems. IBM's AD/Cycle will dominate this segment of the market. The lower tier—a "power to the people" arena—will consist of smaller machines, networked with client/server architecture. The organizations using this technology will not be interested in AD/Cycle.

Newsletter: Is your skepticism about IBM's success in the lower tier of the market due to doubts about the success of OS/2 and LAN/OS?

Schussel: No. The main problem is that IBM as a company does not command much respect from the "PC-bigot" community. As stand-alone products, IBM's lower-range hardware and software offerings are considered to be over-priced and not state-of-the-art. IBM has also been unwilling to take an "open" approach to its PC-based software by refusing to ensure that its software will run on equipment other than IBM PS/2s.

Newsletter: How do you expect OS/2 to fare in the market?

Schussel: I expect OS/2 to be a success, but not as quickly as most market analysts have predicted. OS/2 will become the dominant operating system in commercial IS organizations using client/server computing. At first it will be used primarily to support servers. As the price of 386 and 486 processors

Publication Notice

Object-Oriented Concepts, Databases, and Applications, by Won Kim and Frederick H. Lochovsky, Addison-Wesley, Reading, MA, 602 pp.
A broad introduction, covering both general concepts and instructive applications.

A Guide to DB2, Third Edition, by Chris J. Date and Colin J. White, Addison-Wesley, Reading, MA.
An extensive revision covering DB2 Version 2, Releases 1 and 2.

A Guide to the SQL Standard, Second Edition, by Chris J. Date, Addison-Wesley, Reading, MA.
A comprehensive treatment of standard SQL, with examples, and discussion of its implications.

Creating Expert Systems for Business and Industry, by Paul Harmon and Brian Sawyer, John Wiley & Sons, New York, NY, 329 pp.
A comprehensive guide to a methodology for developing expert systems in a commercial environment.

declines, OS/2 will migrate to the client processors as well. Consequently, DOS will continue to be significant in the operating system market for quite some time.

Newsletter: What future do you see for UNIX?

Schussel: UNIX will also be quite successful. It will be especially attractive to that segment of the market that is willing to buy "anything but" IBM products. UNIX will evolve from being solely an operating system to providing a much broader range of functions including database management and graphical user interface. As a result, it will compete with SAA.

Newsletter: What role does UNIX play in the client/server market?

Schussel: Client/server software can already run under UNIX. The majority of Sybase's sales, in fact, have been to UNIX users.

Newsletter: Will the adoption of client/server architecture eliminate the need for Data Administration and Database Administration?

Schussel: Absolutely not. It is erroneous to assume that because client/server software runs on a PC it requires only the level of support necessary for the typical PC. PCs in a client/server network will be used to perform essential business functions and will therefore require the same types of professional support required by any shared data system.

Data Modeling

Newsletter: Many experts feel that the relational data model, although widely accepted as the database standard, is actually a step backward in terms of semantic data modeling. In what ways has adoption of the relational model resulted in a loss of semantic content?

Schussel: One of the major failings is that relational DBMSs have not supported referential integrity, thus forcing enforcement of integrity constraints into application programs. IBM's IMS and CODASYL-based DBMSs have, in fact, provided more built-in referential integrity capabilities than many relational DBMSs.

Newsletter: What is meant by "semantic" data modeling?

Schussel: Several different development disciplines—database semantics, object-oriented programming, and repository management—include elements of semantic data modeling. Each of these areas emphasizes storing data-related knowledge in a central repository rather than in application programs. Central storage of knowledge permits uniform application to system processes and easier maintenance as well. This knowledge is stored as integrity rules, procedures, and triggers.

Newsletter: In current systems, integrity rules are usually enforced by executing procedures coded by the system developer. Isn't it preferable to be able to state integrity rules declaratively rather than procedurally?

Schussel: Yes. Some developers are currently working to make it possible to state integrity rules in natural language. An example of such a statement would be "Don't accept orders from any customer who has an outstanding balance of more than $10,000." The intent is to have an expert system interpret natural language statements in order to enforce the rules.

Newsletter: Isn't the current state of natural language interpretation rather primitive?

Schussel: Natural language interpretation is a slowly evolving field. In the first generation of natural language interpreters, the user was required to define every term used. Newer products like Intelligent Business Systems' Easytalk use expert system capabilities to interpret natural language statements. Even so, the software is somewhat fragile and can be "tricked" by the user. Expert system-based database query is a field with promise, however.

"Although IBM is something of a late-comer to the E-R arena, I think that its conventions for E-R modeling, as embodied in AD/Cycle and as supported by its business partners, will quickly become the industry standard."

Newsletter: Most of the rules required by semantic data modeling could be expressed formally. Although the ability to express them in natural language would be attractive, it is not a necessary condition for describing rules. Wouldn't it be more profitable in the short-term for vendors to concentrate on supporting enforcement of formally declared rules?

Schussel: Yes.

Newsletter: Although formerly not a major player in the E-R arena, IBM, with the introduction of the repository and AD/Cycle, seems to be moving toward adoption of the E-R approach. What do you think IBM's intentions are with respect to E-R modeling?

Schussel: AD/Cycle was developed by IBM in response to requests made by the Application Development Joint Project (ADJP), an international group of IBM customers, during the mid 1980s. In February of 1989, ADJP published a white paper in which they called for development tools that would give them productivity improvements of at least one order of magnitude. They also requested an open architecture, SAA compatibility, common user access, a migration path from existing dictionaries and development tools, and phased delivery of products with the first products delivered within one year.

Although IBM is something of a late-comer to the E-R arena, I think that its conventions for E-R modeling, as embodied in AD/Cycle and as supported by its business partners, will quickly become the industry standard. The same scenario

was seen when IBM entered the PC market. IBM was not among the first competitors, but once it entered the market, it began to set industry standards.

Newsletter: How will the E-R modeling standard be communicated?

Schussel: A definition of IBM's repository model is due to be made public in June of 1990. It will also include a statement and definition of modeling nomenclature and the definition of the application interface to IBM's repository manager.

Newsletter: In the CASE market, IBM has created alliances with market leaders who have already established their approaches to data modeling. Won't this relationship lessen IBM's influence on data modeling standards?

Schussel: No. The CASE vendors that are aligning themselves with IBM—KnowledgeWare, Index Technology, and Bachman Information Systems—will have to rewrite their software. By 1995, they will have adopted identical approaches to graphics and data modeling in keeping with the direction set by IBM.

"The market is unlikely to perceive the repository as a general DBMS, because it will lack any function at application-execution time."

Newsletter: If these companies no longer offer distinctive approaches to modeling, what will they be selling?

Schussel: They will be providing key pieces of AD/Cycle not supplied by IBM.

Newsletter: In the late 1970s, IBM actively opposed the adoption of the CODASYL database model, claiming that the technology was too immature to warrant establishment of standards. Today innovative approaches to data modeling are still being developed, among them the object-oriented approach. Isn't it perhaps too soon in the evolution of data modeling to establish a standard?

Schussel: A tradeoff is always involved in establishing standards. Standards encourage popular acceptance of a technology. On the other hand, establishment of standards may inhibit innovation.

Given that E-R modeling has been slowly evolving over the past ten years, this is probably the right time to establish standards. The CASE market needs standards for data modeling and repository management. No standards exist yet and no one sets standards more effectively than IBM.

Newsletter: Viewing IBM's establishment of standards as an advantage assumes, of course, that the standards it sets will be satisfactory. How confident are you that IBM will do a good job in choosing the approaches it will support?

Schussel: IBM's establishment of modeling standards depends primarily on delivery of a decent set of products. How well the products will be accepted in the market remains to be seen. Since a reasonable set of products won't be available and in use until mid-1991, it will be a while before an evaluation can be made.

Repository and AD/Cycle

Newsletter: At the 1989 Database World Conference in Boston, Charles Bachman described the repository manager as a "special-purpose database management system [DBMS] designed to store and maintain the complex information structures required by AD/Cycle applications." Is the repository manager a DBMS rather than a database?

Schussel: The repository manager can be viewed as a "conceptual DBMS" for managing application development objects. The physical management, of course, is handled by DB2.

Newsletter: Is the repository manager properly classified as a data dictionary?

Schussel: The repository manager is a significant extension of the traditional data dictionary concept.

Newsletter: In order to support the AD/Cycle products, the repository manager will provide extensibility, i.e., the ability to define new types of information to be stored in the repository. Given its extensibility and thus its ability to permit storage of a wide variety of data types, isn't the repository manager actually a DBMS?

Schussel: Providing some of the functions supported by a DBMS does not make the repository manager a DBMS. The repository is a database of application development logic. The repository provides database access support for application development in the same way that DB2 provides database access support for business applications.

The market is unlikely to perceive the repository as a general DBMS, because it will lack any function at application-execution time.

Newsletter: One of the major differences between a data dictionary and a DBMS is that DBMSs provide data distribution support. Would a data dictionary characteristically support distribution?

Schussel: I would not make that distinction between dictionaries and DBMSs. The key difference is their content—data dictionaries and repositories store metadata whereas DBMSs store data.

Newsletter: Isn't ensuring data integrity also a key element of database support?

Schussel: Integrity support is important, but I would not include it in the top three DBMS functions. What matters most is that the DBMS provide controlled, secure, multi-user access to a common set of data in real time. Integrity is nice to have, but it is not a fundamental requirement in today's market. Over the next few years, as expectations of DBMS support rise, integrity support will become an essential function.

Newsletter: In the repository, IBM seems to be experimenting with numerous forms of integrity assurance. Are these integrity assurance functions the earmarks of a DBMS or of a data dictionary?

Schussel: If the repository's integrity assurance functions work well, IBM will in essence be redefining what these traditional products do. The DBMS will be viewed as the physical layer underlying products like the repository and AD/Cycle.

Newsletter: If IBM is successful in introducing the repository, what other products will its users be required to buy?

Schussel: The development workstation will be the PS/2. The repository will run on a mainframe, initially only under MVS. It will later be extended to run under VM and AIX (IBM's version of UNIX).

Newsletter: Are there plans to deliver the repository for use in the OS/2 environment?

Schussel: No.

Newsletter: SAA requires that every SAA-compatible product run on each of the hardware/software platforms. Is the repository therefore not an SAA-compatible product?

Schussel: The repository manager will be a large multi-user application. Access to it will be from PC workstations, but

the single user definition of a PS/2's functions (as defined by IBM) eliminates the PS/2 from consideration as a repository manager host.

"IBM, in fact, may not want to offer a version of the repository that will run on a PC. Doing so would remove some of the incentive for IBM's customers to purchase mainframes."

Newsletter: Given the transaction processing rates already possible in the client/server environment, what characteristic of the CASE application precludes the repository from running on a server in a network of PCs?
Schussel: No technical limitation precludes a repository from running on a PC-based server. The fact that IBM is not providing this capability means that there is an opportunity for the introduction of a new product into the market.

IBM, in fact, may not want to offer a version of the repository that will run on a PC. Doing so would remove some of the incentive for IBM's customers to purchase mainframes.
Newsletter: CASE seems to be an application ideally suited to the client/server architecture. Given the relative economies of client/server computing versus mainframe computing, isn't there some doubt that a mainframe-based solution like AD/Cycle is really the best approach?
Schussel: Absolutely. AD/Cycle is going to require up-front expenditures for hardware, software, and training. The question is whether its benefits will outweigh these costs.
Newsletter: At this stage, is AD/Cycle a tangible architecture or is it primarily a marketing description?
Schussel: What exists now is IBM's endorsement of the CASE concept and its stated intention to enter the CASE market with a set of standards and functions. IBM has also announced to its customers that they must adopt AD/Cycle if they want to gain significant improvements in application development productivity.
Newsletter: How much control does IBM really have over the future success of the repository and AD/Cycle?
Schussel: IBM has traditionally had total control of the products it has introduced and as a result has had total responsibility for the products' success or failure. In this case, key roles are being played by other vendors (IBM's partners in the CASE arena). IBM, then, becomes dependent on those vendors to supply adequate tools that integrate appropriately with IBM's products. This dependency will probably not prove fatal, but it does present some interesting problems for IBM to manage.

In the PC market, IBM has depended to a certain extent on Microsoft and there is no reason to believe that similar successful cooperation will not be achieved with its partners in the CASE market.
Newsletter: What would be the consequence to IBM if the whole repository effort failed?
Schussel: IBM has many different products on the shelf and if one of them doesn't work, another will. On the other hand, some products are more crucial than others. The most central, visible products in the SAA architecture are the repository and AD/Cycle. As a result, IBM has a lot riding on these products. IBM has tended to be quite successful with its critical products.
Newsletter: Given the scope and integration potential of AD/Cycle, some market analysts predict that existing CASE vendors will move away from developing integrated products. Do you agree with this prediction?
Schussel: Absolutely. AD/Cycle's relationship to the repository is like a bus architecture. Although some of the CASE vendors may develop interfaces among their products, in the long run a bus architecture will be much more stable than one-to-one couplings.
Newsletter: When VSAM was introduced, it was said that storing all data in VSAM files would promote compatibility and integration. That, of course, did not happen. Similarly, if vendors store information in the same repository, they have not necessarily integrated their products. Will using the repository guarantee true integration among CASE products in the AD/Cycle framework?
Schussel: No. Without a common data model, there is no simple way to truly integrate the various CASE tools.
Newsletter: What will happen to CASE vendors like Texas Instruments that have differentiated their products by offering an integrated set of development tools?
Schussel: Conceptually, IBM's approach is superior because it offers the ability to "snap in" other vendors' products. Its superiority depends, of course, on the CASE tools being able to pass data to one another.

Nothing as comprehensive as AD/Cycle has ever been tried before and it remains to be seen if it will work, both technically and economically.
Newsletter: Hasn't the industry's experience with database proven that passing data from one system to another is not an effective means of integration?
Schussel: Yes. If there is no common, underlying data model, nothing much is gained by interfacing tools.
Newsletter: Do you believe that the CASE tools that will plug into AD/Cycle will share a common data model?
Schussel: I think that KnowledgeWare's data model will be the closest to what IBM endorses. Over time other vendors will either adopt this approach in order to fit into AD/Cycle or adopt an "open" approach like Index Technology's in order to offer a product that can be used more generally.
Newsletter: KnowledgeWare's entity modeling approach does not support type hierarchies (a necessity for object-oriented support) and its association types are inadequate. Do these inadequacies concern you?
Schussel: I am not expert enough in that area to know what impact those inadequacies will have.
Newsletter: Do you anticipate that the ANSI IRDS standard will have any significant impact on this market?
Schussel: No. I think it will generate considerable discussion, but have little real impact. From my discussions with key players both inside and outside IBM, ANSI IRDS seems to be primarily useful for documentation, more oriented to Data Administration than to CASE, and inferior to the ISO definition developed in Europe.

"Effective Data Administration and IRM envision a new role for the "user" community, too. In fact, the term "user" is no longer even appropriate because information-age workers are part of the process."

I. Cola

DATABASE RESEARCH GROUP, INC.
31 State Street, Suite 800
Boston, MA 02109
(617) 227-2583

Address Correction Requested

DATA BASE NEWSLETTER

In this Issue:

An Interview with George Schussel 1

Object Orientation and
Its Software Implementation
By James J. Odell 1

Object Database
By Mary E.S. Loomis 1

A Quick Survey of IBM
Repository "Data" Concerns 3

Logical Conclusions 2

Publication Notice 17

Newsletter: Looking at the direction taken in AD/Cycle, would you agree that the primary thrust is a retrofit to the tools and techniques of the 1970s and 1980s?

Schussel: The CASE market is about to embark on its third generation. The first generation consisted of stand-alone dictionaries and normalization tools. These products were interesting, but not very widely accepted in the industry. The second generation began in the mid 1980s with the popularization of the term CASE itself. Numerous products were introduced during this time, addressing modeling and code generation. Although there was considerable activity, there was little penetration into the application development community.

The second generation of CASE ended on September 19, 1989 with the announcement of AD/Cycle. This new generation should continue through the 1990s.

Newsletter: Given that IBM is basing AD/Cycle on existing techniques, do you expect innovation in CASE methodology to stagnate?

"You can enforce integrity, but you can never completely eliminate stupidity."

I. Cola

Schussel: The standards that IBM will establish relate to representation of designs, not to the methods of their development.

Newsletter: Was there any reduction in the level of interest in CASE during 1989?

Schussel: No. Purchasing of CASE products slowed in anticipation of IBM's introduction of AD/Cycle, but interest in CASE itself did not decline.

Newsletter: Is the slowing of the CASE market also due in part to failure of the current generation of CASE products to provide the productivity gains sought by their users?

Schussel: Yes. The major issue, however, is IBM's redefinition of the market.

Newsletter: Should the typical, large mainframe MIS organization begin preparing now for the repository and AD/Cycle?

Schussel: Yes.

Newsletter: Most of the architecture of AD/Cycle and repository has yet to be announced. Given that most of the functionality of AD/Cycle will be developed by other vendors, and that client/server computing may offer an attractive alternative to mainframe-based approaches, how can companies prepare intelligently for the future?

Schussel: Companies can make use of their existing data dictionaries to clean up data definitions in anticipation of entering information into the repository. If they intend to adopt AD/Cycle, they can install DB2 and begin to develop organizational expertise with the product. Companies should also create a Data Administration function if they have not already done so and begin training IS personnel in E-R modeling. ■