April 30, 2003

Noted in Passing — The Author of a Great Idea, Edgar F. Codd (1923 – 2003)

Lou Agosta

Catalyst

Dr. Edgar Codd, IBM Fellow (retired), computer pioneer, creator of the relational database, passed away Friday, April 18, 2003 at the age of 79 of heart failure.

Question

What is the significance of Codd's life and work ?

Answer

The relational database is the dominant design for the storage and manipulation of data relevant to commercial business applications, data warehousing systems and business intelligence, and it is likely to continue to remain so for the foreseeable future.

Specific niche markets in emerging domains such as genomics and the representation of bio-informatics may require or benefit from a new paradigm for the representation of data, but the dominance of the relational model in commercial business operations is secure. Bank accounts, credit cards, stock trading, travel reservations, online auctions and innumerable other now-routine data transactions as well as data warehouses that address them all rely on Codd's model. This model was first implemented by Larry Ellison in an early version of what would become the flagship **Oracle** database. This immediately captured **IBM**'s attention and helped to overcome internal skepticism about Codd's work, which then provided the basis for IBM's own prototype SQL/DS (1981) and DB2 (1983). Other early commercial products based on the relational model included **Sybase** SQL Server, which originally shared code with what is now **Microsoft** SQL Server, **Teradata** (NCR) and **Ingres** (Computer Associates).

One source of the power of this model — the relational model — is that it says nothing about the physical implementation of the data, but provides a simple set of logical operations — union, intersection, negation — from which virtually all other transformations of the data can be derived. Another source is its simplicity. The model is supposed to be implemented by means of a symbol system that employs a small, simple set of English language statements (declarations such as Select, Insert, Update, Delete) to data organized intuitively into rows and columns (tables). The irony is that a system's strengths can also becomes its weaknesses. If anything, the challenge is that the relational model is so simple and elementary. SIMPLE does not always mean EASY — it gets under the radar of common sense and the everyday sloppiness with which most people are comfortable. When taken to an elementary level, even simplicity can be challenging. In addition, the subsequent choice of the term RELATIONAL ALGEBRA by the colleagues at the Lab to describe the syntax of the relational model was obviously coined by software developers and never vetted by the marketing department. It retains a certain aura of mystery — and can inspire fear in the heart of many people for whom algebra was not a good experience in high school. Yes, logical abstraction and mathematical (set) theory is to be found in the background. Yet the intuitiveness of the rows and columns of the table is clearly in an entirely different context as spreadsheet has come to dominate business analysis on the desktop.

What the relational model has had all along (in contrast to the desktop) is integrity — data integrity. Another advantage is that the form of the relational SQL statements is declarative — the user tells the system what to do, not how to do it, which is left to the underlying database itself. This is in contrast to procedural programming languages such as C, which require experts to disentangle the convoluted syntax and so will remain the domain of specialists. **S**TRUCTURED **Q**UERY **L**ANGUAGE (SQL) has always been available to business analysts who were willing to make a modest extra effort without having to become full-fledged developers. It is now an interface for data mining, ETL (extract, transform and load) and a variety of business intelligence analyzes. Thus, if you are looking for a Great Idea that has still not been exhausted — and, by the way, that is one way to tell a Great Idea, its inexhaustibility — see Codd's 1970 paper "A Relational Model of Data for Large Shared Data Banks," *Communications of the ACM* 13, no. 6 (June 1970). Reprinted in *Communications of the ACM* 26, no. 1 (January 1983). For text of the paper see www.acm.org/classics/nov95/.

This assertion can be appreciated in the prediction that all other paradigms without exception will be assimilated to the relational one (for example, see Planning Assumption[1]). That has *already* happened with object-oriented databases (one or two of which have found a niche in specialty verticals such as publishing or avionics). Object-relational extensions are now a feature of the standard relational database in the form of user-defined data types, user-defined functions and inheritance mechanisms. This will likewise happen with in-memory databases, OLAP databases and XML databases. This leads to a clear recommendation for clients — absent very specific industry-specific requirements, do not rush to purchase one of these special-purpose data stores, but rather wait for the functionality to be assimilated to the next release of your standard relational database.

So, like so many touched by genius, Codd has gone from being a voice crying in the wilderness, to being merely impractical, to stating something so obvious that everyone knew it all along. Some readers will take heart from the example of Codd with their own struggles in the corporate jungle in that he received a less than satisfactory review from superiors at IBM at Poughkeepsie, NY, early in his career, leading him to move west to California to seek new opportunities at the IBM Santa Teresa Lab. As they say, the rest is history

Related Giga Research

1. Planning Assumption, IT Trends 2001: Data Warehousing, Lou Agosta

Additional Research

For more details on Codd's life and work, including a standard obituary, see "Computer Engineer, Dead at 79, Revolutionized a Computer Database System, Lisa M. Krieger, *The Mercury News*, April 20, 2003, www.bayarea.com/mld/mercurynews/news/local/5676110.htm