

Pioneers in Database and Information Systems

Nocker Martin

June 10, 2015

Abstract

This paper gives a brief overview about four pioneers in the field of database and information systems who have all won ACM's Turing Award.

Contents

1	Introduction	1
2	Charles W. Bachman	1
3	Edgar F. Codd	2
4	James N. Gray	3
5	Michael Stonebraker	3
6	Conclusion	4

1 Introduction

Databases are used nonstop around the world today. Database management and information processing has become a multibillion dollar business. This paper presents four men that dedicated their lives to the science of database development so that we can use today's technologies that would not be the same without these personalities. All of them won the A.M. Turing Award and they are presented in the order they received the award.

2 Charles W. Bachman

Bachman, born 1924 in Kansas, started his career at Dow Chemical in Michigan working as a chemical engineer. When the company had to face financial problems and canceled the order of the IBM 709 (the first computer which should have been used at Dow) Bachman moved to New York City where he started to work for General Electrics. There his group was working on the "MIACS" project, an integrated management and calculation system.[4] But the important part was the underlying "Integrated Data Store" (IDS) which is considered as the first database management system. IDS, Bachman's most famous contribution to the DBMS development, introduced the network data model and

provided commands to define and manipulate data within the database, today known as Data Definition Language (DDL) and Data Manipulation Language (DML) respectively.[2]

He also came up with the first data structure diagrams known as Bachman Diagrams. During the late 60s he joined the Data Base Task Group (former List Processing Task Group) of CODASYL (Conference on Data Systems Languages), which is famous for introducing the COBOL programming language. They standardized IDS and supported Bachman's network data model against the less flexible hierarchical data model by IBM.

In 1973 Bachman received ACM's Turing Award "for his outstanding contributions to database technology"¹ surprisingly not knowing who Alan Turing had been at that time.[2] Bachman was the first winner for work in the database area [6], the first without a mathematical or physical background and the first who spend his entire career in the industry rather than in academia. Another fact that shows Bachman is a special Turing Award winner is that he is one of few winners without a Ph. D.

In his award accepting Turing Lecture "The Programmer as Navigator" he reviews the database management revolution encountering a change from a computer-centered view to a data-centered view within database management.

3 Edgar F. Codd

Bachman's and his network data model's probably biggest rival, although in a friendly manner [4], has been Edgar F. Codd. Codd was born in England in 1923. In 1948 he graduated from Oxford University where he studied chemistry and mathematics.

In 1949 he joined IBM working at the Selective Sequence Electronic Calculator (IBM's first electromechanical computer) starting his computing career. Codd contributed to important IBM products before attending University of Michigan where he acquired a M. Sc. and Ph. D. in communication science in 1965.

By moving to the IBM Research Laboratory in San Jose, California, he got to the place of his biggest impact. There he concentrated on database problems. He realized that existing databases were only able to be used by special skilled people and above all they didn't have a solid theoretical foundation. Codd used his knowledge of mathematical logic and brought up the invention which turned out to be a key turning point in the history of databases. His idea of the relational data model changed database management due to its theoretical framework into a scientific and academic discipline.[3] Codd's initial paper "A Relational Model of Data for Large Shared Data Banks" where he introduces the relational model is considered as the single most important paper in database management.[7] The following years Codd described and explored the model's implications in several further papers.

Codd and Bachman, the creators of the two competing data models, even took part in a public debate, trying to convince the other.[7] Bachman's concern was that relational query languages are impossible to implement efficiently.

Along with the relational model he defined probably the first relational language "Alpha" by realizing to use predicate logic as a basis for database languages. His further achievements are the definitions of the relational calculus,

¹http://amturing.acm.org/award_winners/bachman_9385610.cfm

the relational algebra, the first three normal forms and the concepts of functional dependence.

Codd died in 2003 in Williams Island, Florida.

4 James N. Gray

After Codd's revolutionary idea of a relational data model, implementations started to emerge where Ingres (at Berkeley, University of California) and System R (IBM, San Jose) were the leading ones. One of the developers of System R was James Gray.

Born 1944 in San Francisco, California, he got a Ph.D. at the University of California at Berkeley in 1969 for his work on context-free grammars and formal language theory in the newly formed computer science department.

His main work within System R, also involving the initial development of SQL, covered concurrency control, crash recovery, system start-up, security and administration.[5] In 1976 he released his pioneering paper with his revolutionary but actually simple ideas. Gray suggested to divide DBMS actions into units of work, so-called transactions. Further, every transaction has to follow the "ACID-properties" which stands for atomic, consistent, isolation and durable. To make it short, ACID means that a transaction may only happen entirely or not at all, that the database has to be in a consistent state after the transaction, that two concurrently working users cannot see any intermediate state of the database of the other user and finally that the effect of a transaction will stay in the database no matter what failure might happen afterwards.[8] The terms transaction and ACID are well known for every computer scientist who gets in touch with databases today.

In 1998 he received the A.M. Turing Award "for seminal contributions to database and transaction processing research and technical leadership in system implementation"².

In 2007 James Gray didn't return from a boat trip and was named dead five years later.

5 Michael Stonebraker

As mentioned in chapter 4 Ingres was one of the first relational database implementations. The leading architect of Ingres, developed at the University of California in Berkeley, was the American Michael Stonebraker. Within this project he set milestones at query language design and query processing techniques.[1]

Stonebraker also succeeded with Postgres, a database following the object-relational data model paradigm he introduced. This model implements the important ideas of object-orientated-programming including abstract data types which may have user defined operators and procedures, while still keeping the database apart from the programming language³.

In 2000 Stonebraker left Berkeley where he worked as a professor for 29 years. Today, he works as Adjunct Professor at M. I. T. and recently worked on column-oriented storage which is optimized for complex queries.

²http://amturing.acm.org/award_winners/gray_3649936.cfm

³http://amturing.acm.org/award_winners/stonebraker_1172121.cfm

Michael Stonebraker is the winner of the 2014 A.M. Turing Award “for fundamental contributions to the concepts and practices underlying modern database systems”. [1]

6 Conclusion

This paper gives a short summary of four pioneers of the database and information system field. Of course, within this article not every achievement of the characters above can be mentioned in detail but the most important contributions are discussed.

References

- [1] AcM turing award goes to pioneer in database systems architecture. *ACM News Release*, 2015.
- [2] C. W. Bachman and T. Haigh. Charles w. bachman interview: September 25-26, 2004; tucson, arizona. *ACM Oral History interviews*, 2006.
- [3] C. Date. Edgar f. codd: A tribute and personal memoir. *Sigmod Record*, 32(4):4, 2003.
- [4] T. Haigh. Charles w. bachman: Database software pioneer. *Annals of the History of Computing, IEEE*, 33(4):70–80, 2011.
- [5] D. A. Patterson. A conversation with jim gray. *ACM Queue*, 1(4):53–56, 2003.
- [6] R. Ramakrishnan and J. Gehrke. *Database management systems*. Osborne/McGraw-Hill, 2000.
- [7] B. Shneiderman. *Database Management Systems: A Review with References*. Computer Science Department, Indiana University, 1976.
- [8] M. Stonebraker. Why did jim gray win the turing award? *SIGMOD record*, 37(2):33–34, 2008.