

Week6 Paper Summary——Relational Model

By Yi Zhou

Codd's paper is considered as one of the most important paper in the history of computer science due to the fact that it introduces a relational model of data for large shared data banks which is the base of many modern databases.

At Codd's time, there were many different types of model for storing data such as network model, ordering model, tree model and so on. Whereas, in the first and second part of this paper, Codd pointed out that these model were not flexible enough to tackle with complex queries from clients because they introduced many dependences such as ordering dependence, indexing path dependence and access path dependence. The aim of relational model was to reduce such dependences and provided a means of describing data with its natural structure only. Relational model is based on a mathematical concept called relationship which could be regarded as subsets of Cartesian product of several sets. Such model of data is proposed as a basis for protecting users of formatted data systems from potentially disruptive changes in data caused by the growth of data amount and change in traffic. It's a normal form for the time-varying collection of relationship. Besides, Codd introduced the key concept in such model and used mathematical tools which we called relational algebra to show how to build a robust relationship model.

As far as I am concerned, Codd's paper shows the power of mathematics. I do believe that he might see such concept on a mathematical book one day and then he was inspired to introduce such concept to tackle with the problems in that-time database system. Such model brought a lot of convenience to human beings and gave Codd a Turing award. Larry Ellison implemented such model and then built Oracle after reading such paper. However, relational model based database might not be that powerful when it comes to big data and more advanced NoSQL databases such as MongoDB have been created to tackle with such problem.