Jong Tai Kim

10/8/2018

Info 330

Assignment 1

Database Design

Introduction

In this paper, we are going to look into the significance of database system in today’s life, the importance of the relational database, and the strength and brief useful properties of normalization process of data.

Question 1: Purpose of Database

As we live in the generation of information, we are dealing with a lot of information floating around us. It indicates that, to deal with those information, we need some kind of system which allows us to use those information into meaningful way. Moreover, the electronic system is preferred as it will manage a lot of information faster and more efficient than system by us humans only. Especially, the large companies and organizations are dealing with millions of different information for decision making which are impossible to be managed only by human brains, such that they inevitably need for some electronic system. The electronic system above is called the database system, which allows us “to organize, process and manage information in a structured and controlled manner” (Kimani D., Introduction to Databases, techopedia, 2018). In other words, it is a key of modern life of business and organization to find such a way that can structure collections of information given to their benefits. As an example of my life --even if it is not done by electronic--, as a student who is looking for jobs and graduate school, it is crucial for me to organize the jobs and grad schools by their requirements, locations, application due dates, and so on, so that I can use those organized information to see what I have to do to apply to the jobs and graduate school and to find which job and grad school I can apply. Thus, the database of list of jobs and graduate schools allows me decide what to do for applying jobs and grad school.

Question 2: Benefits of Relational Database

The relational database is a kind of database where we organize the data by relations to each data. To do so, we create group of tables, also known as entities, which, simply saying, are created to represent relatable data present in each table. To create simple relationship between the tables, we simplify the data in the table by assigning keys, which can relate the data in one table with the same data in other table. Thus, the relational database can simplify large amount of data by the data’s relations to each other. As you might notice, this is huge benefit of relational database design compared to other simple databases, such as paper-based system. For paper-based system, it is very hard to process the data based on its relational property even though we sort the data by its relational properties, indicating that the paper-based system has inevitably inefficient. For instance, in paper-based system, it is hard to find and compare each data to other, and to analyze relational data based on physical appearance --paper. Thus, the relational database design allows us to create quicker and more efficient database design than just paper.

Question 3: Purpose and Benefits of the Normalization

The normalization process is one of key processes to create tables (entities) based on relational database design. The main purpose of the normalization process is to ensure that the data structures are optimally representing data. That is, the normalization can “reduce data redundancy and improve data integrity” of the database, which simplify “the design of a database so that it achieves the optimal structure (<https://en.wikipedia.org/wiki/Database_normalization>, 2017). For instance, the normalization process requires us to represent columns as small as possible so that represents only single value, and to represent each row of table depended on unique identifiers, and to set that unique identifiers only identifier for that row. Furthermore, to follow the rules, we need to create tables that represent single subject, such that only unique identifiers in each table represent rows of the table. Those rules benefit us in a way that we can reduce unnecessary data presented in each table to simple unique identifiers in one table presented in other tables; this not only makes the database look simple but also efficiently managing relations of tables to be referred in others.

Question 4: “The Key, The Whole Key, and Nothing But The Key”

“The Key, The Whole Key, and Nothing but the Key” is the summary of normalization process in relational database design. (Kent W., A Simple Guide to Five Normal Forms in Relational Database Theory, Communications of the ACM, 1983). “The Key” represents the first rule of normalization that the table is in atomic form --each column has single value-- to have the key, “The Whole Key” represents the second rule of normalization that each row of tables should depend on a unique representing key (or set of keys), “Nothing But The Key” represents the third rule of the normalization that the row would be dependent on the chosen primary key, but not the other values in the key.

Question 5: Purpose of Associative Entities

Associative entities are to ensure that there would not be any violation on rules of normalization in the many-to-many related table. In other words, as many-to-many relations are in the situation that both tables would have many instances of each other, indicating that table A would have many instances of table B’s unique primary keys, and table B would also have many instances of table A’s unique primary keys. In other words, only two tables for the many-to-many relations inevitably violate the rules of normalization by table representing more than one subject, or rows having more than one set of keys representing themselves. For instance of associative entities, similar to the bridge tables in lecture, we can consider the case of classes and students, which is many-to-many relations as students can take many classes as well as classes can have many classes. Then, if we create bridge table which take both students primary keys and classes primary keys as foreign keys and new idea like in “Sales Line Item” in example of lecture as primary key joint with students primary keys then the bridge table successfully represents both keys without violation of normalization (Root R., Module01Notes, 2018).

Summary

The five topics represented in this paper describe how the database helps managing data, how relational database is better than paper-based system, how the normalization process is necessary for relational database design, and what associative entities are.