Mason Howard

Chapter 2 Homework

Database Management Systems

01/20/2014

RQ4: What languages emerged to standardize the basic network data model, and why was such standardization important to users and designers?

Data manipulation language, data definition language, Structured Query Language, Unified Modeling Language, etc., emerged to create a semi-coherent way of describing data, requests for data, etc. The standardization that occurred prevented the industry from going every which direction when it came to getting a solution for their particular business.

RQ5: Describe the basic features of the relational data model and discuss their importance to the end user and the designer.

The relational model is basically taking the concept of a relation (mathematics) and calling it a tuple, where each item represents an attribute. It's using one or more fields in a table to define relationships between tables, to minimize repetition of data. It makes the data relatively intuitive to think about for users, and it's easy to design for the designers.

RQ8: Why is an object said to have greater semantic content than an entity?

The relationships are included in the object's model, rather than defining it with a relational data set. It's intended to be built like real world objects, like object oriented programming.

RQ11: What is an ERDM, and what role does it play in the modern (production) database environment?

It adds many of the object oriented model's features within the inherently simpler relational database structure. It mixes the two, as well as extending data types based on classes and inheritance. One of the results of the ERDM is XML. XML is used for structuring semi-structured data.

RQ15: What is a table, and what role does it play in the relational model?

A table is an array of tuples that contain attributes. Or, it's one of the easiest way to represent data to a person in a format that they can understand relationships about data. Attributes in the tuples specify relationships with other tables or columns / rows in other tables.

RQ18: Describe the Big Data phenomenon.

I heard this quote somewhere: “Big Data is like teenage sex: Everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it too.” And I could see it being true. Collecting data and running analysis on it has never been easier, which can lead to a lot of faulty assumptions about what the data mean. “Big Data” is just a buzzword. Depending on its use, it can mean anything from youtube's handling of uploaded videos, to managing video streams and connections, to running real time analysis on twitter. Mostly it's a cop out way of saying “Our company has a large online presence. We need to find out what users are doing on our website, what is making us money and how much it's making, what's not making us money and how much is it costing us, and what's trending and how do we take advantage of it?”

If I'm interpreting the book right, it's cost effective SWOT analysis through IT.

RQ19: What's sparse data? Give an example.

Sparse data is data that have a very large number of attributes, but the actual number of data instances is low. It'd be like a hospital medical record. Many, many, possible data points, but only a few are populated.

RQ7 && RQ17: In a separate file

RQ13 && RQ14: What is a relationship, and what three types of relationships exist? Give an example of each of the three types of relationships.

A relationship is a mapping of one set to another set as defined by a set of rules.

* One to One: One student, one Lnumber
* One to Many: One student, many classes
* Many to Many: Many students, many majors

RQ16: What is a relational diagram? Give an example. In what way does a relational diagram differ from an Entity-relationship diagram?

A relational diagram is an image or drawing that illustrates one or more relationships between one or more sets of data, as well as the fields in each data set. I've attached an example in a separate file.

A: Compare Fig 1.3 and Fig 1.4 in Chapter 1. Why is the structure of Fig 1.4 preferred over Fig 1.3?

There's less repeated data, more purpose to each table, the option of extending the database and adding more tables easily, better legibility, clearer relationships among the data.

B: For Fig 1.5 in Chapter 1 … Identify the 3 (or 4) tables which could be created from the file

CUSTOMER, containing name, phone, an address code, insurance type, amount, renewal date, agent code

AGENT, containing name, phone, address code, agent code

ADDRESSES, street, street number, city, zip code,