

What Goes Around Comes Around

Long Nguyen, 6.814 - Database Systems

September 9, 2017

Reading Questions

1. What is the notion of data independence? Why is it important?
2. What are the key ideas behind the relational model? Why are they an improvement over what came before? In what ways is the relational model restrictive?
3. What are the most important differences between the "hierarchical" model (as exemplified by systems like IMS) and the relational model that Codd proposed?

Answers

1. Data independence is, in a sense, abstraction from the details. It's important for future modularity and readiness for change. We have *physical independence* and *logical independence*. The first refers to physical changes to the data representation having little to no effect on the function of the DBMS app. The second refers to changes to logical data views having little to no effect on functionality and physical representation. In essence, we want both types.
2. The relational model aims to work with a simple model, a high-level manipulation language, and no dependence on a particular physical data rep. It's different with its focus on abstraction, and potentially, this means less chance to delve to the core details should a designer wish to do so.
3. The hierarchical model imposes heavy restrictions on data flow, which is helped by the relational model. Also, with a tabular rep, more fields/details can be represented. In the end, each has its benefits.