Introduction
CMSC 5705 Advanced Topics in Database Systems

Yufei Tao

Department of Computer Science and Engineering
Chinese University of Hong Kong

September 7, 2010
A *database* is a system that manages a set of data in order to facilitate a set of operations. Typically, the functionality of the system is to store the data in a proper format, so that a certain type of queries can be answered efficiently.
A *relational database*, which is perhaps the type of databases we are most familiar with from our undergraduate experience, organizes tabular data (as in spreadsheets) in order to support queries defined by *relational algebra*, or almost equivalently, queries that can be formulated using the SQL language.
Although relational databases serve the needs of traditional data management well, they are insufficient in numerous modern, emerging, applications mainly due to two different reasons:

- The nature of the data is no longer limited to simple values that fit in a spreadsheet, and has become much more versatile.
- People need to perform novel types of operations on their data, or to place new requirements on the existing operations.

Techniques beyond relational databases must therefore be developed to meet with these arising demands.
In this course, we will discuss advanced data management methods that were proposed in recent years to resolve important issues in numerous domains. These methods share two common features:

- Their effectiveness has been confirmed by plenty of experiments on real-world data.
- They can be easily implemented in practice, even if the implementation needs to start from scratch.