CPSC 504:
Data Management

Rachel Pottinger
Course Introduction
2009/01/05

What is this class about?

- Understanding how people design database management systems
- Data management research
- Managing your data

What is data management research?

- Research about managing data including:
  - Traditional (relational) database management systems
    - What they are, how to make them work
  - Other kinds of databases
    - Object-oriented, XML
  - Other data management applications
    - OLAP, data mining, etc.

This class is a seminar

A seminar is "a small group of advanced students … under the guidance of a professor who meets regularly with them to discuss..." [dictionary.com]
- 1 or 2 papers to read for most classes. I'll provide:
  - An explanation why we're reading the paper
  - Necessary background beforehand
  - Suggestions on how to read papers where necessary
- Most days students present papers and lead discussions. You'll present once and lead discussion once. I'll provide:
  - The high level goals of reading papers
  - A set of suggested discussion questions
  - Feedback on your plans and answers to questions
  - Possibly a preliminary suggested set of slides
- Sometimes I'll present the papers and lead discussion

But I haven't taken any database classes!

- There are no set prerequisites
- You do not need to have taken a database class
- Assuming you have a solid basis in computer science, I will provide you with all background material you need

What *are* the prerequisites for this class?

- Ability to read and respond to 1 – 2 papers a class
- Ability to do a project (not necessarily implementation based) either in a group or on your own
- Ability and willingness to present papers and lead discussions
- Willingness to discuss your own ideas and questions in class

Other handy things: databases, AI, logic

Where does my grade come from?

- Analyzing the readings – 30%
  - Post a summary/analysis on WebCT Vista
- Presenting/leading class discussions – 20%
  - One person presents the content
  - One person leads discussion
  - You'll sign up for *different* papers for this
- Course project – 30%
  - See website
  - Doesn't have to have an implementation
- Small number (1?) of homeworks – 5%
- In-class participation – 15%

7

Any administrative questions?

8

Introduce your partner

Discuss the answers to the following questions with your partner for next 5 minutes:

- What is your name?
- What is your affiliation with UBC?
- What is your favourite colour?
- Where are you from?
- What is your database/data management background?
- What do you want to get out of this class?

Grab a piece of paper and a marker and write your name on it

9

Some reasons to use a database:

- Large amounts of data
- Structured data
- Persistent data
- Valuable data
- Performance requirements
- Concurrent access to data
- Restricted access to data

10

What data is stored in databases?

This space intensionally left blank

11

What data do you have?

Here is some data *I* have:
- Papers I've read
- Addresses
- Job search data
- Experiments I've run
- Grades
- CDs, DVDs, and books I own
- Powerpoint slides
- Research notes
- E-mail
- Drafts of research papers and notes from students

12

Class outline

- Crash course in databases
- Standard Relational databases
  - History
  - Query Optimization
  - Query Evaluation
  - Transaction processing
- New Relational Databases
  - Distributed Databases
  - Data integration
  - Adaptive query processing
- Other data models
  - Object Oriented & Object Relational databases
  - XML
- Management of other data
  - OnLine Analytic Processing (OLAP)
  - Data Mining
  - Streaming Data
  - Evolution of Databases
  - Role of theory
- Advanced topics/Student request potpurri

13

To do:

Course website:
http://www.cs.ubc.ca/~rap/teaching/504/2008
- Mailing list: mail majordomo@cs.ubc.ca with "subscribe cpsc504" in the body
- Make sure you can access WebCT Vista (see course webpage)
- Think more about data you have
- Think about which topics you'd like to present/lead discussion on (first come first served)
- Read the project description, and think about projects
- If there are any topics you'd like to see covered in class that aren't, let me know
- If you're waiting for me to sign an add form, see me now

14