CS 4400 Database Systems

Meeting 1: Introduction
Brandon Myers
University of Iowa
One of my research projects

select count(*), mass from Particles
where x > 0 and x < 10
group by mass/10

Why is 4400 important?

Data management is critical to
• business operations and innovation
• scientific discovery
eCommerce business

Warehouse operations

Website

Digital media content

products

users

stock

orders

Analysts predicting demand for Harry Potter

"Users who viewed this item also liked..."
Oceanography

original footage; [http://oceanobservatories.org/](http://oceanobservatories.org/); video by Aaron Marburg
Intellectual property

Method for preventing formation of ice slush in an ice maker
www.google.com/patents/US5582018
Grant - Filed Aug 30, 1995 - Issued Dec 10, 1996 • William J. Black - Scotsman Group, Inc.
A method for preventing the formation of icy slush within the sump of an ice maker. The method includes the steps of monitoring the ...
Overview • Related • Discuss

Ice cream sandwich and method of making the same
www.google.com/patents/US5789008
Grant - Filed Jan 31, 1996 - Issued Aug 4, 1998 • Woodrow C. Monte • Neal B. Julien
An ice cream sandwich includes a layer of ice cream intermediate a pair of crispy cookies. The cookies resist migration of moisture from the ice ...
Overview • Related • Discuss

Dry ice delivery method that controls the temperature of a cooling ...
www.google.com/patents/WO200711675A2\%3Aus\%3Den
App. - Filed Oct 19, 2006 - Published Oct 4, 2007 • Robert E. Whewell, Jr. • Whewell Robert E Jr
A cooler/container with a dry ice sublimation regulating system having an insulated dry ice module that encloses dry ice so that the module's ...
Overview • Related • Discuss

Generation of ice-nucleating crystal
www.google.com/patents/US3127107
Grant - Filed May 29, 1961 - Issued Mar 31, 1964 • John Patrick Merryweather
United States Patent 3,127,107 GENERATION OF ICE-NUCLEATING CRYSTALS John Patrick Merryweather, New Castle, Pa., assignor to Canadian Safety ...
Overview • Related • Discuss

Road-surface additive for preventing ice and melting snow
www.google.com/patents/US4094676
Grant - Filed Nov 24, 1976 - Issued Jun 13, 1978 • Robert Dubois - Plastiroute S.A.
An additive for admixture with road surface preparation materials to prevent formation of ice and to melt snow on the finished road surface.
Overview • Related • Discuss

which inventions have been cited in many disciplines?
Let’s look at the syllabus

- ICON > Syllabus > link
- or, directly go to course website:
  - http://homepage.cs.uiowa.edu/~bdmyers/cs4400_sp17/syllabus/
ICON Discussions

• Use the ICON discussions board: post replies, post your own topics, as long as it is related to 4400
  • Teach each other. Your peers might answer the question first
  • Of course, academic integrity policy applies. Intellectual arguments are great. Posting answers is not okay. If in doubt, ask Brandon
Be successful in CS4400

• There are ~8 homeworks, many involving programming
• Come to class; reading slides is a poor substitute
• Active learning, peer teaching, and other activities to replace the lack of labs/discussions
  • This is a small class, which makes it easier to have group activities!
• Attend Debug Your Brain and/or office hours
• Help your classmates in class, on ICON discussions
• Midterm in class (March 3)
• Keep on top of announcements in ICON/website
Peer instruction

• Think, answer, discuss...
• Participation counts, *not* right answers

• Some class meetings involve use of your computer (need at least 1 per pair of students)
  • If bringing a laptop presents a hardship, email me
What to do now

• HW 1 is out
• Read the syllabus online
  • and vote on the class policies
• Check ICON and reply to the discussion question
• figure out what 2 dates you want to present on
(next)
Databases and Database management systems (DBMS)

- Examples of databases

- Examples of DBMSs
An example: online music streaming service

• What data must it contain?

• What capabilities are needed?
Summary of data management requirements

1. Able to describe real-world entities in terms of stored data
2. Persistently store large datasets
3. Efficiently query & update
4. Change structure (e.g., add attributes)
5. Concurrency control: enable simultaneous updates
6. Crash recovery
7. Security and integrity, provenance

DBMS provides these so that users can focus on application logic
Data structures and databases

• In CS2230 (or equivalent) was all about data structures
• What is the difference between a database and a data structure?
Data warehouses to data lakes

• Conventionally, businesses would have:
  1. Business operations supported by: a DBMS for transactions (e.g., sales, supply chain orders)
  2. Business intelligence supported by: a DBMS for storing a structured and indexed archive of recent and historical data (think library) called a data warehouse. Employees analyzed the data to inform decisions.

• Today, companies like Microsoft refer to data lakes, replacing the carefully maintained databases of a data warehouse with enormous quantities of raw data

• When the data needs to be analyzed, it is transformed with parallel processing systems

• in 4400 we’ll explore XML (semi-structured data), parallel processing, and non-relational systems (“NoSQL”)
People and databases

1. App developer: writes programs that update and query the data in the DB
2. DB designer: models the data by choosing tables and their attributes
3. DB admin ("DBA"): operates the database, diagnoses performance problems
4. Data analyst: data mining (inferring useful information), data integration (combining disparate data)
5. DBMS implementer: builds the DBMS

In 4400 we’ll try to give you some experience in all of these roles, although 4 and 5 are huge topics that demand their own courses
Attribution

• Some slides inspired or quoted from UW CSE 344
  • People and databases
  • Data warehouses to data lakes
  • Summary of data management requirements
  • https://courses.cs.washington.edu/courses/cse344/