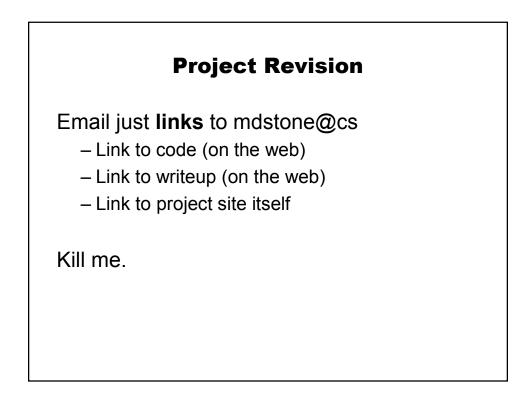


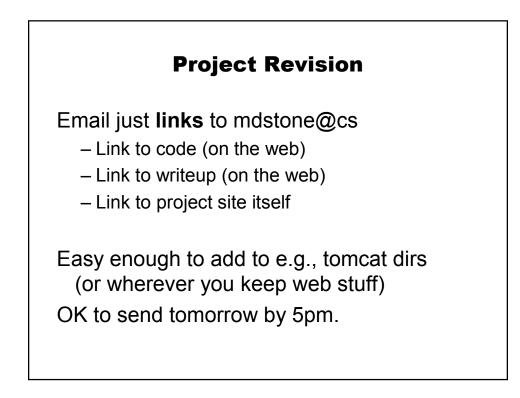
Matthew Stone

Project Revision

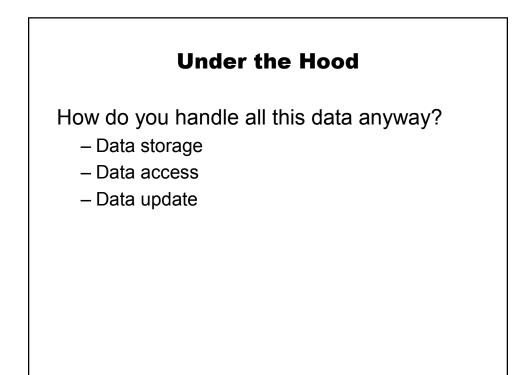
Email just links to mdstone@cs

- Link to code (on the web)
- Link to writeup (on the web)
- Link to project site itself

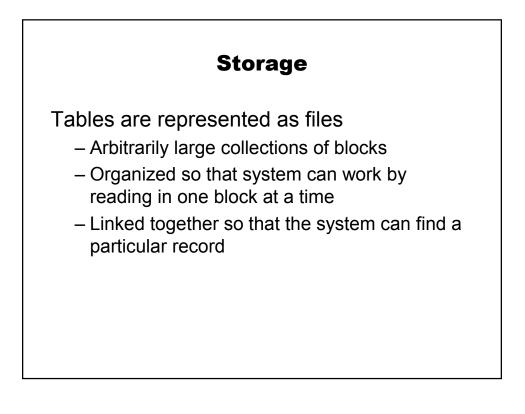




<section-header><section-header><section-header><list-item><list-item><list-item><list-item>



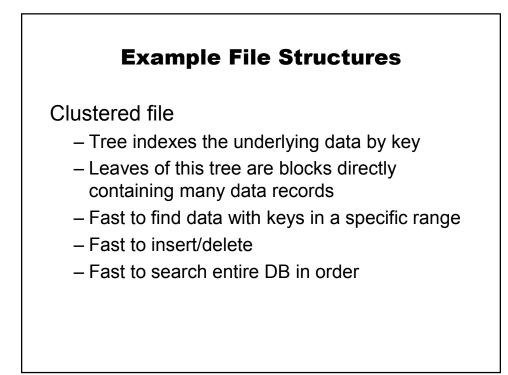
Storage Data is broken up into blocks on disk – Basic units of input/output – Few kilobytes – Stores a few records or a little indexing informaiton

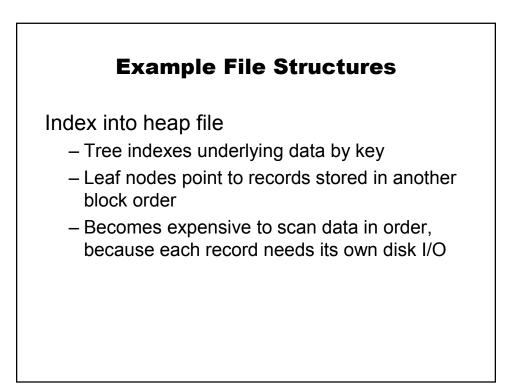


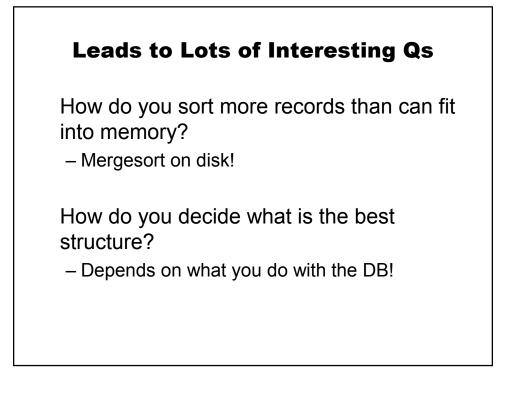
Example File Structures

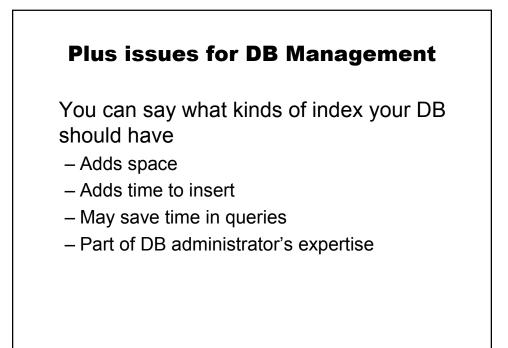
Heap Files

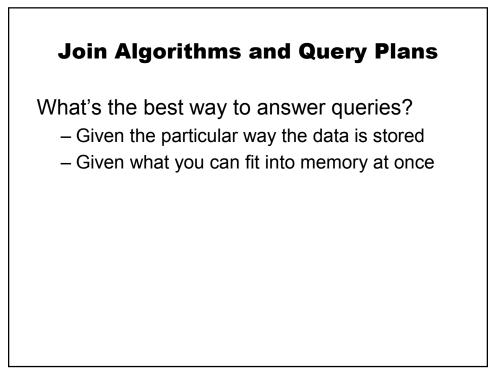
- Records are added in order to the collection of blocks
- Fast to insert an item
- Slow to search for an item

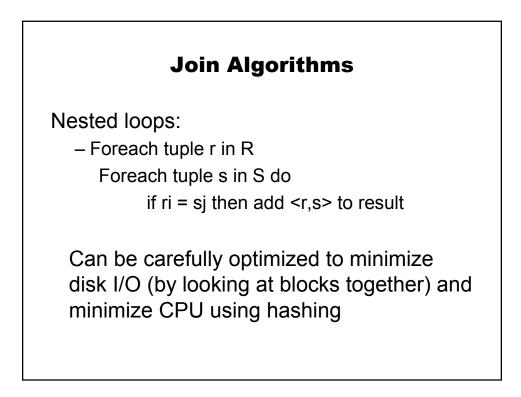


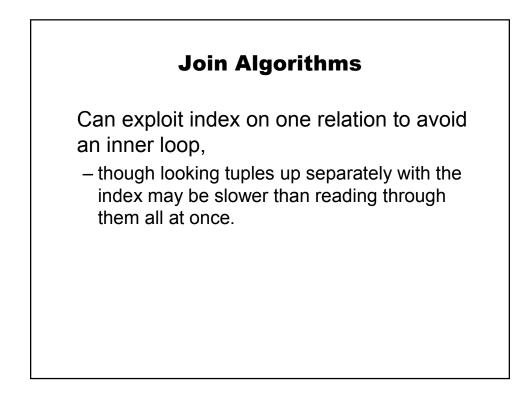












Join Algorithms

Can just sort based on the join key and merge the resulting lists

Works really well if tuples in the relation correspond one-to-one.

Join Algorithms

What's the right strategy?

- It depends
- Optimizers have many different join algorithms at their disposal
- They pick the one they think will make the overall query take the least time.

Leads to lots of interesting issues

Modeling performance Searching for the best query plans

Making Transactions Work

Defining protocols for access

- 2 phase locking
- To read an object, T requests a shared lock
- To write an object, T requests an exclusive lock on it.
- T's locks are released when it is completed.

Making Transactions Work

Tradeoff between safety and throughput

- 2PL asks for many more locks than "strictly" needed in many applications
- Concurrency may be too important

Interesting Questions

Isolation levels:

- Modeling kinds of transaction errors you're willing to tolerate
- The need for speed!

Putting locks into our crazy indexed file reps.

The Point

Lots of cool hairy stuff

– If you like that sort of thing.

Whirlwind tour of Chapters 8, 14, 17.

Final Review

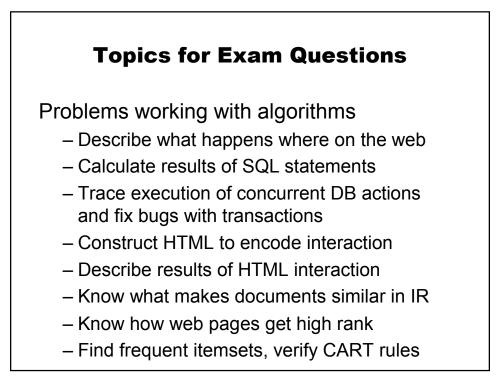
Topics for exam questions Principles and techniques Big Picture

These lists are not exhaustive – but they do prioritize what's most important.

Topics for Exam Questions

Problems in representation

- Draw E-R diagrams
- Formalize networks of concepts
- Design or use relational schemas
- Design semi-structured data representations
- Give information in relational tables
- Represent information in XML, XHTML
- Formulate SQL queries
- Describe XML nodes and paths



Topics for Exam Questions

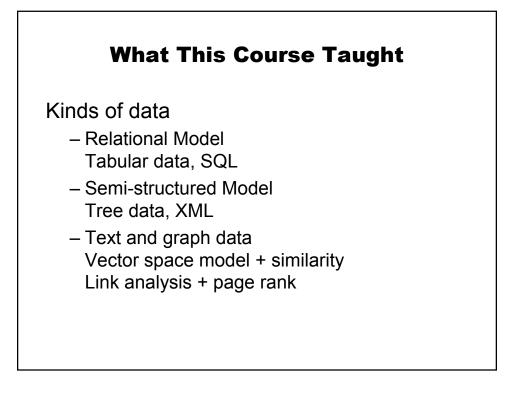
Problems in applying course principles

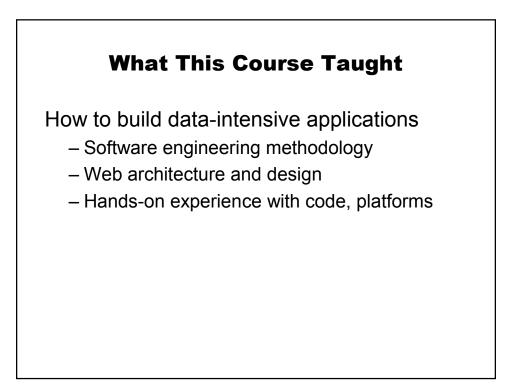
- Characterize system as information manager
- Identify information needed for a problem
- Choose the right technology
- Relate some data, a model and reality
- Fit information into an application/business
- Design a web interaction

What This Course Taught

Interacting with information as CS

- Getting information from the world
- Storing information
- Querying information
- Visualizing and using the results





Big Picture

Useful practical skills

 I see flyers everywhere recruiting ACCESS/PHP/XHTML hackers

Empowerment