

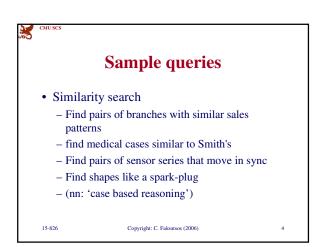


Problem

Given a large collection of (multimedia) records, find similar/interesting things, ie:

• Allow fast, approximate queries, and

• Find rules/patterns



Sample queries –cont'd

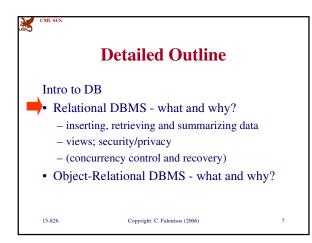
• Rule discovery

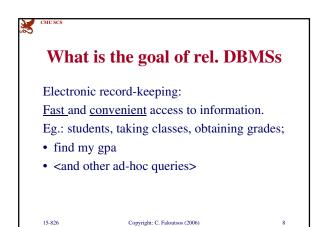
– Clusters (of branches; of sensor data; ...)

– Forecasting (total sales for next year?)

– Outliers (eg., unexpected part failures; fraud detection)



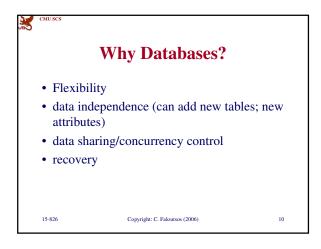




Why Databases?

Copyright: C. Faloutsos (2006)

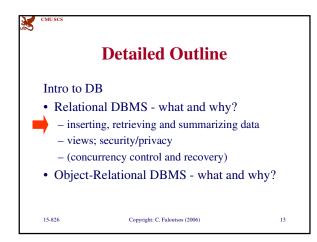
9

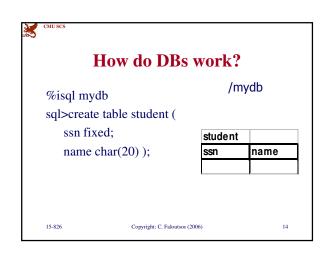


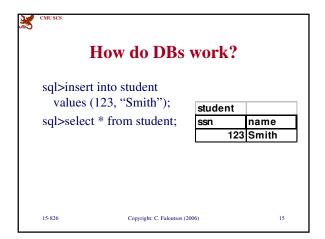
Why NOT Databases?

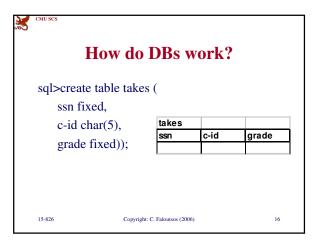
• Price
• additional expertise (SQL/DBA)
• over-kill for small data sets

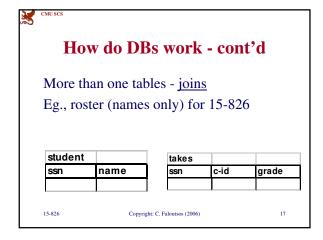


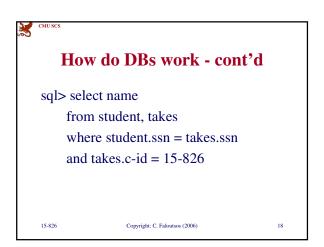


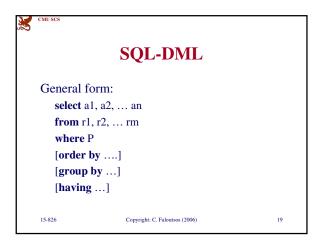


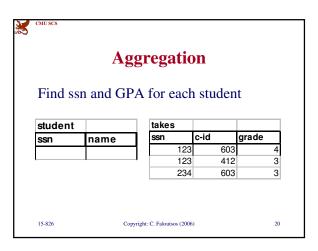


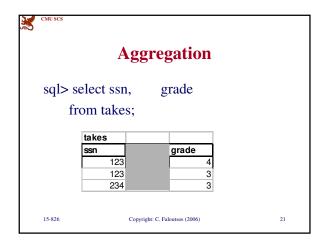


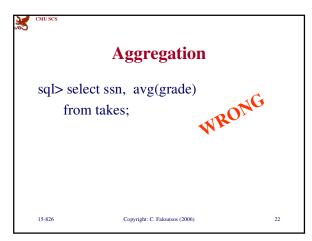


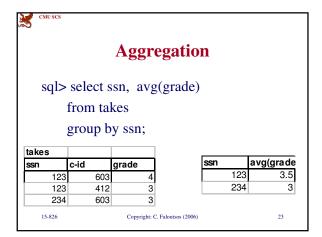


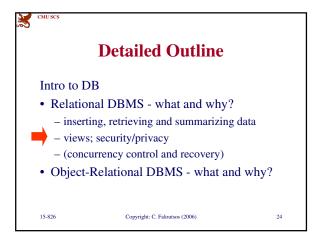


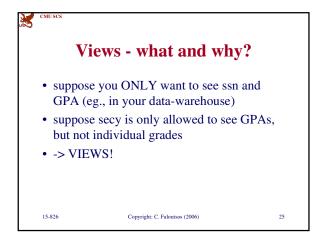


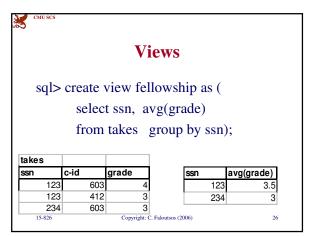


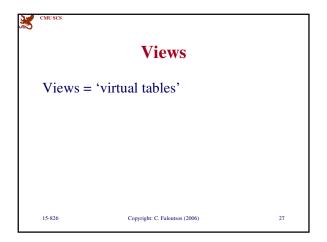


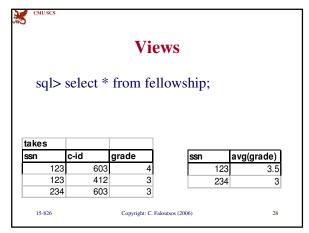


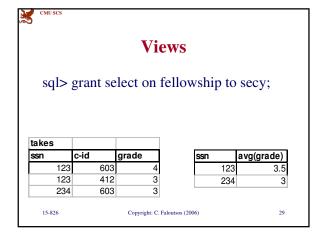


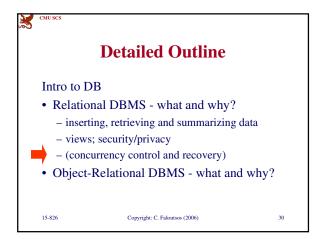


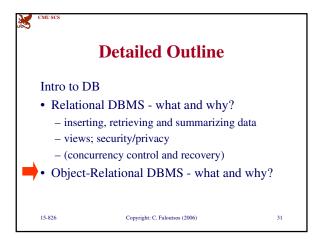


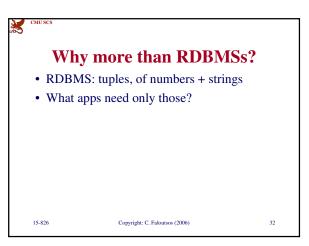












Why more than RDBMSs?

• RDBMS: tuples, of numbers + strings

• What apps need only those?

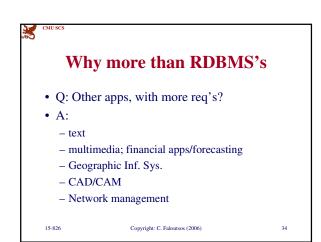
- Banks

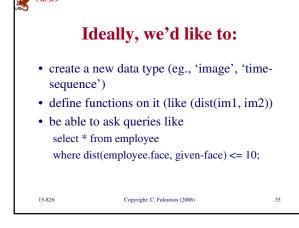
- Airlines

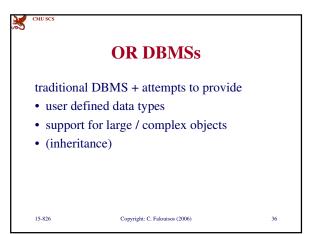
- Retailer stores

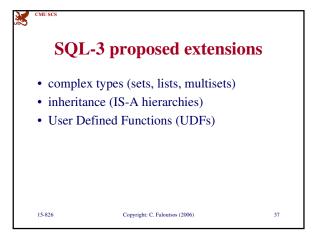
- ...

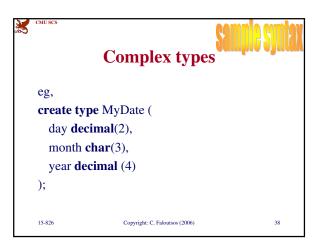
• Q: Other apps, with more req's?











BLObs etc:

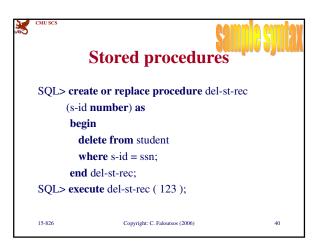
• Large objects, eg., video, images, 3d-MRI scans

• new data types: LOB (=Large OBject)

- BLOB: (up to 4Gb; binary: jpeg, mpeg, ...)

- CLOB: (up to 2Gb; character: english text)

- NCLOB:(.....; multi-byte characters)



Conclusions

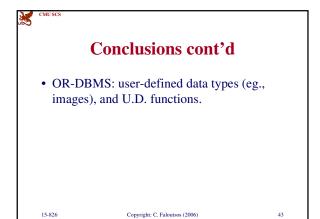
• (relational) DBMSs: electronic record keepers

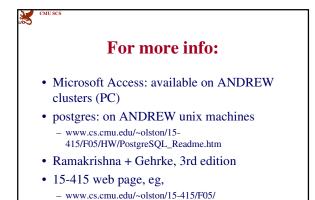
• customize them with create table commands

• ask SQL queries to retrieve info



C. Faloutsos 15-826





Copyright: C. Faloutsos (2006)

8