Fundamentals/ICY: Databases 2013/14

Week 3: Monday
Intro to Some Main Themes

John Barnden
Professor of Artificial Intelligence

School of Computer Science
University of Birmingham, UK

Answer Notes Soon

- **Week 1 exercise answer notes: will be on the web site by Tuesday, via Exercises page**

New Exercises!

- **Extensive *Additional Notes* on starting with SQL**
- **They include the Week 3 *Exercises* (SQL).**
- **On the web site, via Exercises page.**
- **You have TWO weeks: hand in by Mon 28 Oct.**

Let's Think about Some Data Repository Needs in Ordinary Life ...

- What sorts of data are held in things like the following, and how is it structured? What are the needs for extraction of info, updating data, etc.? How frequent and how easy are such operations?
  - **Address books (in physical form).**
  - **Diaries (in physical form).**
  - **Train timetables.**
  - **Cookery books.**
  - **Photo albums.**
  - **Dictionaries (monolingual and bilingual).**
  - **Records of a student society, voluntary society, etc.**
  - **Atlases.**
  - **Encyclopaedias.**
- THINK ABOUT: What about the data in your computer files?!

SOME GENERAL THEMES

Data redundancy, data anomalies (inconsistencies).

Cross-references between places in a data repository.

Associative linking ***versus*** pointing.

Use of tables to represent things.

The mathematical underpinnings of the tables.

DATA REDUNDANCY, DATA ANOMALIES , and CROSS-REFERENCING

(introduction to themes)

Data Redundancy and Anomalies

- Data Redundancy = replicating data in different places in a data repository.
  - ***E.g., in a recipe book, saying how to fry onions every time fried onions are needed in a recipe.***
- Redundancy encourages "data anomalies" and lack of "integrity" – basically, ***inconsistency*** between the different places.

  Such problems arise with insertions, deletions, and modifications in general.
- Redundancy also causes a type of ***inefficiency***: replicated updates.

Redundancy, etc., contd. 1

- Redundancy implies that if you want to modify/delete a piece of information, you need to
  - ***know whether*** there is replication, or ***check*** for possible replications
  - go to the ***effort*** of repeating changes when the item is replicated
  - ***avoid errors*** in such repeated changes.