DATABASE DESIGN

CS2141 - Software Development using C/C++

Goals of Design

- Don’t require on-the-fly modification of structure
- Make database structure easy to understand
- All records should be unique & unambiguous
- Allow the widest range of useful queries possible
- Eliminate repeated fields
- Minimize redundant storage of data
- Only one sort of record per table

No On-The-Fly Modification

- Changing database structure might break existing queries
- All existing records must be updated
  - Potentially huge performance hit during update
  - Might destroy your data
- NOTE: If you're still developing the database, changes are fine

Easy To Understand Structure

- Someone will have to maintain the database eventually
- Provide useful column names
- Avoid weird acronyms and abbreviations when possible
- Avoid using spaces and escaped characters in names
- Pick a consistent naming standard and stick with it

A Simple Table

Activities

Student	Activity1	Cost1	Activity2	Cost2
John Smith	Tennis	$36	Swimming	$17
Jane Bloggs	Squash	$40	Swimming	$17
John Smith	Tennis	$36		
Mark Antony	Swimming	$15	Golf	$47

Prevent Ambiguity

Students

StudentID	Student
84	John Smith
100	Jane Bloggs
182	John Smith
219	Mark Antony

Activities

StudentID	Activity1	Cost1	Activity2	Cost2
84	Tennis	$36	Swimming	$17
100	Squash	$40	Swimming	$17
182	Tennis	$36		
219	Swimming	$15	Golf	$47

Maximizing Useful Queries

Students

StudentID	FirstName	LastName
84	John	Smith
100	Jane	Bloggs
182	John	Smith
219	Mark	Antony

Activities

StudentID	Activity1	Cost1	Activity2	Cost2
84	Tennis	$36	Swimming	$17
100	Squash	$40	Swimming	$17
182	Tennis	$36		
219	Swimming	$15	Golf	$47

Eliminate Repeated Fields

Students

StudentID	FirstName	LastName
84	John	Smith
100	Jane	Bloggs
182	John	Smith
219	Mark	Antony

Activities

StudentID	Activity	Cost
84	Tennis	$36
84	Swimming	$17
100	Squash	$40
100	Swimming	$17
182	Tennis	$36
219	Swimming	$15
219	Golf	$47

Minimize Redundant Storage

Students

StudentID	FirstName	LastName
84	John	Smith
100	Jane	Bloggs
182	John	Smith
219	Mark	Antony

Participants

StudentID	ActivityID
84	3
84	15
100	18
100	15
182	3
219	15
219	22

Activities

ActivityID	Activity	Cost
3	Tennis	$36
15	Swimming	$17
18	Squash	$40
22	Golf	$47

Data Types

- Each column has an associated data type
- Most database engines enforce types
- SQLite supports five types (but does not enforce them):
  - NULL, INTEGER, REAL, TEXT, BLOB
- Dates/times may be stored as text, reals, or integers

Primary Keys

- Mechanism to guarantee a record is uniquely identifiable
- One primary key per table
- Could be any type, but integers are common
- Enforces uniqueness of key on insert
- INTEGER PRIMARY KEY will supply a value if not specified

SQL Syntax

- CREATE TABLE [IF NOT EXISTS] table_name (column_name type[, column_name type]);
- eg:

```
CREATE TABLE Students
(StudentID INTEGER PRIMARY KEY,
 FirstName TEXT,
 LastName TEXT);
```