


Announcements

- Grace Hopper Week ~ Dec. 6 – 12

Whereas Grace Murray Hopper, one of the first females in the field of computer science, engineered new programming languages (COBOL) and pioneered standards for computer systems which laid the foundation for many advancements in computer science; and Whereas the week of December 7, in honor of Grace Hopper's birthday, is designated as "National Computer Science Education Week


It's followed by some appropriate 'be it resolved' points about encouraging education and opportunities for females and underrepresented minorities in the field.



D.A. Clements, UW Information School 1

Rear Admiral Grace Hopper

- Computing pioneer (1906-1992)
- Invented the compiler and COBOL



12/7/2009

Announcements

- This week
 - Lab 11 and Project 3A due Wednesday at 10pm
 - Best GoPosters survey due Friday at 10pm
- Finals week
 - Lab 12 and Project 3B due Wednesday, Dec. 16 at 10pm

D.A. Clements, UW Information School 3

Announcements

- No final for this class

D.A. Clements, UW Information School 4

Announcements

- Read ch 17 for today
- Read the second half of ch 13 for Lab 12

D.A. Clements, UW Information School 5

Announcements

- Quiz this week in lab
 - Wednesday and Thursday
 - Topics
 - Spreadsheets
 - XML
 - Privacy
 - Cookies
 - Identity Theft

D.A. Clements, UW Information School 6

Announcements

- This week's GoPost discussion is posted

D.A. Clements, UW Information School 7

Just a thought...

- *Dijkstra: Whether a computer can think is about as interesting as whether a submarine can swim.*

D.A. Clements, UW Information School 8

Project 3

- Part A
 - Build database tables
 - Brief sample data
 - Answer queries
 - Get information out of the database


D.A. Clements, UW Information School 9

Project 3

- Part B
 - Imagine a year has gone by. . . .
 - The database has grown
 - *Download* the much-larger database
 - Boat Club members had some accidents


D.A. Clements, UW Information School 10

Project 3



Announcements

- Project 3B
 - You're called in to extend the database to make sure sailors are qualified to run the boats they rent
 - Build the queries
 - Answer the questions
 - Save database and queries, and upload to Collect It
 - Submit WebQ



Important!

- querySailorAge
 - current age: `DateDiff("yyyy",[tableSailors].[birthdate],Now())`

FIT 100—Fluency with Information Technology

A Table with a View

Primary keys, normalization, and SQL

D.A. Clements

D.A. Clements, UW Information School 14

Video

- Relational databases and tables

D.A. Clements, UW Information School 15

Video

- Primary Keys (5 min.)

D.A. Clements, UW Information School 16

Primary/Foreign Key

- Controlled redundancy:
 - Stores relationship between tables
 - Database tables share common attributes only to enable the tables to be linked
 - True redundancy exists only when there is unnecessary duplication of attribute values


D.A. Clements, UW Information School 17

Problem Fields (Don'ts)

Last Name	First Name	Full Name	City State Zip	Calculated Field		Multipart Field	Calculated Field	Multivalue Field
				Hourly	Weekly	Invoices		
Sullivan	Frank	Frank Sullivan	Kent, WA 98032	20.07	802.85	123		
Silby	Judy	Judy Silby	Yakima, WA 98902	16.73	669.04	127, 217, 319		
Harding	Joel	Joel Harding	Auburn, WA 98001	13.38	535.23	124, 297		
Rathke	Nicole	Nicole Rathke	Benton, WA 98005	9.37	374.66	116		
Lee	Allen	Allen Lee	Kent, WA 98032	16.73	669.04	151, 905		
Allert	Maria	Maria Allert	Yakima, WA 98902	8.03	321.14	143		
Young	Jim	Jim Young	Spokane, WA 99242	18.06	722.57	161, 181		


- Calculated field – can be computed by mathematical calculation or text concatenation
 - Waste of storage space (redundant),
 - No assurance the calculated value is updated when the user changes the input field(s)
- Multipart field – contains that should be two or more fields
 - Extra work when you want to analyze your data
- Multivalue field – multiple correct entries for the field
 - Create a separate subset table with each value in its own record.
- Derived field – contents of one or more fields absolutely predicts the contents of another
 - Should be dropped from the table

D.A. Clements, UW Information School 18

 **Video**


- Redundancy and Normalization
(5 min.)

D.A. Clements, UW Information School 19

 **Entities**


- Entity
 - Anything that can be identified by a fixed number of its characteristics (*attributes*)
- Attributes have
 - Names—field name, attribute, or column name
 - Values—the data stored in the table

16-20 D.A. Clements, UW Information School

 **Entities**

- An entity defines a table
 - Name of the entity is the name of the table
 - Each attribute of that entity
 - The column heading is the attribute name


16-21 D.A. Clements, UW Information School

 **Island table**

Island		
<i>Name</i>	<i>Area</i>	<i>Elevation</i>
Isabela	4588	1707
Fernandina	642	1494
Tower	14	76
Santa Cruz	986	846


Figure 16.4 A table instance for the island entity.

D.A. Clements, UW Information School 16-22

 **Properties of Entities**

- A relational database table can be empty
- Instances Are Unordered
 - Order of the rows and columns does not matter in databases
 - Freedom to move the data is limited to exchanging entire rows or exchanging entire columns

16-23 D.A. Clements, UW Information School

 **Properties of Entities (cont'd)**

- Uniqueness
 - No two rows can be the same
 - Two rows can have the same value for some attributes, just not all attributes

16-24 D.A. Clements, UW Information School

Properties Of Entities (cont'd)

- Atomic Data
 - Not decomposable into any smaller parts
 - Separate fields for street, city, state, postal code
 - "Only atomic data" rule relaxed for certain types of data
 - Dates, times, currency

16-25

D.A. Clements, UW Information School

Database schemas

- Database schema – way to define a table
 - Collection of table definitions that gives the name of the table, lists the attributes and their data types, and identifies the primary key

```

Island
iName      Text      Island Name
area       Number   Area in square kilometers
elevation  Number   Highest point on the island
Primary Key: iName
    
```

Figure 16.5 Database table definition for an Island table.

Database Tables Summary

- Tables in databases have a structure that is specified by metadata
- The structure is separate from its content
- A table structures a set of entities
 - Things that we can tell apart by their attributes
- The entities of the table are represented as rows
 - Rows and columns are unordered
- Tables and fields should have names that describe their contents
 - Fields must be atomic (indivisible)
 - One of more attributes define the primary key

16-27

D.A. Clements, UW Information School

TABLE OPERATIONS

D.A. Clements, UW Information School

28

Operations on Tables

- A database is a collection of tables
- Main use of database is to look up information
 - Users specify what they want to know and the database software finds it
- We can perform operations on tables to produce new tables
- The questions we ask of a database are answered with a whole new table, or view

16-29

D.A. Clements, UW Information School

```

Nations
Name      text      Common rather than official name
Domain    text      Internet top-level domain name
Capital   text      Nation's capital
Latitude  number   Approx. latitude of capital
NS        Boolean   Latitude is N(orth) or S(outh)
Longitude number   Approx. longitude of capital
E_W       Boolean   Longitude is E(ast) or W(est)
Interest  text      A short description of the country
Primary Key: Name

Name      Dom      Capital   Lat  NS  Lon  EW  Interest
Ireland   IE       Dublin    52  N   7   W   History
Israel    IR       Jerusalem 32  N   35  E   History
Italy     IT       Rome      42  N   12  E   Art
Jamaica   JM       Kingston  18  N   77  W   Beach
Japan     JP       Tokyo     35  N   143 E   Kabuki
    
```

Figure 16.6 The Nations table definition and sample entries.

16-30

D.A. Clements, UW Information School

SELECT Operation

- Takes rows from one table to create a new table
 - Specify the table from which rows are to be taken, and the *test* for selection
Syntax: `SELECT Test FROM Table`
 - Test is applied to each rows of the table to determine if it should be included in result table
 - Test uses attribute names, constants, and relational operators
 - If the test is true for a given row, the row is included in the result table; otherwise it is ignored

```
SELECT Interest='Beach' FROM Nations
```

16-31 D.A. Clements, UW Information School

Name	Dom	Capital	Lat	NS	Lon	EW	Interest
Australia	AU	Canberra	37	S	148	E	Beach
Bahamas	BS	Nassau	25	N	78	W	Beach
Barbados	BB	Bridgetown	13	N	59	W	Beach
Belize	BZ	Belmopan	17	N	89	W	Beach
Bermuda	BM	Hamilton	32	N	64	W	Beach

Figure 16.7 Part of the table created by selecting countries with a Test for Interest equal to Beach.

16-32 D.A. Clements, UW Information School

Animation

- A natural join

D.A. Clements, UW Information School 33

Physical and Logical Database

TABLES AND VIEWS

D.A. Clements, UW Information School 34

Structure of a Database

- Physical database and logical database
 - Physical database is the files, records in any order, no logical organization other than tables
 - Logical database is a view of database that shows only the rows and fields needed by the users
 - Solves Information Overload:
 - Users see only what they need
 - Users see only what they have permission to see

16-35 D.A. Clements, UW Information School

Physical vs. Logical

Figure 16.15 Structure of a database system. The physical database is the permanent repository of the data; the logical database, or view of the database, is the form of the database the users see. The transformation is implemented by the query processor, and is based on queries that define the logical database tables from the physical database tables.

D.A. Clements, UW Information School 16-36

Physical Database

- Designed by database administrators
 - Fast to access
 - No redundancy/duplicating information
 - Multiple data can lead to inconsistent data
 - Backup copies in case of accidental data deletion or disk crash

16-37 D.A. Clements, UW Information School

Logical Database

- Creating specialized views of the data for different users' needs
 - Creating a new "result set" from the current data each time
 - Fresh
 - Accurate

16-38 D.A. Clements, UW Information School

Defining Physical Tables

- Database schemes (schema)
 - Metadata specification that describes the database design

16-39 D.A. Clements, UW Information School

Figure 16.16 Table declarations from Microsoft Access 2007: (a) Home_Base table declaration shown in the design view; and (b) students table declaration. Notice that the key is specified by the tiny key next to Student_ID in the first column.

16-40 D.A. Clements, UW Information School

The Idea of Relationship

- A **relationship** is a correspondence between rows of one table and the rows of another table
 - Because the key Student_ID is used in each table, can not only find the address for each student (*Lives_At*), but can also find the student for each address (*Home_Of*)
- Relationship examples

16-41 D.A. Clements, UW Information School

Figure 16.17 The Relationships window from the Microsoft Access database system; the 1-to-1 *Lives_At* and *Home_Of* relationships are shown between Home_Base and Students.

16-42 D.A. Clements, UW Information School



Defining Logical Tables

- Constructing a View Using Join

- Match on the common field of Student_ID

```
Master_List = Student JOIN Home_Base  
On Student.Student_ID = Home_Base.Student_ID
```

```
Student_ID  
First_Name  
Middle_Name  
Last_Name  
Birthdate  
On_Probation  
Street_Address  
City  
State  
Country  
Postal_Code
```

16-43

Figure 16.18 Attributes of the Master_List table. Being created from Student and Home_Base tables. Student_ID from the component tables.