# Sets and Set Theory

CS1F

IM Lecture 7

Craig Macdonald

# Overview



This lecture

- Sets and Set Theory
- Relations and the Cartesian Product

Next lecture (lecture 8)

- Relational Algebra
  - The foundations for SQL

## Where to go for more info…

3

- Rosen, Discrete Mathematics and Its Applications
  - sets - sections 1.4 & 1.5
  - relations - sections 6.1 & 6.2
  - http://www.mhhe.com/math/advmath/rosen/
- Rolland, section 3.3

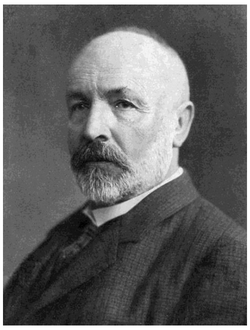
## Sets and Set Theory

4

**Set theory** is the branch of mathematics that studies **sets**

**Sets** are collections of objects

- The set of all numbers
- the set of all animals with tails
- the set of letters in the alphabet
- the set of all students in a class
- …………………

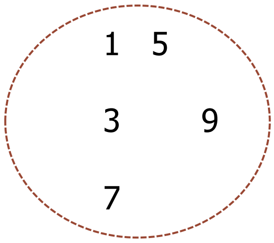


Georg Cantor (1845-1918)

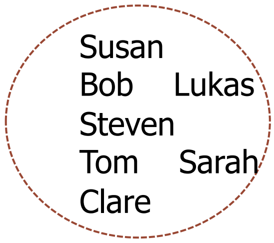
# Sets

5

Often all members of a set have similar properties




Odd numbers
less than 10




Students in a
Tutorial Group

# Set Theory - Vocabulary

6

- Objects in a set are called ‘***elements***’ or ‘***members***’ of a set
- A set is said to ‘contain’ its elements

- In databases
  - all exam scores make up a '***set***' of exam scores.
  - all employees of a company make up a ‘***set***’ of employees

## Describing Sets

7

- Describing a set
  - List all the members between braces
    - E.g. **{a, b, c, d}**
    - Represents the set with the four elements a, b, c, and d.

## Describing Sets

8

- E.g. The set V of all vowels in the English alphabet
- E.g. The set O of positive integers less than 10

## Describing Sets

9

- E.g. The set V of all vowels in the English alphabet
  - **V = {a, e, i, o, u}**
- E.g. The set O of positive integers less than 10
  - **O = {1, 3, 5, 7, 9 }**
- | | denotes the *cardinality* of a set
  - **|V| = 5, |O| = 5**

## Set Equality

10

- Two sets are ***equal*** if and only if they have the same elements
  - **Order doesn't matter**
    - {1,3,5} = {1,5,3} = {3,1,5} = {3,5,1}= {5,1,3} ={5,3,1}
  - **Repetition doesn't matter**
    - {1,2} = {1,1,2} = {1,2,2,2,2}

## Set Equality

11

A = {1,2,3}

B = {3,2,1}  C = {1,1,2,2,2,3}  D= {1,2,3}

**Which set(s) are equal to A?**

## Set Equality

12

A = {1,2,3}

B = {3,2,1}  C = {1,1,2,2,2,3}  D= {1,2,3}

**A = B, C and D**

## Sets

13

- Sets *usually* group together elements with associated properties
  - but seemingly unrelated properties can also be listed as a set
  - {2, e, Fred, Paris} is also a set
    - We just don't know much about exactly how they are related to each other

## Predicates and Sets

14

- It is sometimes inconvenient or impossible to describe a set by listing all of its elements
  - What is the set of all integers less than 1 million?

## Predicates and Sets

15

- It is sometimes inconvenient or impossible to describe a set by listing all of its elements
  - What is the set of all integers less than 1 million?
    - {1,2,3,4,5......................................!!!!!!!}

## Set Builder Notation

16

- Characterise all those elements in the set by stating the properties they must have to be members

E.g.

- The set O of all positive integers less than 10 in set builder notation is:
  - O = {X | X is an odd integer less than 10}
- More mathematical definitions are also OK:
  - $O = \{X \mid X \in \mathbb{N} \land x < 10 \land x \% 2 == 1 \}$

## Predicates and Sets

17

- A **predicate** is sometimes used to indicate **set membership**
- A predicate *F(x)* will be true or false, depending on whether *x* belongs to a set

## Predicates and set membership

18

An example

{*x* | *x* is a positive integer less than 4}

is the set {1,2,3}

If *t* is an element of the set {*x* | *F*(*x*)}

then the statement *F*(*t*) is *true*

So if F(x) is defined as x % 2 = 0

{x|F (x)} contains…. the set of all even numbers

Here, *F*(*x*) is referred to as the ***predicate***, and *x* the *subject* of the *proposition*

## Some Notation

20

- **$a \in A$**
  - a is an element of set A
- **$a \notin A$**
  - a is not an element of set A
- **$\emptyset$**
  - The empty or null set
  - Also represented by { }

## Graphical representation of sets

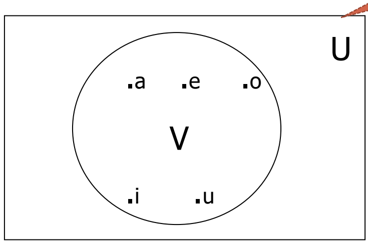
21

- Sets can be represented graphically using ***Venn diagrams***
- The ***universal set U*** (which contains all of the objects under consideration) is represented by a rectangle
- Inside the rectangle, circles are used to represent sets
- Sometimes points are used to represent the particular elements of the set

# A Example Set

22

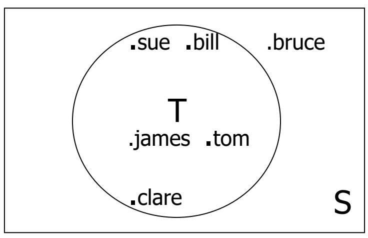
U is the Universe – the set of all possible items




The set V of vowels from all letters U

# A Example Set

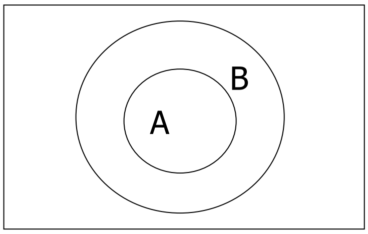
23




The set T of people in tutorial group from all Students S

# Subsets

24



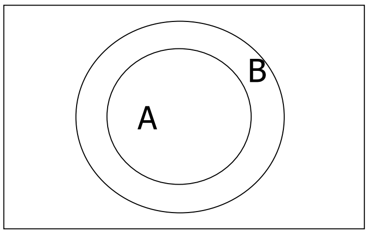

A is a subset of B

$A \subset B$

A test that returns true iff $A \subset B$

# Subsets (2)

25




A is a subset or equal to B

$A \subseteq B$

A test that returns true iff $A \subseteq B$

## The Power Set

26

- Given a set S, the ***power set*** is the set of all subsets of the set S
  - Denoted by P(S) or $\mathbb{P}(S)$
- E.g. the power set of {0,1,2} is

## The Power Set

27

- E.g. the power set of {0,1,2} is
  - P({0,1,2}) = {Ø, {0}, {1}, {2}, {0,1}, {0,2}, {1,2}, {0,1,2} }
  - NB - the **empty set** and the **set itself** are members of this set of subsets
- If a set has n elements, its power set has $2^n$ elements
- The power set does not contain numbers, it contains SETs

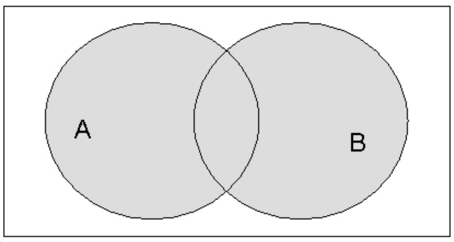
# Set Operations

28

- Two sets can be combined in many different ways
  - The following illustrates some such combinations
    - See Rosen, 1.5 for further explanations

# Union

29




Symbol like Union

The union of A and B

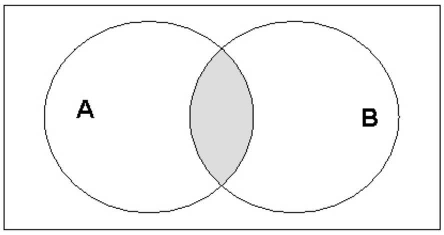
$A \cup B$

$A \cup B = \{x \mid x \in A \vee x \in B\}$

The set that contains those elements that are either in A, B, or in both

## Intersection

30




The **intersection** of A and B

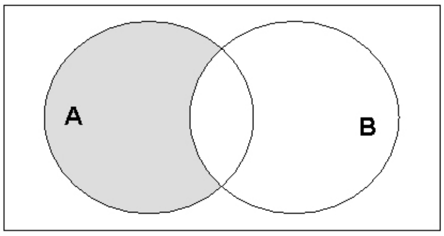
$A \cap B$

Symbol like aNd

$A \cap B = \{x \mid x \in A \wedge x \in B\}$

## Difference

31



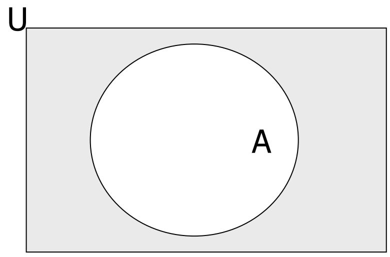

The **difference** of A and B

$A - B$

$A - B = \{x \mid x \in A \wedge x \notin B\}$

# Complement

32



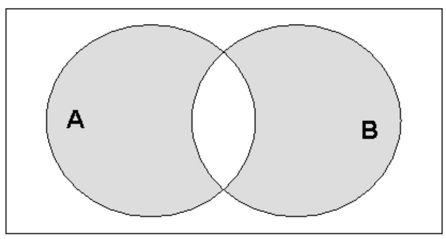

The complement of A

$\overline{A}$

$\overline{A} = \{x \mid x \notin A\}$

# Symmetric Difference

33




The symmetric difference of A and B

$A \oplus B = (A - B) \cup (B - A)$

$A \oplus B = \{x \mid (x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A)\}$

## Summary

34

- What are sets?
- Notation for making sets, comparing sets
- Operators: making new sets from other sets
  - **∪ union**
  - **∩ intersection**
  - **– difference**
  - **$\overline{A}$ complement**
  - **⊕ symmetric difference**

## From past exam(s)

35

Given the following sets:

A = {1,3,7,9}, B = {2,4,6}, C = {7,9}

determine the following (assume that P is the powerset operator):

- |B| =3
- P (B) = { {}, {2},{4},{6}, {2,4}, {4,6}, {2,4,6}, {2,6} }
- |P (C)| = 4
- A ∪ B = {1,2,3,4,6,7,9}
- A ∩ C = {7,9}
- Which of the following are true?
  - C ⊂ A = TRUE
  - C ⊂ B = FALSE

## Summary of Sets

36

- What are sets? {1,3,5}
- Set builder notation for making sets; comparing sets
- Operators: making new sets from other sets
  - ∪ **union**
  - ∩ **intersection**
  - – **difference**
  - $\overline{A}$ **complement**
  - ⊕ **symmetric difference**
- So how does this help with databases?

## So why is this useful?

37

Consider a query:

*"What are the grades of student 8187491?"*

How can we query the database to obtain this information?

## So why is this useful?

38

Two ways of querying a database:

- procedural (relational algebra, Pandas)
  - sequence of operations
  - the output of each operation is the input to the next operation

**result ← F4 (F2 (F1(tableA), tableB), F3(tableC))**

- declarative (SQL)
  - describes the desired results (in terms of conditions)
  - the DBMS works out the operations

**result ← CONDITIONS (tableA, tableB, tableC)**

This is based on set theory

This is internally implemented as RA operations

*RA is key to understanding SQL query processing!*

## A Relation instance

39

**STUDENT**

Schema

| name | matric | exam1 | exam2 |
|---|---|---|---|
| Gedge | 891023 | 12 | 58 |
| Kerr | 892361 | 66 | 90 |
| Fraser | 880123 | 50 | 65 |

- This Student relation instance has:
  - Degree 4 and Cardinality 3
- A **relation** is a **set** of **tuples** (or **n-tuples**)
  - As each tuple has n values (same as degree of the relation)
  - E.g. <Fraser,880123,50,65> is a 4-tuple.

# N-tuples are not sets

40

- The order of elements in a collection is sometimes important
  - But **sets are unordered**, so a different structure is needed
- This is provided by ***ordered n-tuples***
  - *<2,1,5> is an 3-tuple*

# N-tuples

41

- Two ordered n-tuples are equal if and only if each corresponding pair of their elements is equal
  - $<a_1,a_2,....a_n> = <b_1,b_2,.....b_n>$ if and only if $a_i=b_i$ for i=1,2,…n
  - {1, 3, 5} = {3, 1, 5} = TRUE for SETS
  - <1, 3, 5> = <3, 1, 5> = FALSE for N-TUPLES
- NB: We can use <> or () to denote tuples, but not {}

## Cartesian Product

42

Let A and B be sets

- The ***cartesian product*** of A and B (A X B) is
  - the set of all ***ordered pairs*** *(i.e. tuples)*

    <a,b> where a ∈ A and b ∈ B

A={0,1}, B={a,b,c}

A X B =

## Cartesian Product

43

Let A and B be sets

- The ***cartesian product*** of A and B (A X B) is
  - the set of all ***ordered pairs*** *(i.e. tuples)*

    <a,b> where a ∈ A and b ∈ B

A={0,1}, B={a,b,c}

A X B =
{<0,a>,<0,b>,<0,c>, <1,a>, <1,b> ,<1,c>}

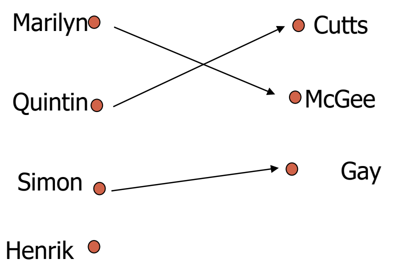
# Connecting to Databases: Representing a relation

44

Forename = {Marilyn, Quintin, Simon, Henrik}
Surname = {McGee, Cutts, Gay}

Domains

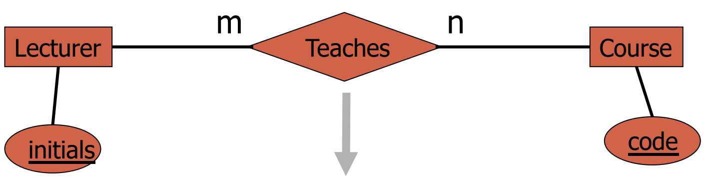
**Names = {<Marilyn,McGee>,<Quintin,Cutts>, <Simon, Gay>}**




| R | Cutts | McGee | Gay |
|---|---|---|---|
| Marilyn | | x | |
| Quintin | x | | |
| Simon | | | x |
| Henrik | | | |

# A relationship and its equivalent relation

45




Lecturer

| Initials |
|---|
| MMcG |
| QC |
| SGay |
| HRight |
| RInnis |

Teaches

| Lecturer | Course |
|---|---|
| MMcG | CS1Q |
| QC | CS1CT |
| SGay | CS1P |
| SGay | CS1Q |

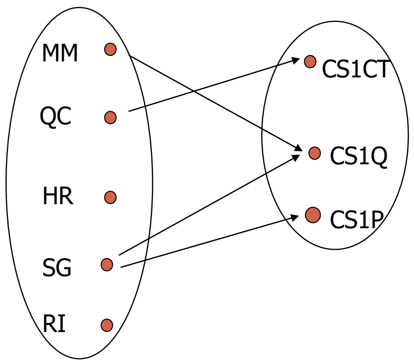
Course

| Code |
|---|
| CS1Q |
| CS1CT |
| CS1P |

# The table as a relation

46

Lecturers = {MM, QC, SG, HR, RI}
Courses = {CS1P, CS1Q, CS1CT}
**Teaches = {<MM, CS1Q>, <QC, CS1CT>, <SG, CS1P>, <SG,CS1Q>}**




| Teaches | CS1P | CS1Q | CS1CT |
|---|---|---|---|
| MM | | x | |
| QC | | | x |
| SG | X | X | |
| HR | | | |
| RI | | | |

# This Week

47

- A tutorial on Sets
- Next lecture:
  - Relations, and Relational Algebra
- You can also continue to work on your database **in your own time**…
  - In the lab, or from home/laptop
  - NEXT WEEK'S LAB – Populating DB with Data

# Where to go for more info...

48

- Rosen, Discrete Mathematics and Its Applications
  - Sets - sections 1.4 & 1.5
  - Relations - sections 6.1 & 6.2
  - see http://www.mhhe.com/math/advmath/rosen/