Database Management Systems

Madhavan Mukund

https://www.cmi.ac.in/~madhavan

Sai University

Lecture 2, 18 August 2023

Administrative issues

Grading

- Assignments (both pen-and-paper and coding), 50%
- Midsemester exam (week of Oct 3–6), 20%
- Final exam, 30%

Textbooks

- Abraham Silberschatz, Henry F. Korth and S. Sudarshan: *Database System Concepts (7th ed)*, McGraw-Hill (2021)
- Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom: *Database Systems: The Complete Book (2nd ed)*, Pearson (2013)

Course material

- Class notes and other supplementary material will be made available online (Moodle)

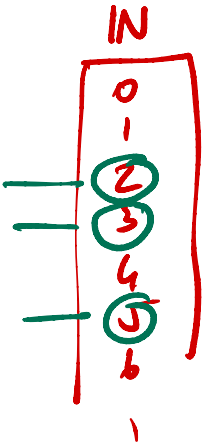
What is a relation?

- A subset of some set
- Natural numbers $\mathbb{N} = \{0, 1, \ldots\}$

$\text{Primes} \subseteq \mathbb{N}$

$17 \in \text{Primes}$ – in relation

$18 \notin \text{Primes}$ – not in relation



Cartesian Product of 2 sets

$A = \{a, b, c\}$

$Z = \{x, y\}$

Left & right

$(a, x) \in A \times Z$

$(x, a) \notin A \times Z$

$$A \times Z = \{(a,x), (a,y), (b,x), (b,y), (c,x), (c,y)\}$$

$\mathbb{N} \times \mathbb{N}$ - Cartesian product - Also a set

$= \{(i, j) \mid i \in \mathbb{N}, j \in \mathbb{N}\}$

$X \subseteq \mathbb{N} \times \mathbb{N}$ Pairs of the form (j, j^2)

$= \{(0,0), (1,1), (2,4), (3,9), \ldots\}$

$(15, 225) \in X$? ✓ $(17, 829) \notin X$

$\mathbb{N}$ - natural number, $\mathbb{R}$ - real number

$S \subseteq \mathbb{N} \times \mathbb{R}$

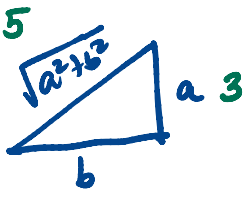
$(i, \sqrt{i})$

$\{(0,0), (1,1), (2, 1.414\ldots), (3, 1.732\ldots)\}$

$\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ $\{(i,j,k) \mid i \in \mathbb{N}, j \in \mathbb{N}, k \in \mathbb{N}\}$

Pythagorean triples

$3^2 + 4^2 = 5^2$



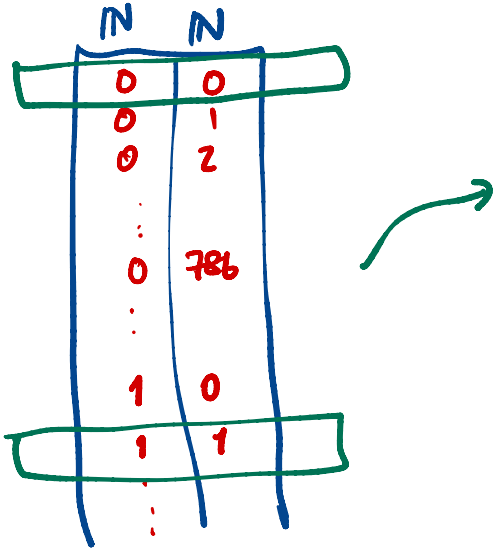
$= \{(3,4,5), (5,12,13), \ldots (6,8,10)$

25 144 169 36 64 100

The relational model

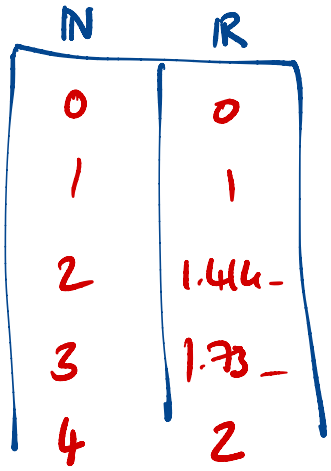
Relation

Square



The relational model

Square root



Tables

Name	DoB
Nikhil	16-04-2011
Ananya	28-03-2015

Name x DoB

Nikhil 28-03-2015

Ananya 16-04-2011

Nikhil 18-01-2042

DoBTable $\subseteq$ Names x DoB

The relational model

Mathematically, a relation can be infinite

Infinite set of pairs, triple, ...

"Real world" table – always finite

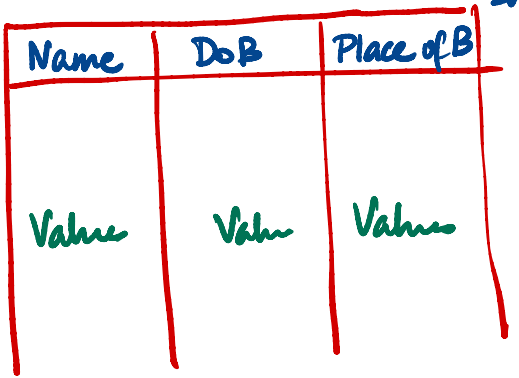
Underlying "universe" may be infinite

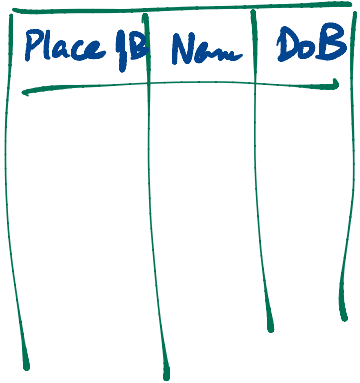
(x, y) is "not in" the table

The relational model

Table

Column Order does not matter since columns have "names"





The relational model

Terminology

Relations		Tables
pair, triple, quadruple, ⋮, n-tuple	Tuple	Rows
Attributes		Columns

When we create a database

- Fix the structure of each table – relation schema
  - column names + value types
- Populate each table – relation
  - always finite

University database

ID	*name*	*dept_name*	*salary*
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Instructor

dept_name	*building*	*budget*
Biology	Watson	90000
Comp. Sci.	Taylor	100000
Elec. Eng.	Taylor	85000
Finance	Painter	120000
History	Painter	50000
Music	Packard	80000
Physics	Watson	70000

Department

University database

course_id	*title*	*dept_name*	*credits*
BIO-101	Intro. to Biology	Biology	4
BIO-301	Genetics	Biology	4
BIO-399	Computational Biology	Biology	3
CS-101	Intro. to Computer Science	Comp. Sci.	4
CS-190	Game Design	Comp. Sci.	4
CS-315	Robotics	Comp. Sci.	3
CS-319	Image Processing	Comp. Sci.	3
CS-347	Database System Concepts	Comp. Sci.	3
EE-181	Intro. to Digital Systems	Elec. Eng.	3
FIN-201	Investment Banking	Finance	3
HIS-351	World History	History	3
MU-199	Music Video Production	Music	3
PHY-101	Physical Principles	Physics	4

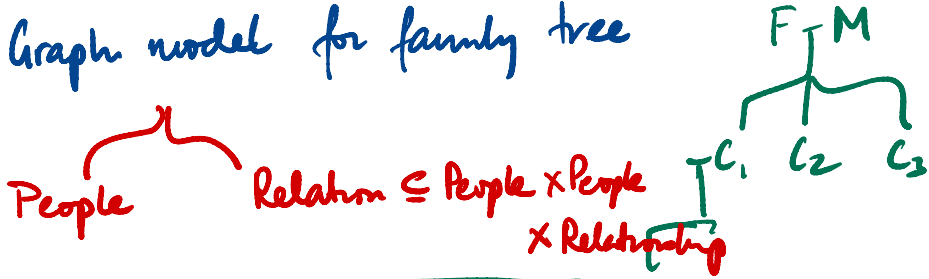
Course

course_id	*sec_id*	*semester*	*year*	*building*	*room_number*	*time_slot_id*
BIO-101	1	Summer	2017	Painter	514	B
BIO-301	1	Summer	2018	Painter	514	A
CS-101	1	Fall	2017	Packard	101	H
CS-101	1	Spring	2018	Packard	101	F
CS-190	1	Spring	2017	Taylor	3128	E
CS-190	2	Spring	2017	Taylor	3128	A
CS-315	1	Spring	2018	Watson	120	D
CS-319	1	Spring	2018	Watson	100	B
CS-319	2	Spring	2018	Taylor	3128	C
CS-347	1	Fall	2017	Taylor	3128	A
EE-181	1	Spring	2017	Taylor	3128	C
FIN-201	1	Spring	2018	Packard	101	B
HIS-351	1	Spring	2018	Painter	514	C
MU-199	1	Spring	2018	Packard	101	D
PHY-101	1	Fall	2017	Watson	100	A

Section

What about other kinds of data?

Graph model for family tree

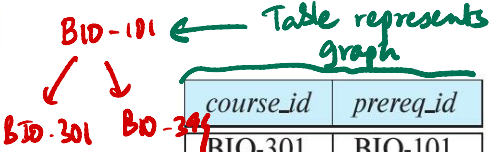


F	M	Spouse
F	C_1	Parent
C_1	M	Child

University database

ID	*course_id*	*sec_id*	*semester*	*year*
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teaches



course_id	*prereq_id*
BIO-301	BIO-101
BIO-399	BIO-101
CS-190	CS-101
CS-315	CS-101
CS-319	CS-101
CS-347	CS-101
EE-181	PHY-101

Prerequisites

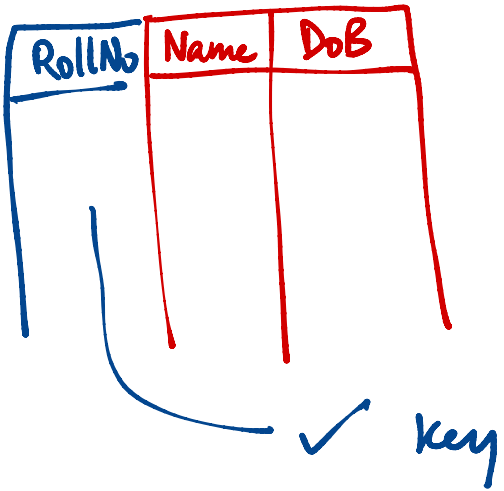
Keys and integrity

DBMS needs to execute "queries" across all possible databases

Along with schema, provide some hints/clues

Key – a column whose value uniquely identifies row/tuple

Keys and integrity



Name is not key

Has to be guaranteed unique in all instances of the table

Keys and integrity

Key may be more than one column

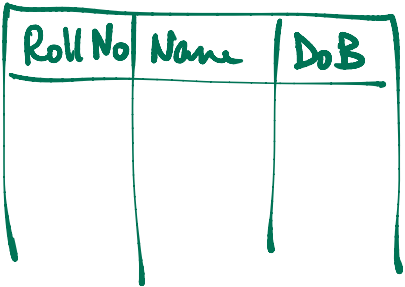
Phone company

Calls

From No.	To No.	Date	Start	End
92018673	22264121	18-08-23	16:02	16:04

"Key"

Keys and integrity



Roll No – key ✓

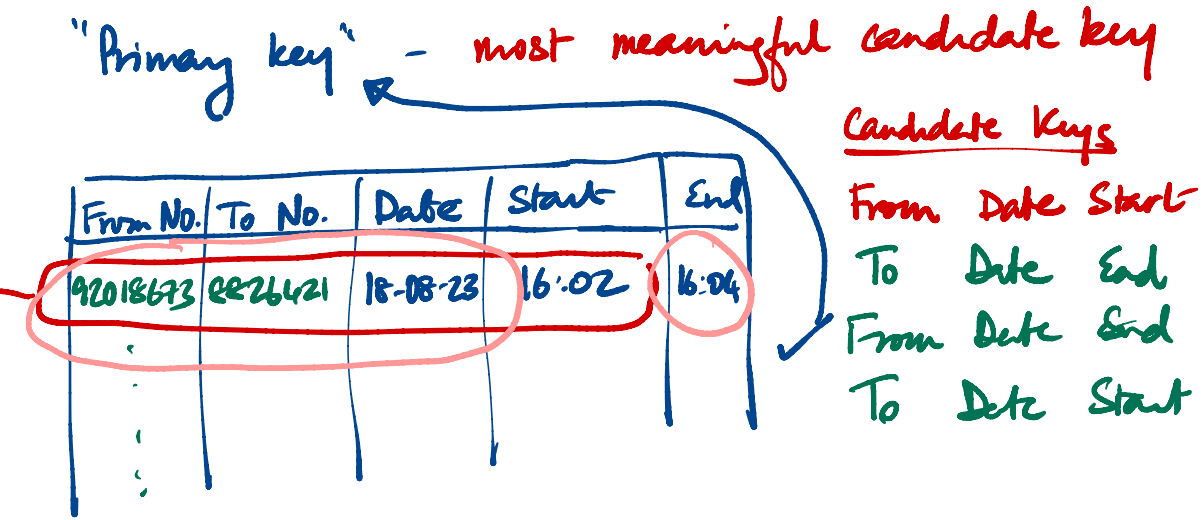
(Roll No, Name) – key

(Roll No, DoB) – key

"Superkey" – Any combination of cols that acts as a key

"Candidate Key" – Minimal combination

Keys and integrity



Keys and integrity

Course codes in Prerequisite table must exist in Course table

Constraint across tables

University database

ID	*course_id*	*sec_id*	*semester*	*year*
10101	CS-101	1	Fall	2017
10101	CS-315	1	Spring	2018
10101	CS-347	1	Fall	2017
12121	FIN-201	1	Spring	2018
15151	MU-199	1	Spring	2018
22222	PHY-101	1	Fall	2017
32343	HIS-351	1	Spring	2018
45565	CS-101	1	Spring	2018
45565	CS-319	1	Spring	2018
76766	BIO-101	1	Summer	2017
76766	BIO-301	1	Summer	2018
83821	CS-190	1	Spring	2017
83821	CS-190	2	Spring	2017
83821	CS-319	2	Spring	2018
98345	EE-181	1	Spring	2017

Teaches

Should exist in Course table

course_id	*prereq_id*
BIO-301	BIO-101
BIO-399	BIO-101
CS-190	CS-101
CS-315	CS-101
CS-319	CS-101
CS-347	CS-101
EE-181	PHY-101

Prerequisites