

Normalisation

Business Data Management and Analytics

Contributions by Arthur Adamopoulos and Vince Bruno

Database Lecture 8

1

1

Normalisation

- Normalisation [none, 1st, 2nd, 3rd forms]
- Functional Dependence
- Partial/Transitive dependencies
- Primary Keys
- Referential Integrity
- EERD

Database Lecture 8

2

2

Normalisation - Steps

- Each stage is a normal form
- Normal forms relate by applying simple rules about dependencies.

Database Lecture 8

3

3

Normalisation - steps

```

graph TD
    A[bunch fields] -- "remove repeated groups" --> B[1st]
    B -- "remove partial dependencies" --> C[2nd]
    C -- "remove transitive dependencies" --> D[3rd]
    D --> E[ ]
  
```

Database Lecture 8

4

4

Functional Dependence

- A relationship between two attributes
- One field is dependant on another, the first field value would not come into existence unless the second field value does.
- IF A depends on B there is only 1 value for A for each value of B
- A only exists when B exists first.
- $B \rightarrow A$

Database Lecture 8

5

5

Partial Dependency

- A field that depends on part of the primary or candidate key.
- If tables do NOT have multi-part keys, NO partial dependency can exist.

CONNECT

Tower Id +

Call Id ←

Orig Phone No —

Dest Phone No —

Seconds

Database Lecture 8

6

6

Transitive Dependency

- A field that depends on a NON-Key field(s).



7

Zero Normal Form



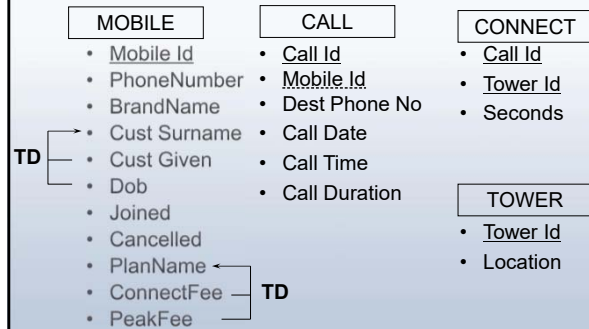
8

1st Normal Form



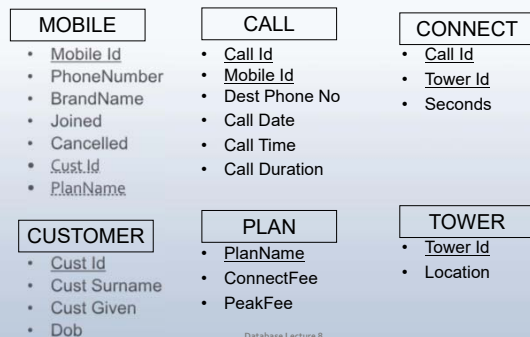
9

2nd Normal Form



10

3rd Normal Form



11

Other Normal Forms

- Boyce-Codd Normal Form
 - remove remaining anomalies resulting from functional dependencies;
- Fourth Normal Form
 - remove anomalies that result from a multi-valued dependencies;
- Fifth Normal Form
 - designed to cope with dependency known as join dependency.

12

Keys

- Primary Key
 - Unique identifier (field or fields) of a table
 - properties of a primary key are:
 - Uniqueness
 - Availability
 - Stability
 - Minimality
- Candidate Key
 - A field or fields that could be a primary key
- Composite Key
 - a primary key containing more than one field.

Database Lecture 8

13

13

Constraints - Domain

- input checking
 - type
 - length
 - formats
 - allowable values
 - min/max ranges
 - Optional/mandatory

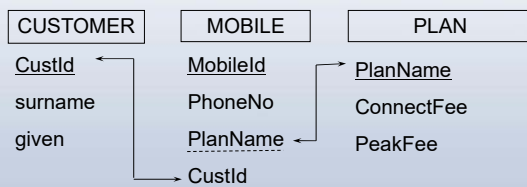
Database Lecture 8

14

14

Referential Integrity

- related to foreign keys only
- values in foreign key must exist in primary key of related file.



Database Lecture 8

15

15

RI - Other Issues

- Insert/Update
 - value inserted/change in foreign key must already exist in primary key of other file.
- Delete - three options:
 - not allow;
 - null out the corresponding foreign key(s);
 - cascade delete - remove entire record and any related foreign keys.

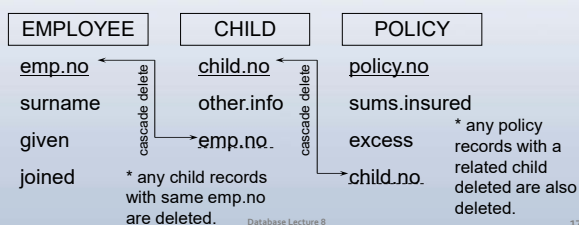
Database Lecture 8

16

16

RI - Cascade Delete

- A delete is issued to delete an employee from the EMPLOYEE file.



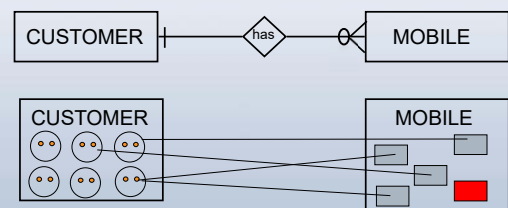
Database Lecture 8

17

17

Relationships

- One other aspect to relationships, which can further define the allowed associations between entities is:
 - Mandatory – at least one association must exist



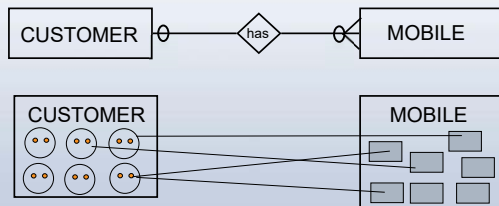
Database Lecture 8

18

18

Relationships

- Optional – zero or more association can exist



Database Lecture 8

19

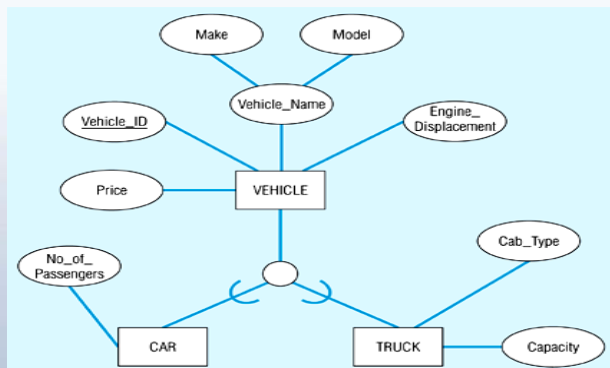
Enhanced ER Model (EER)

- Super and Sub types
- The ER model has been enhanced by various people to include inheritance.
- Concepts have simply been borrowed from the Object Oriented model but drawn differently.
- Both supertypes and subtypes can participate in relationships.
- Page 153 in the Hoffer (textbook) describes the various ways this is represented.

Database Lecture 8

20

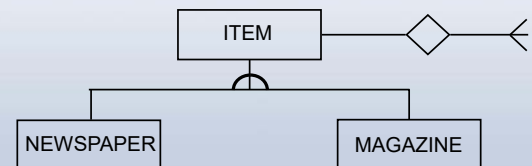
EER Model Super/Sub Types



21

EER: Example PEOPLE

- An ITEM delivered by newsagent has a primary key of ITEM No and each of the sub-types have their own primary key.



Database Lecture 8

22

EER: Example PEOPLE



Database Lecture 8

23

EER Model Super/Sub Types

- Three options to translate super/sub types into a relational model
 - Keep all levels
 - Roll Up
 - Roll Down
- Hoffer (text) describes keep all level transformation, page 224-225

Database Lecture 8

24

EER: Example PEOPLE

- Keep All Levels

ITEM	NEWSPAPER	MAGAZINE
<u>item no+</u>	<u>newspaper no</u>	<u>magazine no</u>
type	Field1	Field3
Name	field2	field4
frequency		
price		

Database Lecture 8

25

25

EER: Example PEOPLE

- Roll up

ITEM
<u>item no</u>
type
Name
Frequency
Price
Field1-4

Database Lecture 8

26

26

EER: Example PEOPLE

- Roll down

NEWSPAPER	MAGAZINE
<u>newspaper no</u>	<u>magazine no</u>
Name	Name
Frequency	Frequency
Price	Price
Field1	Field3
field2	field4

Database Lecture 8

27

27

Exercise 1: Normalise

- What is wrong with the following set of tables, designed for the Kangaroo Holiday Park ? How would you redesign them ?

CABINS	TOURIST
<u>Cabin No</u>	<u>Tourist No</u>
PeopleCatered	Tourist Name
Ensuite (Y/N)	Tourist Address
<u>Tourist_No</u>	Tourist Phone
Kitchen (Y/N)	Date Booked
Date Arriving	<u>Cabin_No1</u>
Date Leaving	<u>Cabin_No2</u>
Type Of Shelter	<u>Cabin_No3</u>
	Total Cost
	Duration Of Stay

Database Lecture 8

28

28

Exercise 2: Normalise

- What is wrong with the following set of tables, designed for a small antique book library? How would you redesign them?

BOOK	BORROWER
<u>Catalog_No</u>	<u>Borrower_No</u>
Author	Borrower_Name
Title	Borrower_Address
<u>Borrower_No</u>	<u>Catalog_Numbers*</u>
Borrower_Address	Fines_Owing
Date_Due	Fines_Paid
	Date_Returned

Database Lecture 8

29

29