Good Database Design

- no redundancy of FACT (!)
- no inconsistency
- no insertion, deletion or update anomalies
- no information loss
- no dependency loss

Data Redundancy

- **Major aim of relational database design is to group attributes into relations to minimize data redundancy and reduce file storage space required by base relations.**

- **Problems associated with data redundancy are illustrated by comparing the following Staff and Branch relations with the StaffBranch relation.**

Data Redundancy

Staff

staffNo	sName	position	salary	branchNo
SL21	John White	Manager	30000	B005
SG37	Ann Beech	Assistant	12000	B003
SG14	David Ford	Supervisor	18000	B003
SA9	Mary Howe	Assistant	9000	B007
SG5	Susan Brand	Manager	24000	B003
SL41	Julie Lee	Assistant	9000	B005

Branch

branchNo	bAddress
B005	22 Deer Rd, London
B007	16 Argyll St, Aberdeen
B003	163 Main St, Glasgow

Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

Data Redundancy

- **StaffBranch relation has redundant data: details of a branch are repeated for every member of staff.**

- **In contrast, branch information appears only once for each branch in Branch relation and only branchNo is repeated in Staff relation, to represent where each member of staff works.**

Update Anomalies

- **Relations that contain redundant information may potentially suffer from update anomalies.**

- **Types of update anomalies include:**
  - **Insertion**
  - **Deletion**
  - **Modification**

Data Redundancy

Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

Insertion Anomalies

- **New member of staff joins branch B005**
  - – Insert new row into StaffBranch table
  - – Type wrong address: 163 Main St, Glasgow.
  - – Database is now inconsistent!

- **Establish new branch with no members of staff**
  - – B008, 57 Princes St, Edinburgh
  - – No staff members, so staffNo must be NULL
  - – But staffNo is the primary key of the StaffBranch table, so cannot be NULL!

Deletion Anomaly

- **Mary Howe, staffNo SA9, leaves the company**
  - **Delete the appropriate row of StaffBranch**
  - **This also deletes details of branch B007 where Mary Howe works**
  - **But no-one else works at branch B007, so we no longer know the address of this branch!**

Modification Anomaly

- **Branch B003 has transferred to a new location**
  - **New address is 145 Main St, Glasgow**
  - **Must change *three* rows of the StaffBranch relation**

Functional Dependency

- **Functional Dependency**
  - **Describes relationship between attributes in a relation.**
  - **If A and B are attributes of relation R, B is functionally dependent on A (denoted A ->B), if each value of A in R is associated with exactly one value of B in R.**
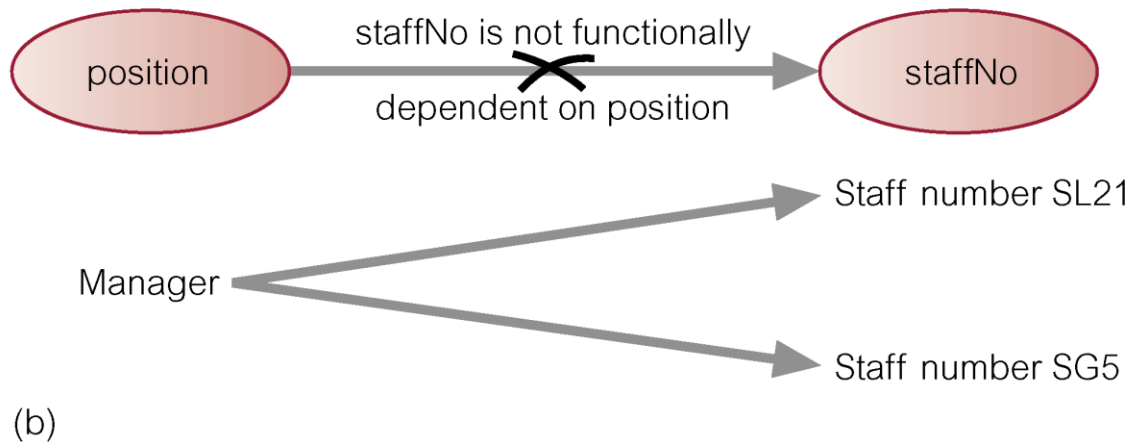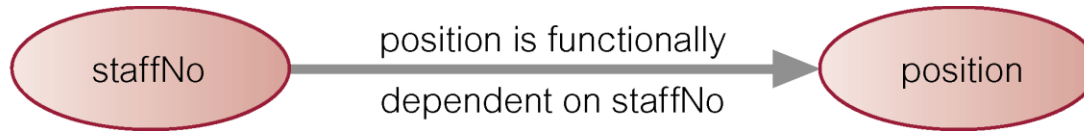
Functional Dependency

• **Diagrammatic representation:**

A → B is functionally dependent on A → B
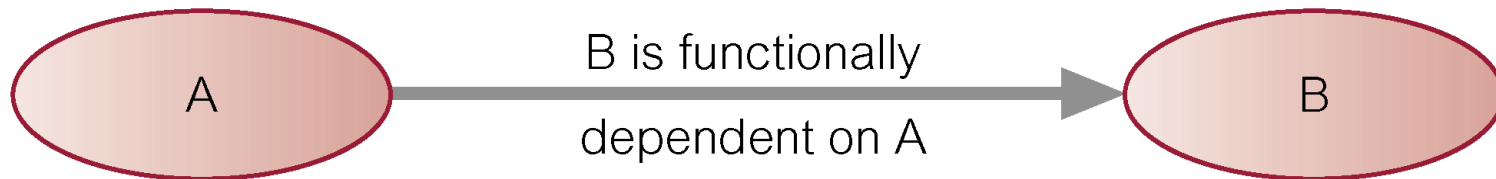
☐ *Determinant* **of a functional dependency refers to attribute or group of attributes on left-hand side of the arrow.**

Example - Functional Dependency

staffNo → position is functionally dependent on staffNo

position

Staff number SL21 → Manager

(a)

position → staffNo is not functionally dependent on position ✗ staffNo

Manager → Staff number SL21

Manager → Staff number SG5

(b)

Identifying Candidate Keys

- **A candidate key is an attribute, or set of attributes, that uniquely identifies a row**
  - **Must be *irreducible***
  - ***No part of a candidate can ever be NULL***

- **An attribute A that functionally determines *every other* attribute of the relation is a candidate key**
  - **For each value of A there is exactly one value of each of the other attributes**
  - **So each value of A must identify a single row**

Identifying Primary Keys

- **A primary key is a candidate key chosen to identify rows uniquely within a table**
  - **Other candidate keys called *alternate keys***

- **Some guidelines on choosing the primary key**
  - **Pick the candidate key with fewest attributes**
  - **Pick the candidate key that makes most sense for the business!**

Functional Dependency

Formal Definition: Attribute B is functionally dependant upon attribute A *(or a collection of attributes)* if a value of A determines a single value of attribute B at any one time.

Formal Notation: $A \rightarrow B$ This should be read as 'A determines B' or 'B is functionally dependant on A'. A is called the *determinant* and B is called the *object of the determinant*.

Example:

staffNo	job	dept	dname
SL10	Salesman	10	Sales
SA51	Manager	20	Accounts
DS40	Clerk	20	Accounts
OS45	Clerk	30	Operations

Functional Dependencies

staffNo $\rightarrow$ job

staffNo $\rightarrow$ dept

staffNo $\rightarrow$ dname

dept $\rightarrow$ dname

Partial Functional Dependency

Partial Functional Dependency: This is the situation that exists if it is necessary to only use a subset of the attributes of the composite determinant to identify its object uniquely.

Example:

Student#   Crs#   CrsName   grade

Full Functional Dependencies
(Student#, Crs#) $\rightarrow$ grade

Partial Functional Dependencies
Crs# $\rightarrow$ CrsName

Full Functional Dependency

Full Functional Dependency: Only of relevance with composite determinants. This is the situation when it is necessary to use all the attributes of the composite determinant to identify its object uniquely.

Example:

order#	Pd#	qty	price
A001	001	10	200
A002	001	20	400
A002	002	20	800
A004	001	15	300

Full Functional Dependencies
(Order#, Pd#) $\rightarrow$ qty
(Order#, Pd#) $\rightarrow$ price