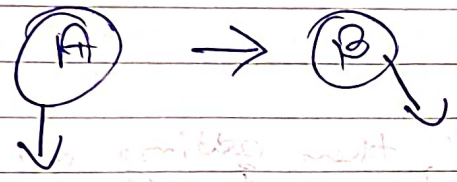
Normalisation

- It is an step towards Database optimisation.
- Functional dependency (FD)

When value of B is dependent on A. Like from taking value of A we can determine B.



→ Types of FD

① Trivial FD

A → B ; B is subset of A.

{Emp-ID, name} → {Emp-ID}

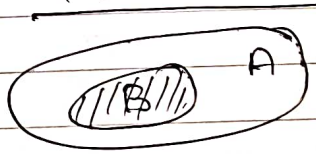
② Non-trivial FD

A → B ; B is not a subset of A.

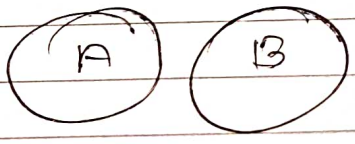
{emp-ID, name} → {Emp-address}

- Venn diagram of the both types of FD above:-

Trivial FD



Non-trivial FD



$A \cap B = NULL$.

→ **Rules of FD**

① **Reflexive**

:- If "A" is a set of attribute and "B" is a subset of "A", then A → B.

② **Augmentation**

:- If "A" can determine "B", then adding an attribute to this functional dependency won't change anything.

A → B

AX → BX.

③ **Transitivity**

If A determines B and B determines C, then A determines C.

A → B, B → C, A → C.

④ Why Normalization?

:- To avoid redundancy from DB.

⑤ What happen if we have redundant data?

:- Insertion, deletion and updation anomalies arises.

→ Anomalies / abnormality

① Insertion anomaly

- When certain data (attribute) can not be inserted into the DB without the presence of other data.

② Deletion anomaly

- The delete anomaly a situation where the deletion of data results in the unintended loss of some other important data.

③ Updation anomaly

- A single data update required multiple rows of data to be updated.

→ Due to these anomalies, DB size increases and DB performance become very slow.

→ To rectify these anomalies, we use database optimisation technique called Normalization.

→ Normalisation divides the composite attributes into individual attributes or larger table into smaller and link them using relationships.

→ We decomposes the table until we achieve "SRP"

SRP ⇒ Single Responsibility Principle,

Like one table constaint info of one entity only.

→ Types of Normal forms

① 1NF

- Every relation cell must have atomic value.
- Relation must not have multi-valued attributes.

② 2NF

- Relation must be in 1NF.
- There should not be any partial dependency
- All non-prime attributes must be fully dependent on Primary key.
- Non prime attributes can not depend on the part of the primary.

③ 3NF

- Relation must be in 2NF
- No transitivity dependency exists
- Non-prime attribute should not find a non-prime attribute.

④ BCNF (Boyce-Codd normal Form)

- Relation must be in 3NF.
- FD:- A → B, A must be a Super key.
- We must not derive prime attribute from any prime or non-prime attribute.