Tutorial 8

Arvind Viswanathan
viswaa2@mcmaster.ca

Agenda

- Functional Dependency
- Closure
- Minimal Basis
- Schema Decomposition
- Lossy Decomposition

Functional Dependency

- What is an FD?
- “It is a kind of an Integrity Constraint that generalizes the concept of a key.” - Ramkrishnan & Gehrke

- Let’s look at an example to simplify this:

FD: Street, City, Country ->Pincode

Street	City	Country	Pincode
Broadway Ave	Hamilton	Canada	L8S2V6
Emerson	Hamilton	Canada	L8S2V4
Main St	Hamilton	Canada	L8S4L8
Sussex St	Hamilton	Canada	L8S3L4

FD’s

- Let’s break it down further to understand why FD?

- Functional: It’s a mathematical way of showing that for the values of the attributes on the L.H.S we get a unique value on the R.H.S.

- Dependency: The value on the R.H.S. depends on the value in the L.H.S. and hence, is called a dependency.

Basic Terminology to know in FD

- Trivial FD's: These are some FD's which are very obvious and are not necessary to be stated as such.

- Examples of Trivial FD's:
- A -> A
- B-> B

Superkeys

- What is a super key?
- It can be used to determine all other attributes.
- Can you give an example of a Superkey from the table given below?
- Mac_Student(name, addr,nationality, dept, year)
- {name, addr, dept} can be a superkey}
- There are always exceptions and you must choose the keys wisely based on the situation.

Closure

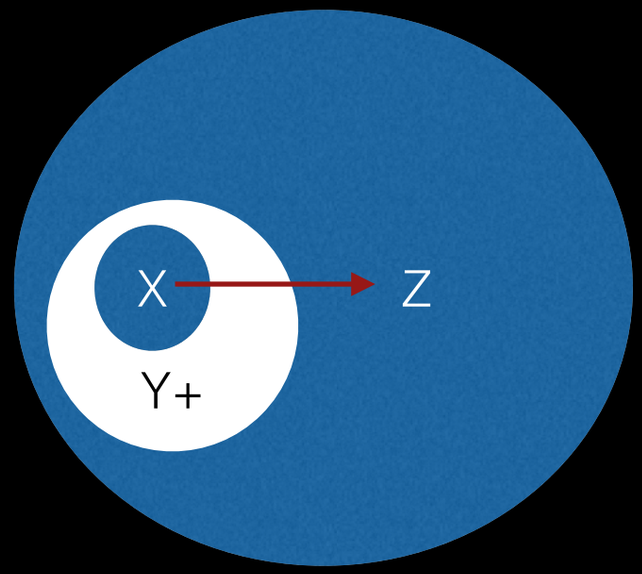
- Think of this in the form of Set Theory. If you can visualize it will become easy to solve.

An example would be a

closure test by which we can

say that the closure of the initial

elements given is equivalent to the result.



- Given R(a,b,c,d,e,f)

ab ->c, ac->d, c->e, ade->f

Find the Closure of {a,b}?

Answer: {a,b,c,d,e,f}

Closure Test

- Let’s try out another example of the closure test:

R{x,y,z,a1,a2,a3,a4}

x->a1, y->a2, a1->a3, z->a4, a4->a2, a2->x, a2->a1, y->a3
Find the closure of {x,z} and give suitable reasons why?

Minimal Basis

- This is also known as Minimal Cover.
- Some points to note with relation to FD's:

    R.H.S. are single attributes

    No FD can be removed

        No attribute can be removed from the L.H.S.

- How to construct a minimum cover?
    - Decompose the R.H.S
    - Remove FD's and see if the remaining is still consistent with the original set.
    - Repeatedly try to remove an attribute from LHS and see if it can be derived back using the remaining FDs.

Construction of Minimal Basis

- Split R.H.S. into singletons
- For all f in F’ test if J= (F’-f) is closure of F
- For all i in L.H.S(f) test if (F’-f+f’) is equivalent to the closure of F
- Repeat steps 2 and 3 till there is no further progress

Minimum Basis: Example

- R(V,X,Y,Z,) <—Relation
- The FD’s defined are:
- F = {V->VY, X->VXY, Z->VXY}
- What is the minimal basis of F.

V->Y, X->V, Z->X

Schema Decomposition

- The goal:
- To avoid redundancy and anomalies that are created because of it.
- Update & Deletion Anomaly
- Update: An occurrence of a fact is changed but not all have been updated accordingly.
- Deletion: When a tuple is deleted then along with it a fact is lost in the process.

Anomaly is a consequence of Bad Design

- Let’s design a database with redundancy and bad design and see what the result is?

An Anomaly!

Splitting Relations

- The tuples have to be divided in such a way that on ‘joining’ them again there is no information loss.

Street	City	Country	Pincode
Broadway Ave	Hamilton	Canada	L8S2V6
Emerson	Hamilton	Canada	L8S6V4
Main St	Hamilton	Canada	L8S4L8
Sussex St	Hamilton	Canada	L8S1F3

Information Loss

- What is this?
- Information Loss is when we have a lack of information or redundant information which leads to ambiguity and thereby in us not able to come to a result.
- To put it in layman's terms, if we were to reach a goal and if there was information loss we would not be able to reach the same goal due to misleading information or some missing information.

Types of Lossy decomposition?

- FD Loss : When a decompostion is done and then rejoined the attributes are no longer enforced by the same FD's.

- Join Loss: On joining decomposed tuples the result is supposed to get back the original tuples.The lack of being able to come back to this original state is called join loss. There may be extra tuples or a lack or tuples depending on the schema and the tables in use.

Questions?

Thank You!

Contact: viswaa2@mcmaster.ca

Office Hours: Thursdays, 3-4 PM, ITB 128