INFO20003: Database Systems

Dr Renata Borovica-Gajic

Lecture 23

Overview, sample exam questions

Part I

Semester 1 2018, Week 12

WHY ARE DATABASES COOL ?

Featuring people from industry

Part 4

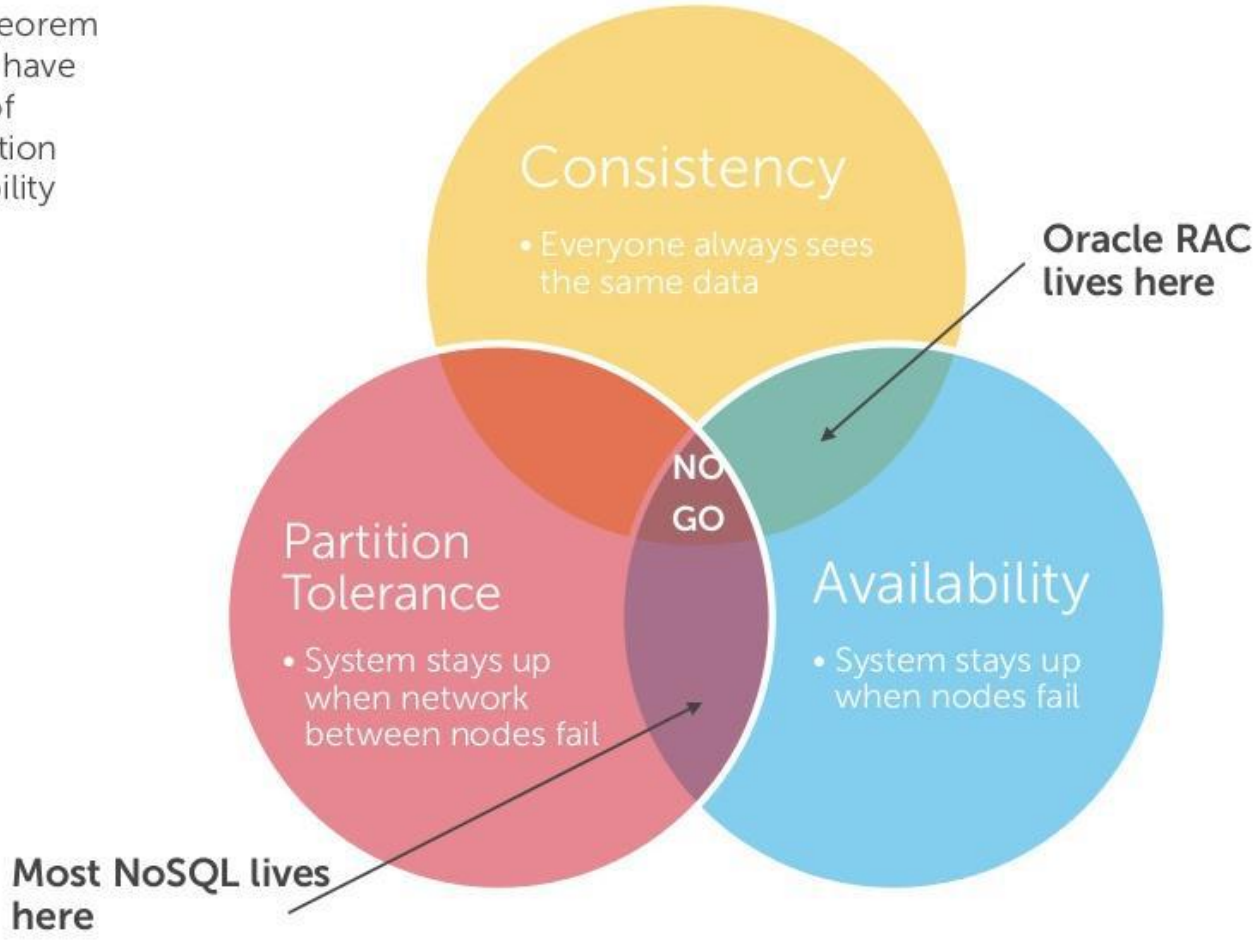
IBM Research, USA

From the previous lecture…

The CAP theorem: Revisited

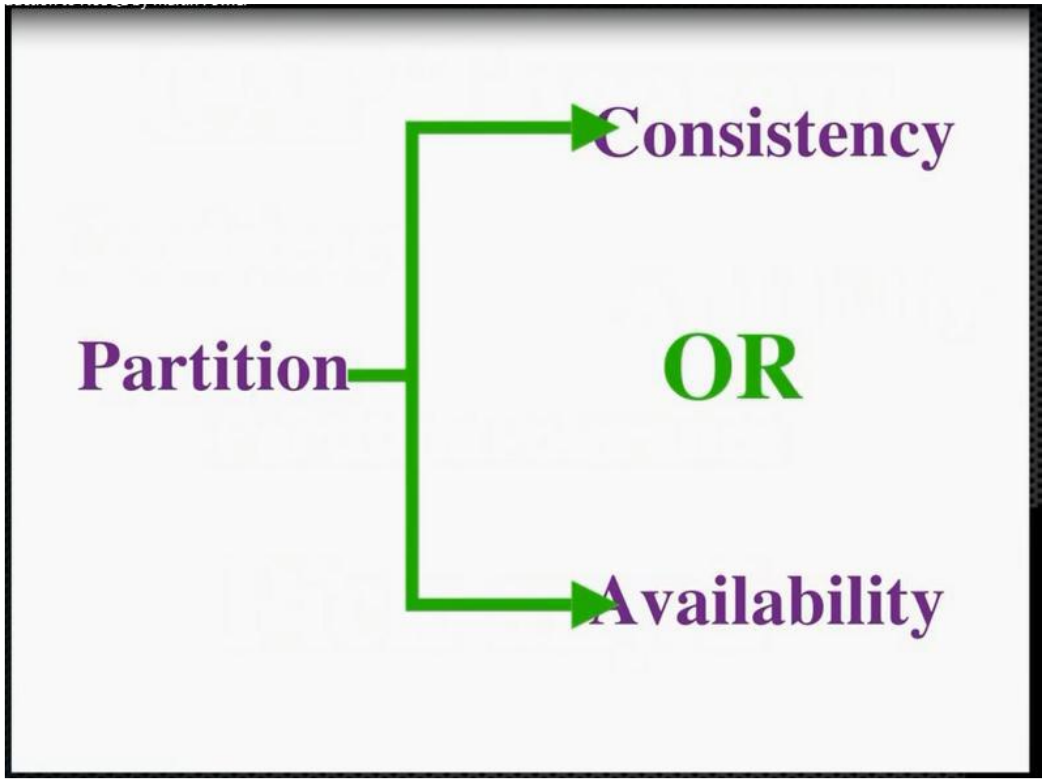
CAP Theorem says something has to give

- CAP (Brewer's) Theorem says you can only have two out of three of Consistency, Partition Tolerance, Availability



CAP theorem: alternative presentation

- Fowler's version of CAP theorem: *If you have a distributed database, when a partition occurs, you must then choose consistency OR availability.*



ACID vs BASE

ACID (Atomic, Consistent, Isolated, Durable)
vs
Base (Basically Available, Soft State, Eventual Consistency)

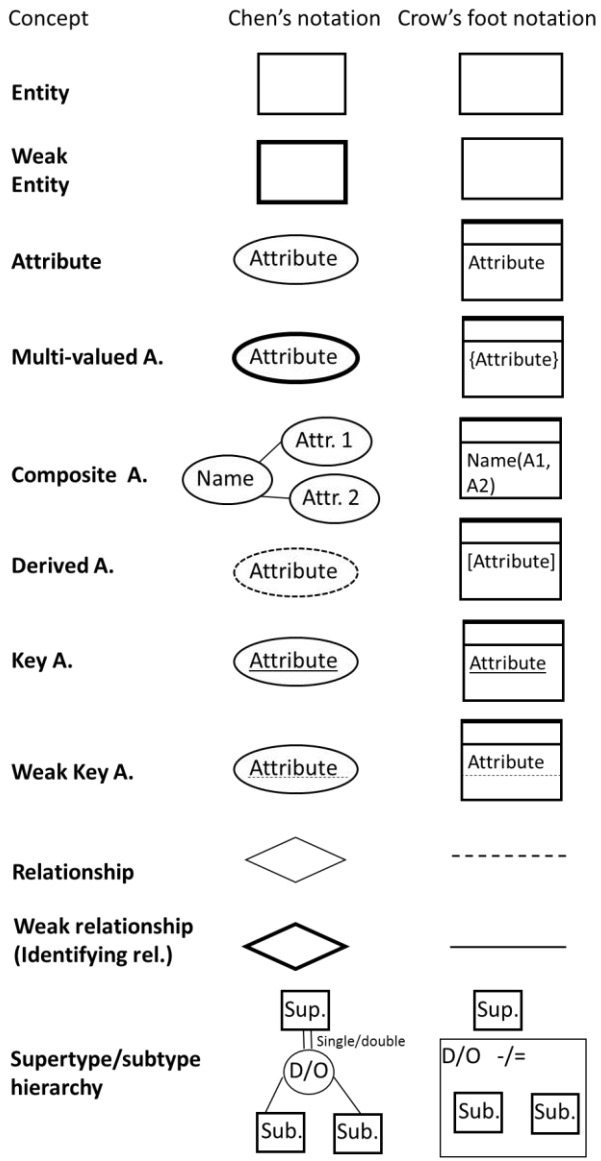
- **Basically Available**: This constraint states that the system does guarantee the *availability* of the data; there will be a response to any request. But data may be in an *inconsistent* or *changing* state.

- **Soft state**: The state of the system could change over time -even during times without input there may be changes going on due to 'eventual consistency'.

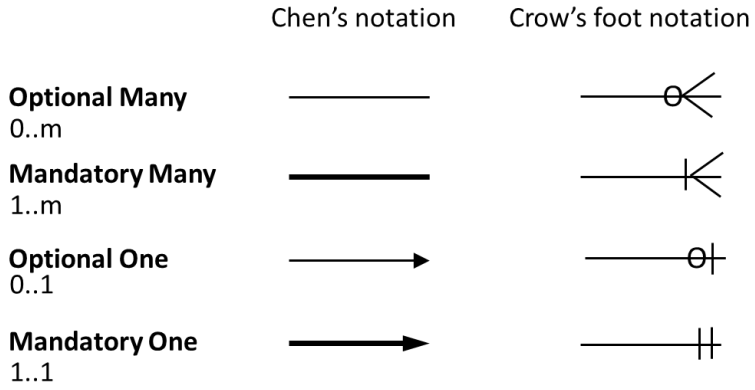
- **Eventual consistency**: The system will eventually become consistent once it stops receiving input. The data will propagate to everywhere it needs to, sooner or later, but the system will continue to receive input and is not checking the consistency of every transaction before it moves onto the next one.

Overview

Modelling – learn notations

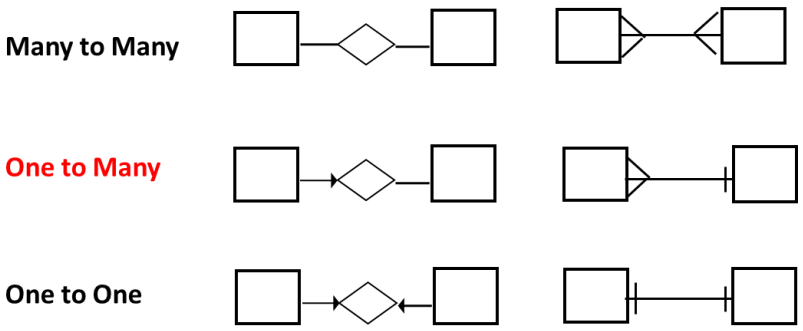


Relationship Cardinality/Constraints



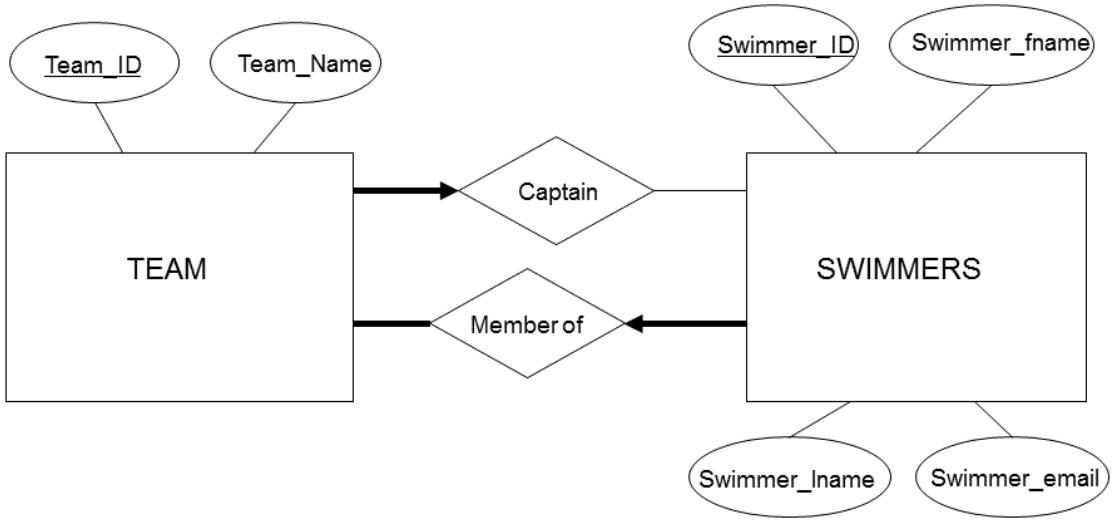
BINARY Relationship Cardinalities

Here we just looked at cardinalities and omitted participation constraints (optional/mandatory) for clarity

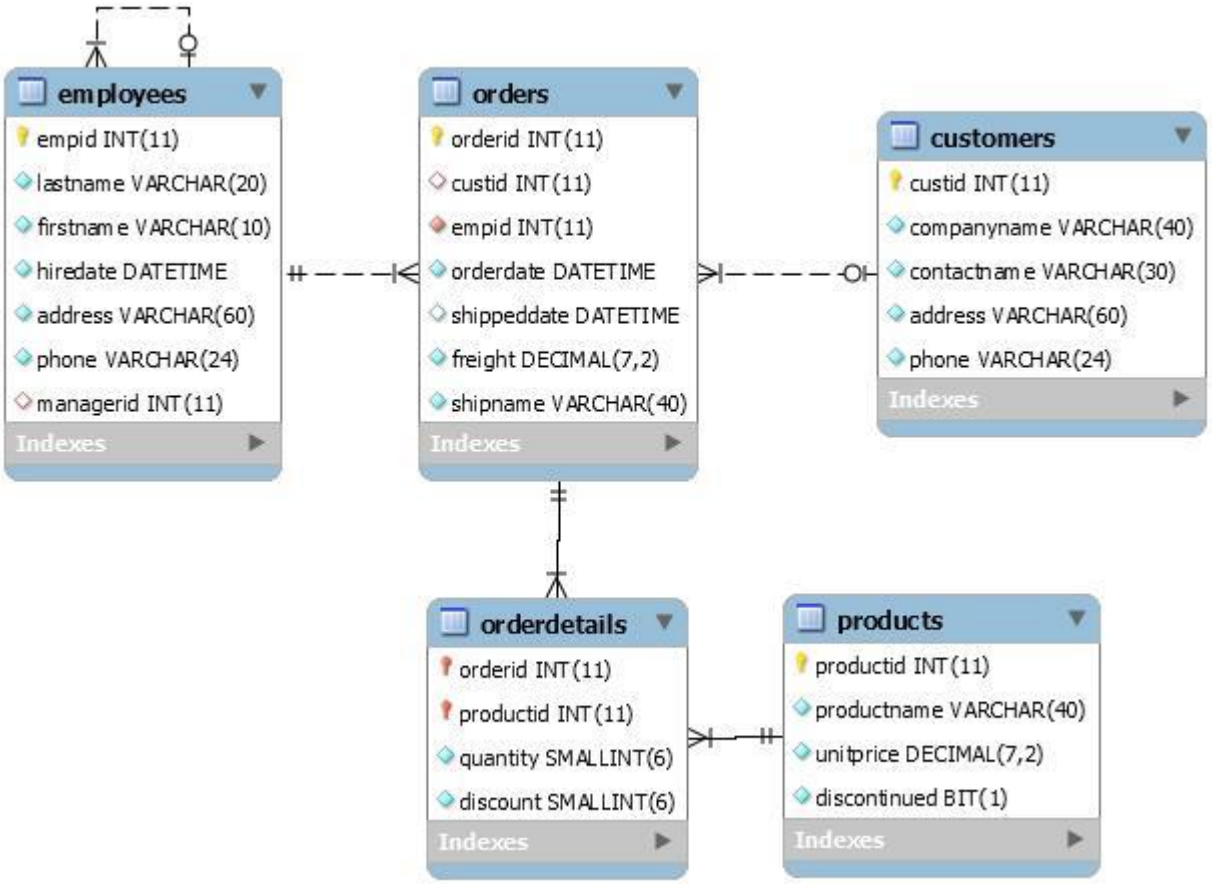


Example

Q.2. Write SQL statements to create the tables for the above data model. Be sure to specify primary and foreign keys.

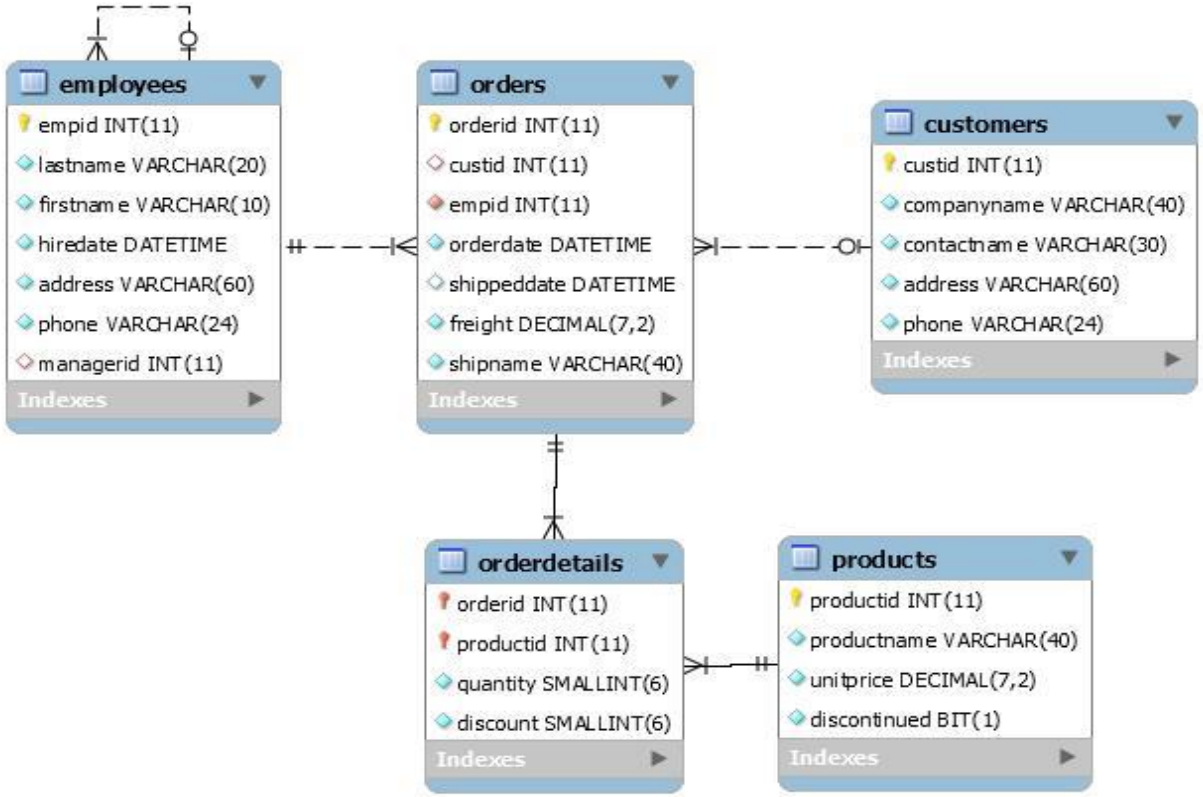


Q3. Given the following schema, write a *single* SQL statement to correctly answer each of the following questions (3A – 3D). DO NOT USE VIEWS to answer questions.



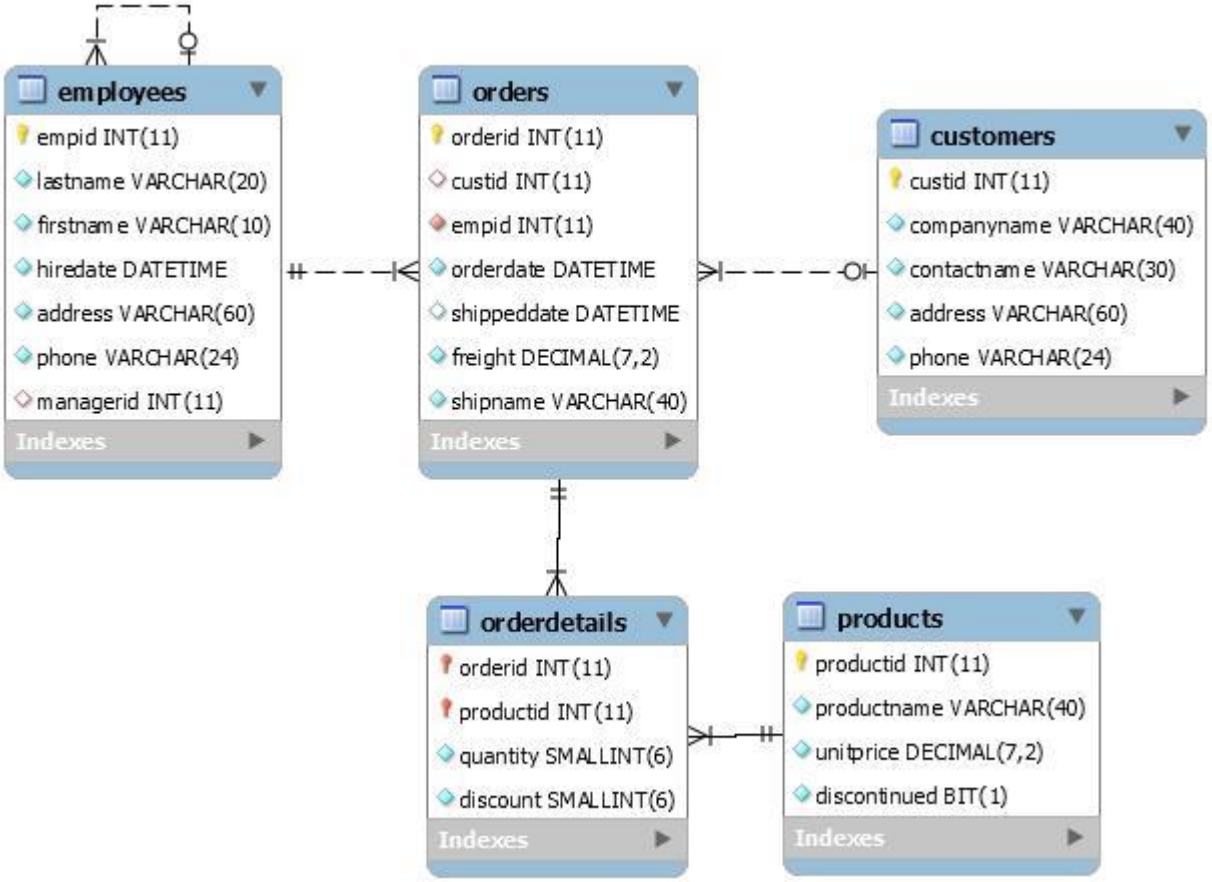
SQL: Q1

Write a query that returns the first name and last name of employees whose manager was hired prior to 01/01/2002.



SQL: Q2

Write a query that returns customers (IDs) whose company name is ‘Google’, and for each customer return the total number of orders and total quantities for all products that were not discontinued (‘1’ means discontinued, ‘0’ not discontinued).



Query Processing – learn formulae

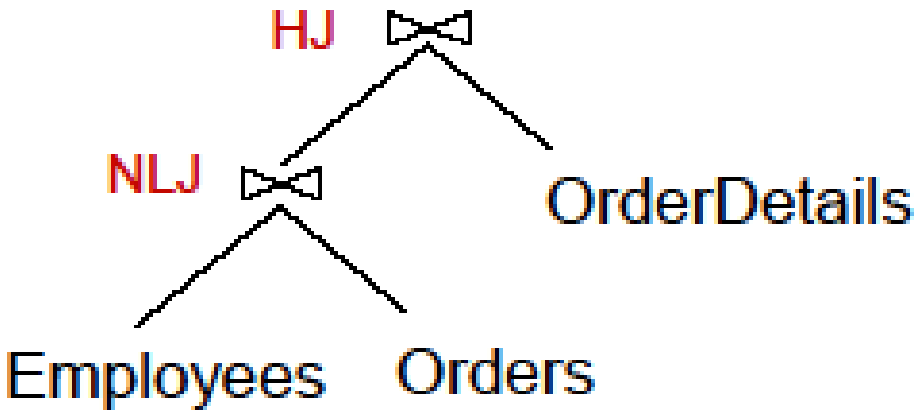
- Resources/Query Processing Cost Formulae

Query Processing 1/2

Consider relations *Employees*, *Orders* and *OrderDetails*. Imagine that relation Employees has 1,000 pages, relation Orders 5,000 pages, and relation OrderDetails 10,000 pages. Each page stores 100 tuples, and neither relation has any indexes built on it. Consider the following query:

```
SELECT *
FROM Employees as E, Orders as O, OrderDetails as OD
WHERE E.empid = O.empid AND O.orderid = OD.orderid;
```

Compute the cost of the plan shown below. NLJ is a *Page-oriented* Nested Loops Join. Assume that *empid* is the candidate key of Employees, *orderid* is the candidate key of Orders, and 100 tuples of a resulting join between Employees and Orders can fit on one page.



Query Processing 2/2

Consider the query presented below. Does the following equivalence class hold? Yes/No and Why?

```
SELECT firstname, lastname
FROM Employees NATURAL JOIN Orders NATURAL JOIN OrderDetails
WHERE quantity > 5 AND freight < 100
```

$$\Pi_{\text{firstname, lastname}}(\sigma_{\text{quantity}>5 \wedge \text{freight}<100}(\text{Employees} \bowtie \text{Orders} \bowtie \text{OrderDetails}))$$

$$\leftrightarrow \sigma_{\text{quantity}>5 \wedge \text{freight}<100}(\Pi_{\text{firstname, lastname}}(\text{Employees} \bowtie \text{Orders} \bowtie \text{OrderDetails}))$$

Summary

- Read questions *carefully*
- Be aware of the time (~mark a minute)
- Tackle the questions you know first
- If you are spending too much time on something, leave it and come back later to it

Next lecture

- Sample exam questions Part II