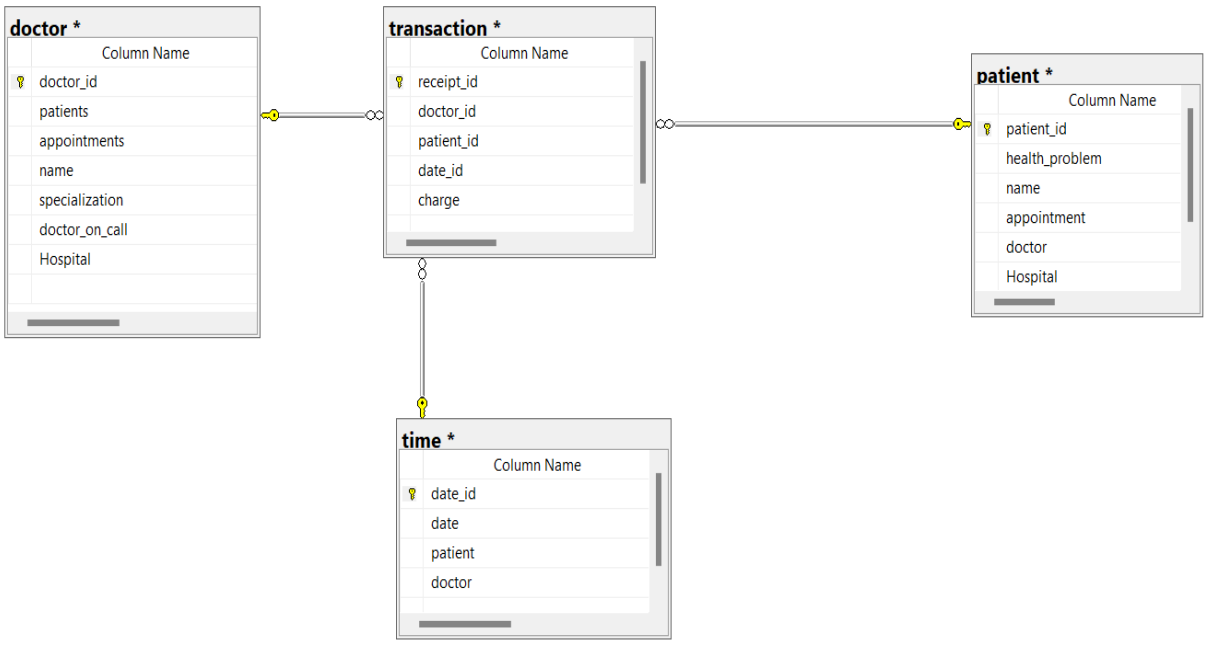
STUENT: ANTONIOS PROMPONAS

ID: 03400232

Question 2

Part A



In the above diagram, you can see a star schema data warehouse. In particular, this schema has 3 dimensions and 1 fact table(as it is supposed to be). The 3 dimensions are refered below:

i. Doctor: The table “doctor” has as primary key the variable “doctor_id “. It has 6 more attributes that are associated with the patients and their appointments , the name , the specialization and the hospital in which each doctor works.
ii. Patient: The table “patient” has as primary key the variable “patient_id “. It has 5 more attributes that are associated with their appointments ,

the doctor and the and the hospital in which each patient is hospitalized. In addition, it includes attributes about the names and the health problems of the patients.

iii. Time: The table “time” has as primary key the variable “date_id” . It has 3 more attributes that are associated with the doctors,patients and their appointments.

Both of the above dimensions are connected with the fact table “transaction”. This fact table includes the foreign keys “doctor_id”, “patient_id” and “date_id” from the corresponding tables. It obviously contains its own primarey key that is “receipt_id”. In addition, it includes the variable “charge” which is a measure that refers to the the fee that a doctor receives from the patient for the visit.

Part B

This exercise demands from us to find the OLAP operations that are necessary in order to list “the total fee collected by each doctor for 2022”. Taking this sentence and breaking it in small pieces helps us to find more easily the appropriate operations. In particular, the part “total fee” means that we need to do a sum. Consequently, we will need to do pivoting or in other words aggregation. The part “collected by each doctor ” lead us to execute a group by. So , it is demanded to perform a roll up in variable “doctor”. Finally, the part “for 2022” refers to a particular period of the variable day which means that the slice and dice operation is required.

So, we will need to perform these 3 operations:

i. Roll up
ii. Slice and dice
iii. Pivoting (Aggregation)

Now, I have to specify the order of the above operations.

Firstly, I perform slice and dice (and to be more specific , I execute slice)in variable “day” in order to seperate and take all the data of the days that belong to the year of 2022. Then , I execute 2 roll up. The first one, is executed

in variable “doctor ”and the other one is executed in variable “date” that now contains only days of 2022 as a value. The result of these OLAP operations is that we have eliminated the variable “patient” and we have grouped by doctors and days of 2022. That means that we have gathered information about how much money have the doctors received from each day of 2022. What remains to do now, is to sum the fee so, I execute pivoting so as to take the final amount of money that are collected by each doctor in 2022. Thus, I perform aggregation in variable (measure) “charge” for all the days of 2022. As a result I am finally able to get and appear the list of the total fee collected by each doctor for 2022.

In conclusion, I would like to point out that there many different ways to modify or perform the desirable result. The method depends on the form of the data and how we handle it. But , what’s for sure, is that we firstly have to slice on “day”, then we have to roll up on “doctor” (and maybe on the “day”) and finally we have to pivot on the measure variable “charge”.

Question 1

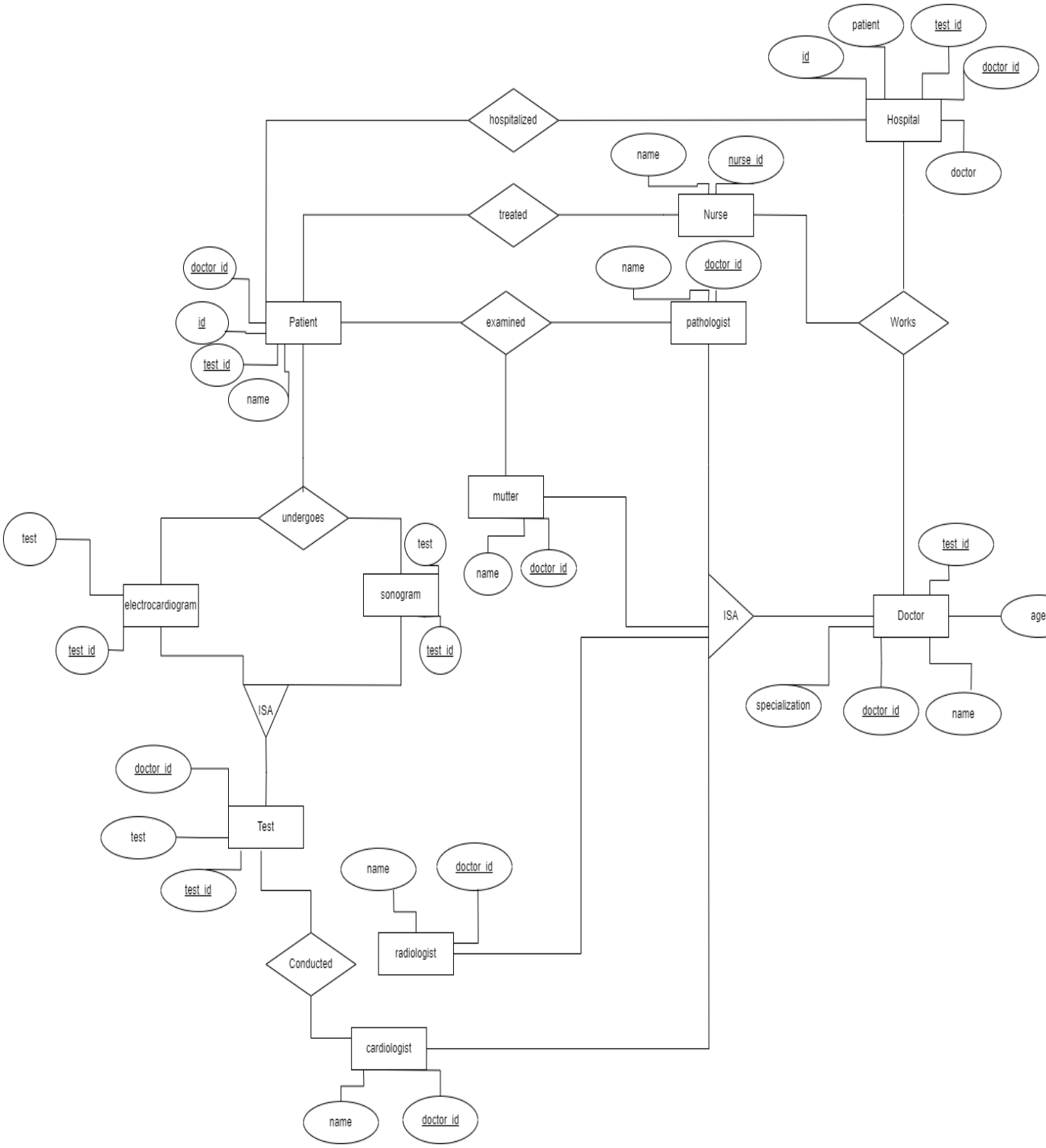
Part A

The diagram below, depicts an ER diagram that is associated with a patient who is hospitalized in hospital. The patient undergoes a variety of many different medical tests and is examined by many different doctors.

In particular, the nurses take care of the patient. The patient undergoes medical tests such as electrodiagram and sonogram which are conducted by particular doctors. To be more spefic cardiologist conduct electrodiagram and radiologist conduct sonogram.The patient is also examined by a pathologist and a mutter. The common factor of the nurses and all of these doctors is that they work in the hospital that the patient is hospitalized.

At this point, I would like to point out that in the below er diagram, I have created 2 subclasses. The first subclass is refered to the doctors that examine the patient and the other subclass is refered to the tests that the patient undrgoes.In particular, electrodiagram and sonogram belong to class Test, while nurses, pathologist,mutter, cardiologist and radiologist belong to class doctor. Their relation is expressed by the triangle the contains the word "ISA".The subclasses contain attributes of the class that belong to.

The circles of each entity , are the attributes of the entity and the underlined atributes are the primary and foreign keys of the entities.


patient
test_id
id
doctor_id
hospitalized
Hospital
name
nurse_id
doctor
treated
Nurse
doctor_id
name
doctor_id
id
Patient
examined
pathologist
Works
test_id
name
mutter
undergoes
test
test
name
doctor_id
test_id
electrocardiogram
sonogram
ISA
Doctor
age
test_id
test_id
test_id
specialization
doctor_id
name
ISA
doctor_id
test
Test
test_id
name
doctor_id
radiologist
Conducted
cardiologist
name
doctor_id

Part B

B.A:

SQL:

```
Select e.employee_name, e.street, e.city

FROM employee e, works w, company c, manager m

WHERE e.employee_name=w.employee_name AND
w.company_name=c.company_name  AND

w.salary>22000 AND

c.company_name='NT University';
```

RELATIONAL ALGEBRA:

π employee.employee_name, employee.street, employee.city(σ employee.employee_name = works.employee_name$\wedge$works.company_name = company.company_name$\wedge$works.salary > 22000$\wedge$company.company_name = 'NT University'$\wedge$employee.employee_name = manager.manager_name(employee$\times$works$\times$company))

B.B:

SQL:

```
Select e.employee_name
FROM employee e, works w, company c
WHERE e.employee_name=w.employee_name  AND
w.company_name=c.company_name AND
e.city=c.city;
```

RELATIONAL ALGEBRA:

$$\Pi_{\text{employee\_name}}(\sigma_{\text{employee.employee\_name} = \text{works.employee\_name} \wedge \text{works.company\_name} = \text{company.company\_name} \wedge \text{employee.city} = \text{company.city}}(\text{employee} \times \text{works} \times \text{company}))$$