IFB105 Modelling Information Systems

Assignment 3A Handout
- Information Modelling with ORM -
Semester 1, 2021

Due date: 23 April 2021 (11:59pm)
Weight: 20%

Assignment Declaration

By submitting this assignment, I am/We are aware of the University rule that a student must not act in a manner which constitutes academic dishonesty as stated and explained in the QUT Manual of Policies and Procedures. I/We confirm that this work represents my individual/our team's effort. I/we have viewed the final version and declare that it does not contain plagiarized material.

Full Name	Student No.	Signature
James Norman	n10256989	James Norman

SCENARIO

Use the data as shown in Table 1 and Table 2 to come up an ORM information/data model for Task 1 requirements.

Table 1 - COVID Variants known to be originated with date and country information

Covid Variant	Country Originated	Date known to be originated
SARS-CoV-2	China	Dec. 2019
B.1.1.7	UK	Oct. 2020
B.1.351	South Aftrica	Jan. 2021
P.1	Brazil	Jan. 2021

Table 2 - Patients' details infected by various COVID variants, the country they belong to and outcomes after infections

Patient Name	Date effected	Varient	Country	Recovered	Death
Jason Li	18/12/2019	SARS-CoV-2	China	Yes	
Ramesh Lal	19/01/2020	SARS-CoV-2	India		Yes
John Smith	20/12/2020	B.1.1.7	UK		Yes
Liam Luski	1/01/2021	B.1.1.7	Australia	Yes	
Enzokuhle Dlamini	15/01/2021	B.1.1.7	South Africa		Yes
Robert Smith	20/01/2021	B.1.351	UK	Yes	
Christopher Mission	2/02/2021	P.1	US	Yes	

TASK 1 (15 Marks)

Your task is to design an information (data) model that suggests an effective and efficient way for storing data of an Australian government agency working in the health area. As you are an expert in ORM, you decide to develop an ORM model.

Complete the drawing of the ORM conceptual schema for the universe of discourse as specified in the tables as shown in scenario above by performing steps 1-5 as described below of the Conceptual Schema Design Procedure.

CSDP Step 1: Transform familiar examples into elementary facts and apply quality checks.

CSDP Step 2: Draw the fact types and apply a population check.

CSDP Step 3: Check for entity types to be combined and note any arithmetic derivations.

CSDP Step 4: Add uniqueness constraints and check the arity (length) of fact types.

CSDP Step 5: Add mandatory role constraints.

Note: Refer to the Assessment 3A Template for a complete solution and understand how you/your team is required to submit the working of each CSDP steps.

Step 1: Transform familiar examples into elementary facts and apply quality checks.

Table 1.

Singular elementary facts

- **The Covid variant 'P.1' originated in the country 'Brazil'**
- **The Covid variant 'B1.1.7' originated on the date 'Oct. 2020'**

Combined facts

- **The Covid variant 'Sars-CoV-2' originated in the country 'China' on the date 'Dec. 2019'.**

Check table recreation:

Covid variant	Country originated	Date known to be originated
Sars-CoV-2	China	Dec.2019
P.1	Brazil	
B1.1.7		Oct.2020

As there Are no new rows and the elementary facts provide all the information needed to fill out the tables it seems as though the elementary facts indeed seem to be correct.

Table 2.

Singular elementary facts

- **The patient with patient name 'Jason Li' was affected on the date '18/12/2019'.**
- **The patient with patient name 'John Smith' was infected by the variant 'B.1.1.7'.**
- **The patient with patient name 'Liam Luski' belongs to the country 'Australia'.**
- **The patient with patient name 'Enzokuhle Dlamini' had the 'death' as the outcome of infection.????**
- **The patient with patient name 'Robert Smith' has 'recovered'????**

Combined facts

- **The patient with patient name 'Christopher Mission' belonging to the country 'US' was affected on the date '2/02/2020' by the variant 'P.1' has recovered.**
- **The patient with patient name 'Ramesh Lal' belonging to the country 'India' was affected on the date '19/01/2020' by the variant 'SARS-CoV-2' has died.**

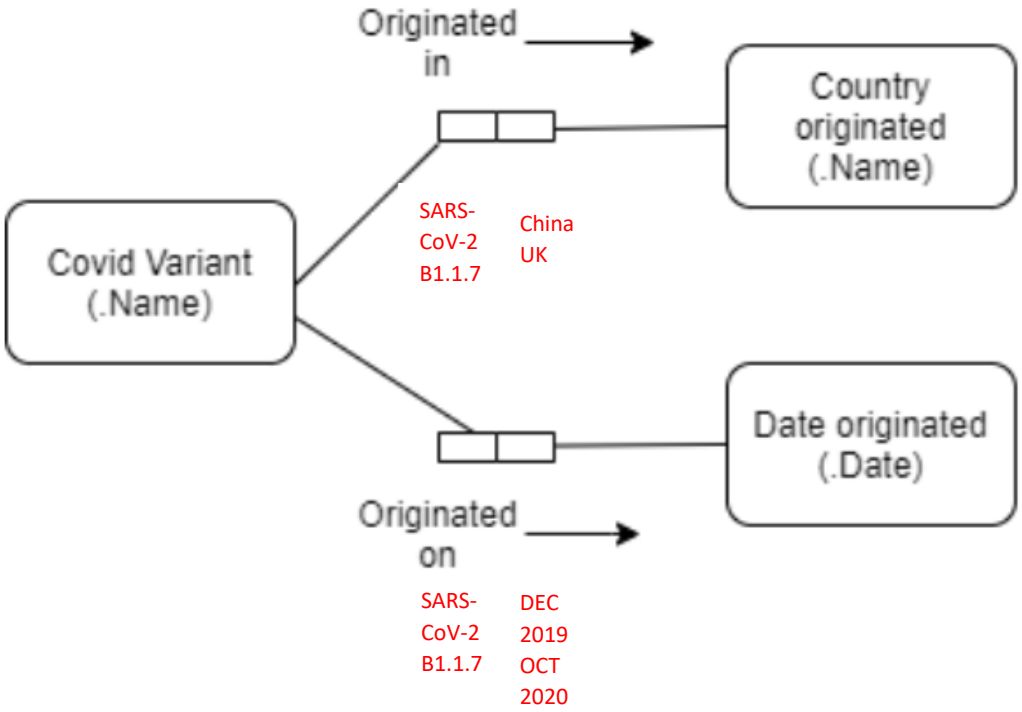
Check table recreation:

Patient name	Date effected	Variant	Country	Recovered	Death
Jason Li	18/12/2019				
John Smith		B.1.1.7			
Liam Luski			Australia		
Enzokhule Dlamini					Yes
Robert Smith				Yes	
Christopher mission	2/02/2020	P.1	US	Yes	
Ramesh	19/01/2020	SARS-CoV-2	India		Yes

As there Are no new rows and the elementary facts provide all the information needed to fill out the tables it seems as though the elementary facts indeed seem to be correct.

Step 2: Draw the fact types and apply a population check.

Table 1.
(population check)



All the relationships above are binary relationships which is decided by the elementary facts, we could use a population check to Whether we lose information in a binary and if a ternary fact type may be best. Assume that we have a ternary fact type with covid variant, country originated, and Date originated the split and joined tables would look like this.

Original table

Covid variant	Country originated	Date known to be originated
Sars-CoV-2	China	Dec.2019
P.1	Brazil	Jan.2021
B1.1.7	UK	Oct.2020

Splitting the above shows two tables

Covid variant	Country originated
Sars-CoV-2	China
P.1	Brazil
B1.1.7	UK

Covid variant	Date known to be originated
Sars-CoV-2	Dec.2019
P.1	Jan.2021
B1.1.7	Oct.2020

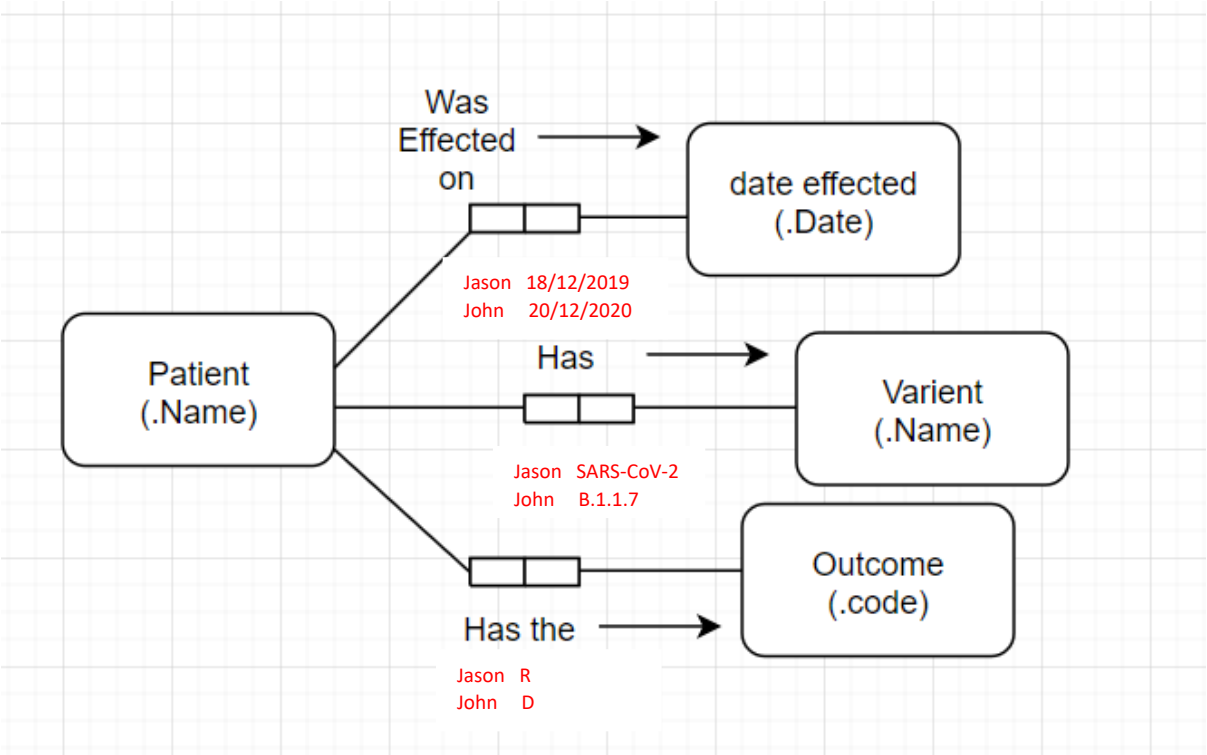
Joining them produces

Covid variant	Country originated	Date known to be originated
Sars-CoV-2	China	Dec.2019
P.1	Brazil	Jan.2021
B1.1.7	UK	Oct.2020

It can be observed that the joined table has no differences from the original table and that a ternary fact type would be incorrect. Meaning there should be two binary fact types among the three entity/ value types.

Table 2.

(Population check)



R signifies recovered and D is dead

Once again with the diagram of table two all the facts are binary facts, and we will conduct a split and join method to make sure that no information is lost and that none of these facts should indeed be ternary or of an even higher arity.

Original table

Patient name	Date effected	Variant	Country	Recovered	Death
Jason Li	18/12/2019	SARS-CoV-2	China	Yes	
John Smith	20/12/2020	B.1.1.7	UK		Yes
Christopher mission	2/02/2020	P.1	US	Yes	
Ramesh	19/01/2020	SARS-CoV-2	India		Yes

When split we get 5 separate tables that look like

1.

Patient name	Date effected
Jason Li	18/12/2019
John Smith	20/12/2020
Christopher mission	2/02/2020
Ramesh	19/01/2020

2.

Patient name	Variant
Jason Li	SARS-CoV-2
John Smith	B.1.1.7
Christopher mission	P.1
Ramesh	SARS-CoV-2

3.

Patient name	Country
Jason Li	SARS-CoV-2
John Smith	B.1.1.7
Christopher mission	P.1
Ramesh	SARS-CoV-2

4.

Patient name	Death
Jason Li	
John Smith	Yes
Christopher mission	
Ramesh	Yes

5.

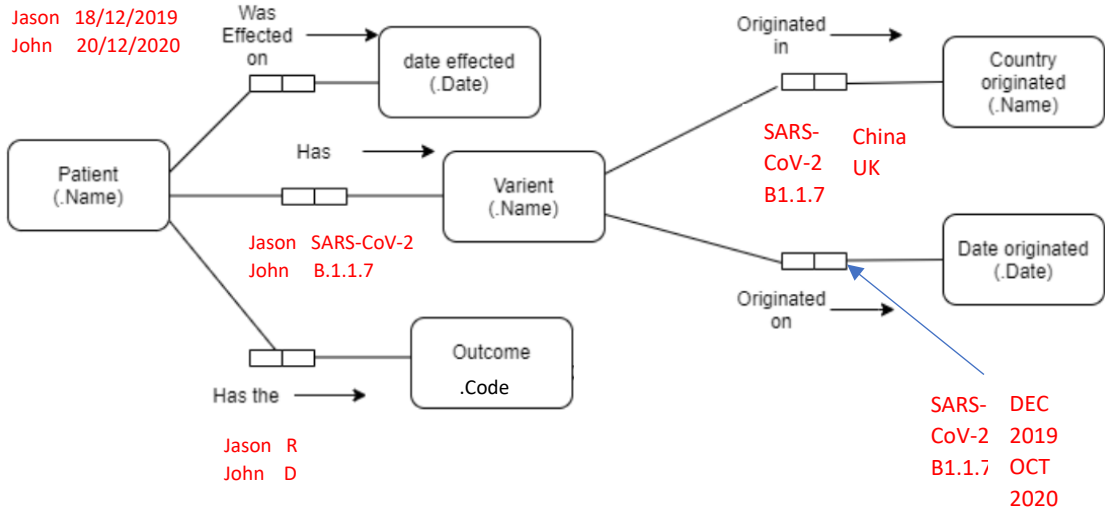
Patient name	Recovered
Jason Li	Yes
John Smith	
Christopher mission	Yes
Ramesh	

Joining them produces

Patient name	Date effected	Variant	Country	Recovered	Death
Jason Li	18/12/2019	SARS-CoV-2	China	Yes	
John Smith	20/12/2020	B.1.1.7	UK		Yes
Christopher mission	2/02/2020	P.1	US	Yes	
Ramesh	19/01/2020	SARS-CoV-2	India		Yes

It can be observed that the joined table once again has no differences from the original table and that the use of any ternary fact types would be incorrect. Meaning there should be 5 binary fact types among the three entity/ value types.

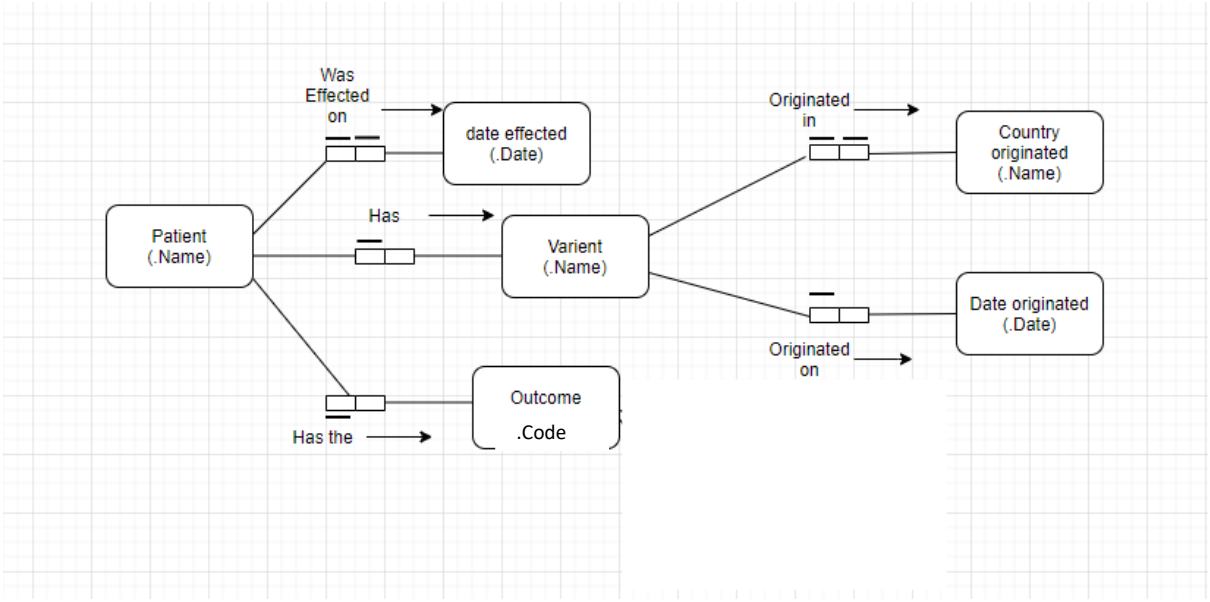
Tables 1 and 2 joined



As shown in the individual tables above when these tables are joined in a full diagram there are no ternary fact types, and all the value type and entity types should be correct.

Step 3: Check for entity types to be combined and note any arithmetic derivations.
The ORM model above whilst having fact types with the same reference scheme do not have any combinable fact types due to the way that the diagram is structured, and therefore have no entities that can be combines into a certain primitive type. Additionally, no arithmetic derivation can be found in the proposed ORM model. (you could maybe join date originated and date effected into 1 entity type “DATE”, but I chose not too as I did not know if it was correct and didn’t want to mess up the rest of the assessment.)

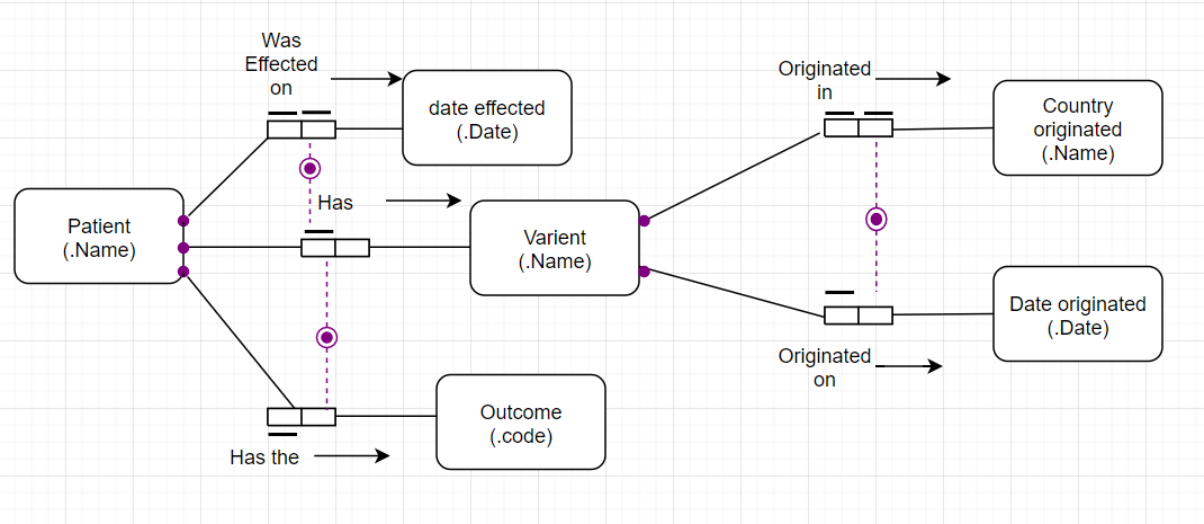
Step 4: Add uniqueness constraints and check the arity (length) of fact types.



Explanations for the identifies UC:

- In the data for each patient there is only one date effected with each patient making it a one-to-one relationship.
- Each patient only has 1 variant however variants can infect multiple patients meaning it is a one-to-many relationship.
- Each patient can only have 1 outcome however one outcome can affect many patients making it a one-to-many relationship.
- In the data each variant has exactly one country originated, and no country originated has more than one variant meaning it as a one-to-one relationship.
- Each variant has only 1 date originated however date originated has more than 1 variant so therefore it is a one-to-many relationship.

Step 5: Add mandatory role constraints, and check for logical derivations.

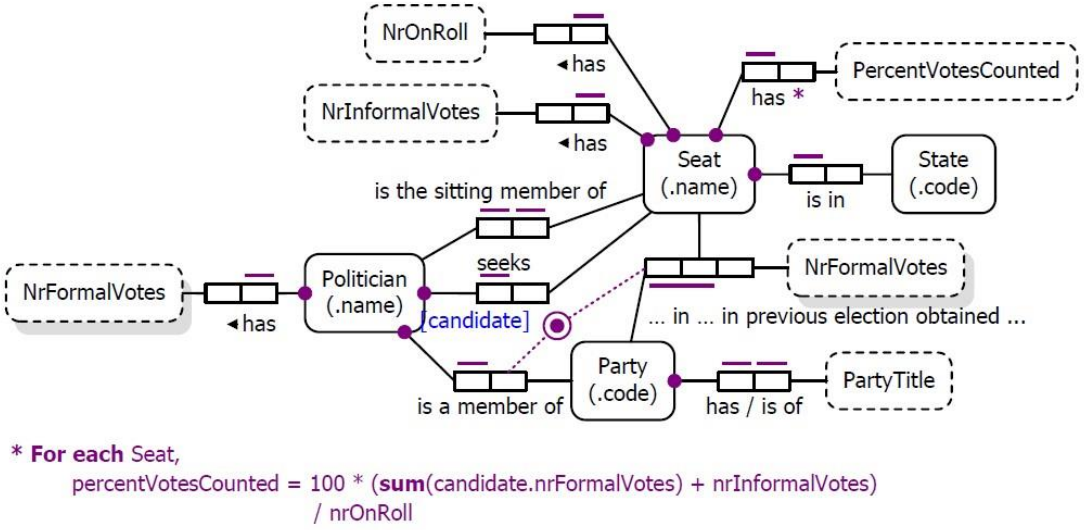


Explanations for the identifies mandatory constraints: (From the data)

- Each patient must have a date effected.
- Each patient must have a variant name.
- Each patient must have an outcome.
- Each patient must have either a variant name or an outcome.
- Each patient must have a variant name, or a date effected.
- Each Variant name must have a country originated.
- Each variant must have a date originated.
- Each variant must have either a country originated, or a date originated.
- Each outcome must either be recovered our death but also cannot be both.
  (all these constraints are using the data alone as if rows are added these constraints might not hold true)

TASK 2 (5 Marks)

Map the following ORM schema to a relational database schema.



Seat (Seat_name, Party_code, [NrOnRoll], PercentVotesCounted, State_Code, NrInformalVotes)

Party (Seat_name, Party_code, Previous_NrFormalVotes, PartyTitle)

Politician (Seat_name, Politician_name, Seat_is_sitting_member, Seat_Seeks, Current_NrFormalVotes, Party_code_ member of)

State (Seat_name, State_code)