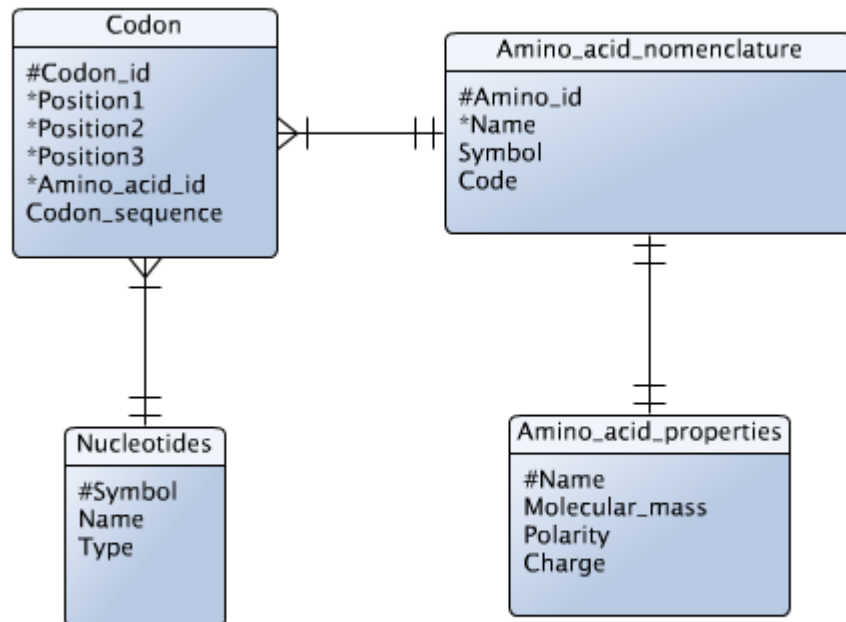


Compulsory exercise 2

Task 1)

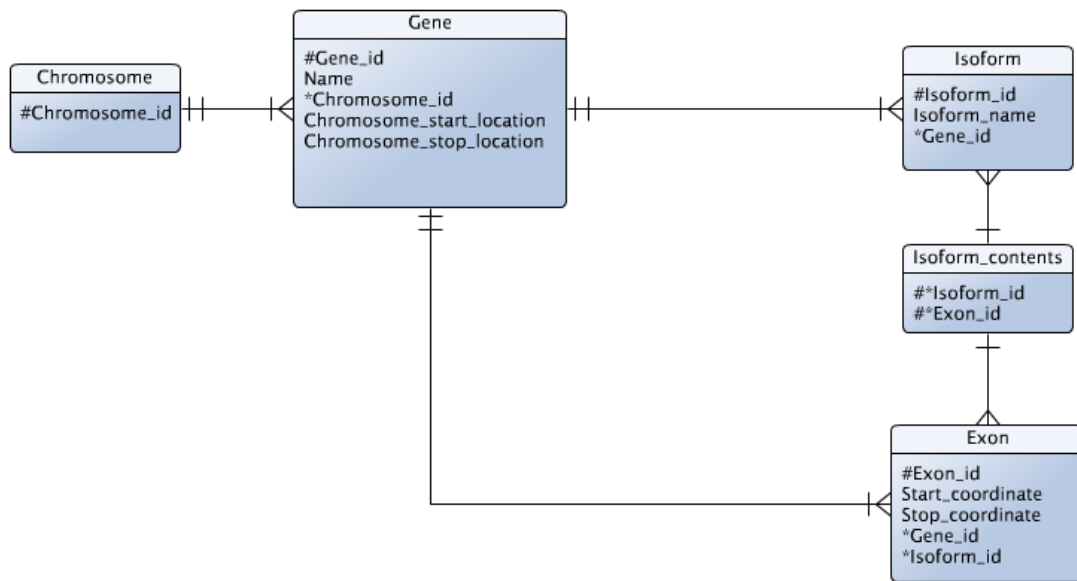


Task 2)

i)

Genes, Exons, Chromosomes and Isoforms are entities. There is also an entity that bridges the Exons and Isoforms, as each Isoform can contain multiple Exons, and each Exon can be inside multiple Isoforms.

ii)



iii)

Chromosome (#Chromosome_id) Gene_location (#*Gene_id, #*Chromosome_id, Chromosome_start_location, Chromosome_stop_location)

Gene (#Gene_id, Gene_name, *Chromosome_id)

Isoform (#Isoform_id, Isoform_name, *Gene_id)

Exon (#Exon_id, Start_coordinate, Stop_coordinate, *Gene_id, *Isoform_id)

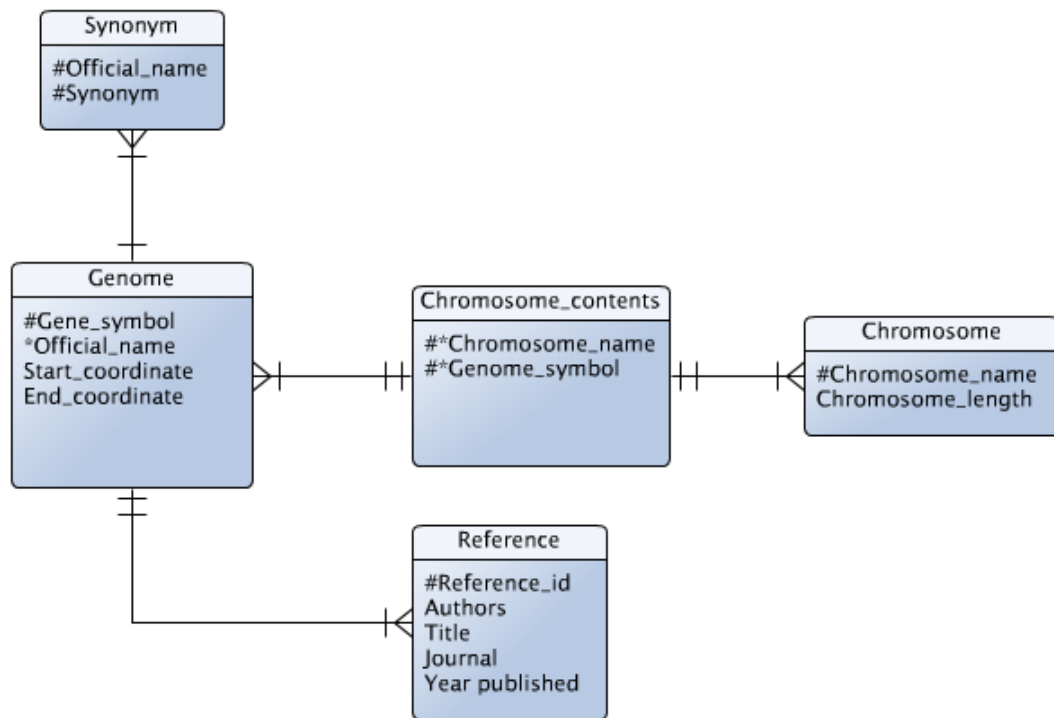
Gene location is moved out of Gene in order to conform to 3rd normal form.

Task 3)

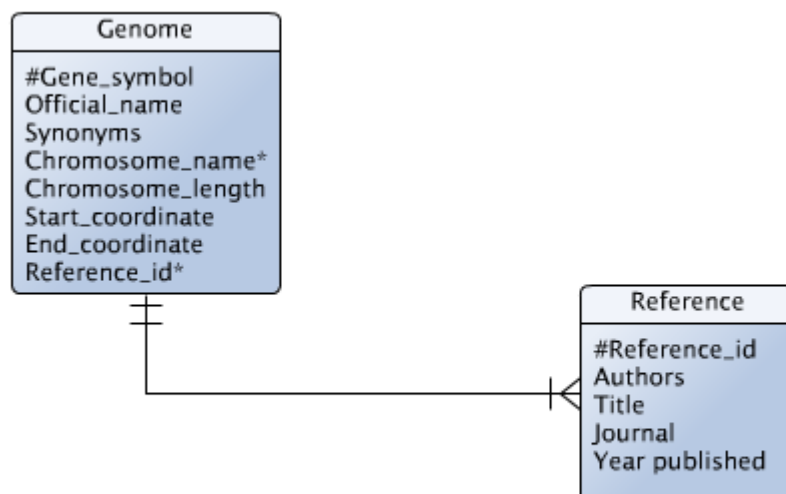
I)

Genome, Chromosome, and Reference are entities, as well as an entity that lists the contents of each Chromosome (assuming a genome is can be found in multiple chromosomes).

ii)

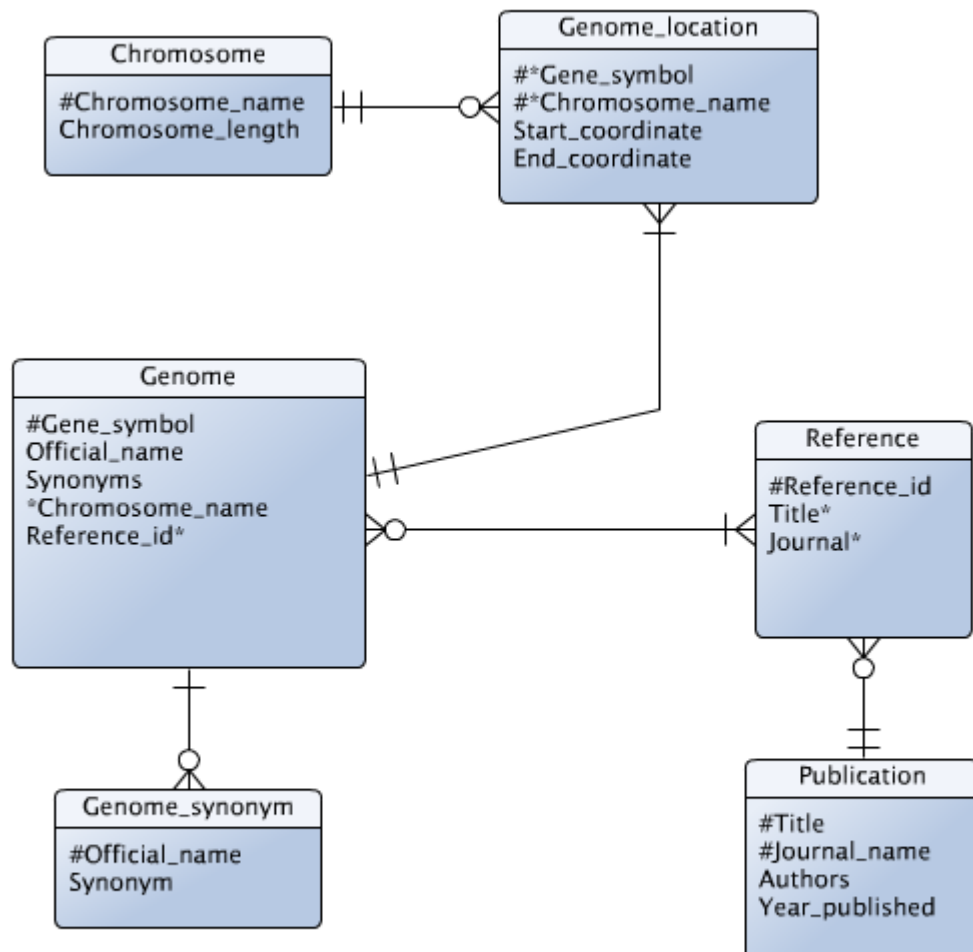


iii)

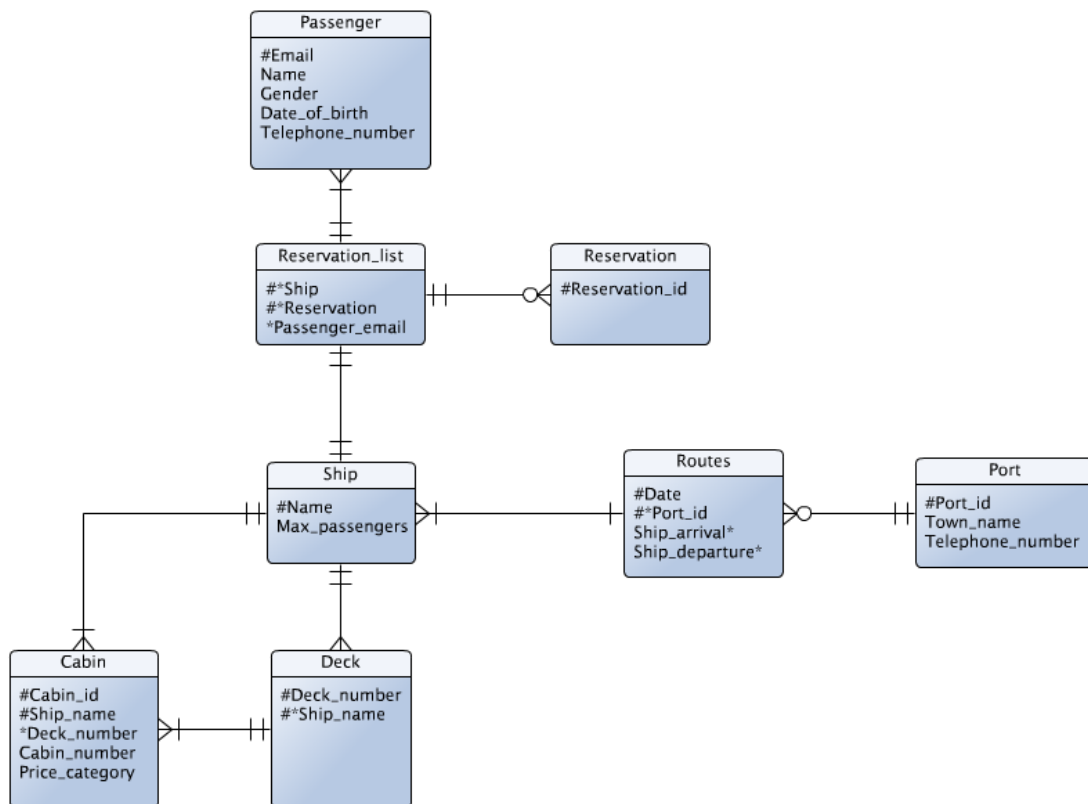


Chromosome now is a part of the Genome entity, and a unique Genome will need both a gene symbol and a chromosome name to be identified. However, chromosome length is information about a subset of the entity, namely chromosome name, and thus, 2nd normal form is violated.

iv)



Task 4)



Task 5)

i)

The problem with the truck table as it is listed is that it doesn't allow a truck to be uniquely identified, because of the assignment number. A truck would most likely be used more than once, and there would be multiple instances of each truck, one for each assignment. Instead, the current assignment of the trucks should be moved to a new table, so that each truck can be identified uniquely, while at the same time be used for more than one assignment.

ii)

Registration_number -> Model Registration_number -> Registration_year

Registration_number -> Maximum_weight Model -> Maximum_weight

Assignment_number -> Registration_number Assignment_number ->

Registration_year Assignment_number -> Model Assignment_number -> Maximum_weight

iii)

In this case Assignment_number is the candidate key. Registration_number isn't necessarily unique as long as the vehicle has been assigned to more than one assignment, and therefore isn't a candidate key unless the table is split up as proposed in 5i).

iv)

Container_type (#Type_id, Type_name, Max_weight, Cubic_quantity*)

Quantity_price (#Cubic_quantity, Nightly_rate)

Container (#Container_number, Type_id*)

Customer (#Telephone_number, Address)

Assignment (#Assignment_number, Telephone_number*, Container_number*, Start_date, End_date)

Truck (#Registration_number, Registration_year, Model, Maximum_weight)

Truck_Assignments (#Assignment_number*, Registration_number*)

The Quantity_price could've been called Weight_price instead, and contained a relationship between Max_weight and Nightly_rate. That would be up to the system the company uses to deduce the rate for each container. I opted for cubic quantity, as that was what seemed most logical to me.