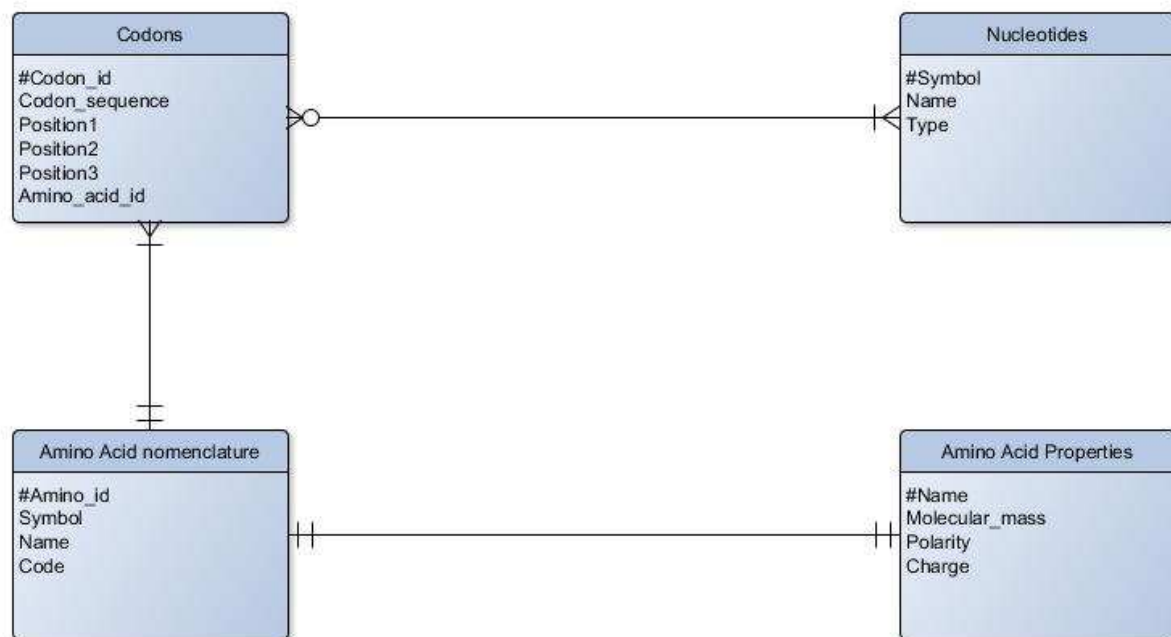
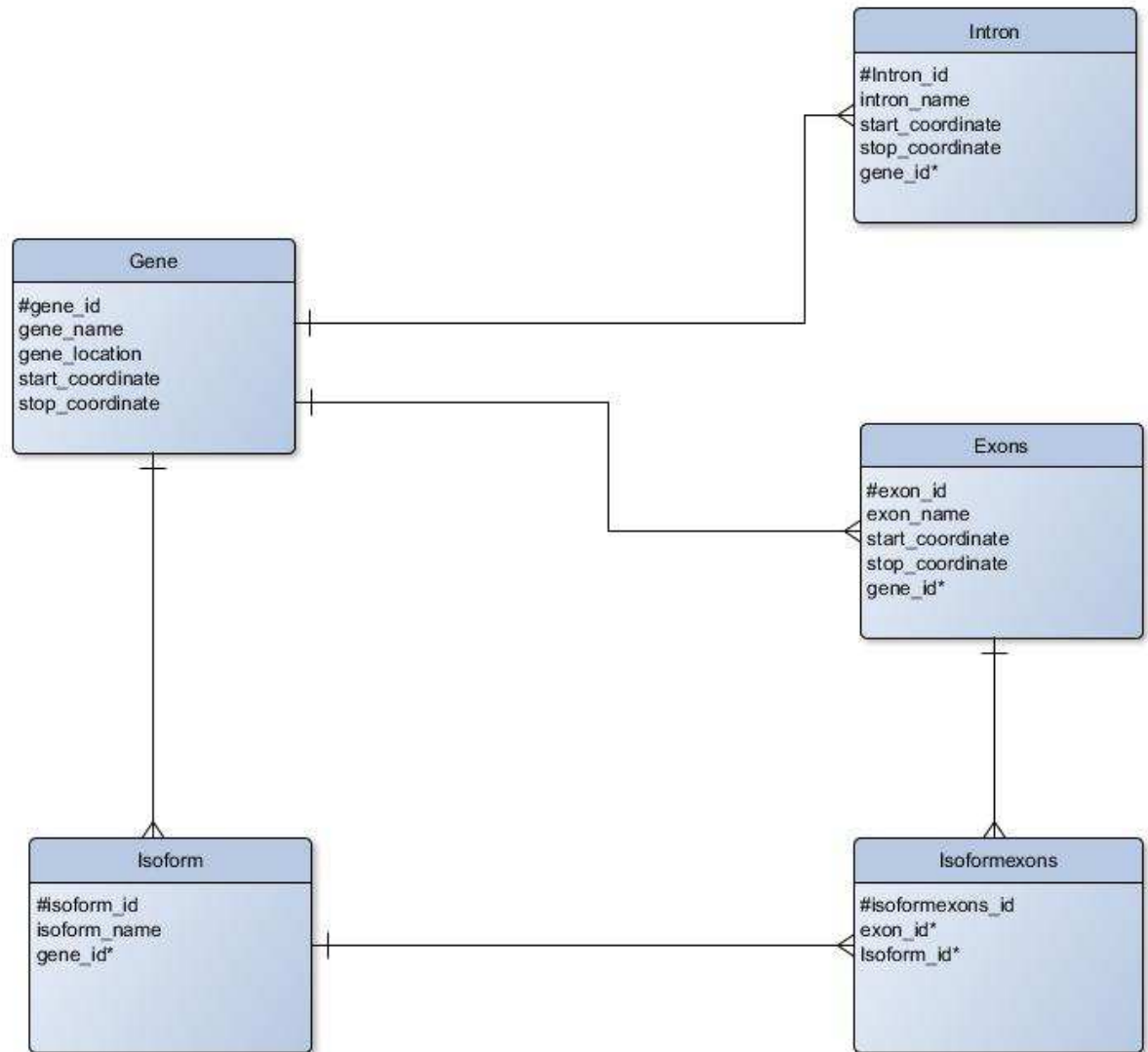


Task 1)**Task 2)**

- i) The entities in the database are: Gene, Exon, Isoform and Intron. I decided to not take Chromosome as an entities here even Gene had an attribute referring to chromosome, because this database was focused on relationship between Gene, Exon, Isoform. I chose to include Intron because it has meaning for how Exons are separated in a gene.
- ii) Genes have one to many relationship with Exons and Isoform, Exons and Isoforms have many to many relationship, which I created a link table "Isoformexons" so we can have two one to many relationship instead of many to many relationship. The diagram is on next page:



- iii) Third normal form (3NF) means within a table, no non-key field is dependent on any other non-key fields. In my ERD from last task do not have any attributes that provokes the 3NF rule.

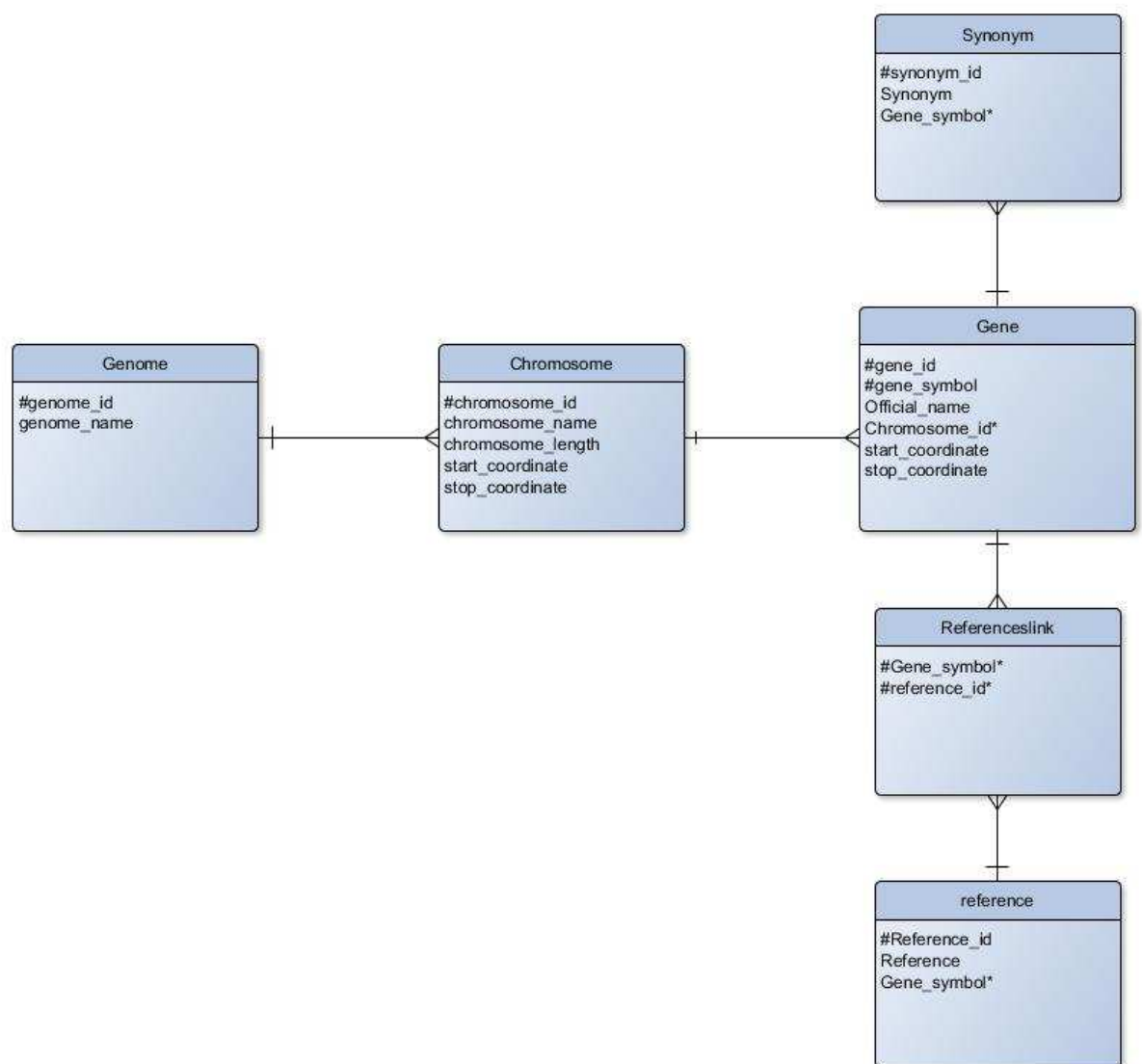
Table:

- **Gene**(#gene_id, gene_name, gene_location, chromosome_name, start_coordinate, stop_coordinate)
- **Exon**(#exon_id, exon_name, start_coordinate, stop_coordinate, gene_id*)
- **isoform**(#isoform_id, isoform_name, gene_id*)
- **isoformexon**(#isoformexons_id, isoform_id*, exon_id*)
- **Intron**(#Intron_id, intron_name, start_coordinate, stop_coordinate, gene_id*)

Task 3

- i) If I had to design a working database for this task, I could also include entities from last task in this one to complete the database. However, since the task do not require us to do that, I decided to make a “new” database just concerning relationships for Genes, chromosomes and genomes. That is why the entities for this task will be *Genomes*, *Chromosomes*, *Genes*, *Synonym*, *Referencelink* and *Refence*. To meet the requirements of the following parts of task 3, there might be some changes to this structure.

ii)



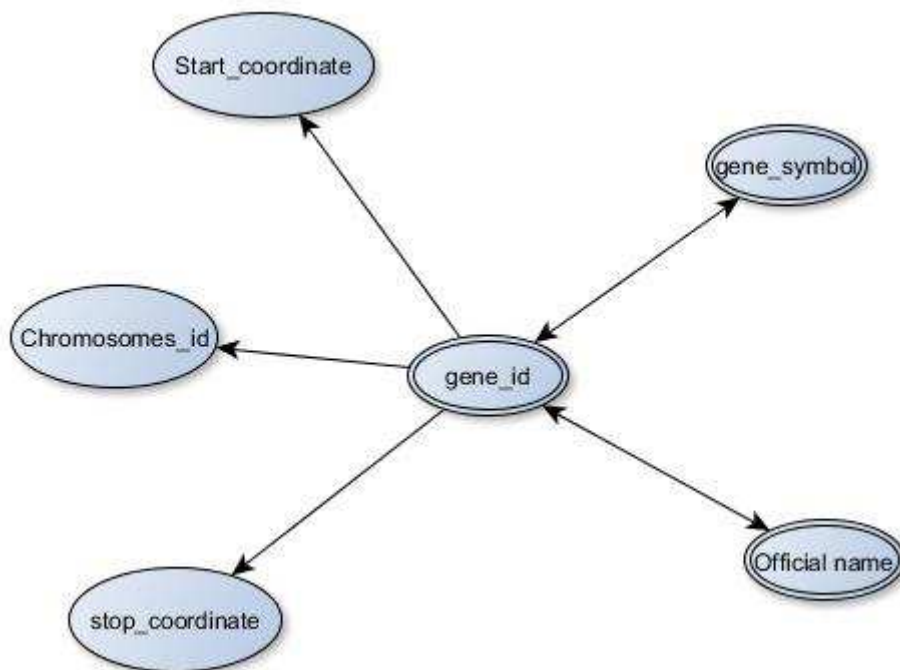
- iii) To conforming to the first normal form, we need to create separate tables for Synonyms and references. However, to not conforming to the second normal form, we have a composite key made by gene_id and gene_symbol in Gene table.

However, you can get any other columns by using either `gene_id` or `gene_symbol`, also `official_name` is not dependent on the entire primary key, in this case the entire composite key. That is why it is not in the second form.

Tables:

- **Genome**(#genome_id, genome_name)
- **Chromosome**(#Chromosome_id, chromosome_name, chromosome_length, start_coordinate, stop_coordinate)
- **Gene**(#gene_id, #gene_symbol, official_name, Chromosome_id*, start_coordinate, stop_coordinate)
- **Referenceslink**(#gene_symbol, #reference_id)
- **Reference**(#Reference_id, reference, gene_symbol*)
- **Synonym**(#synonym_id, synonym, gene_symbol*)

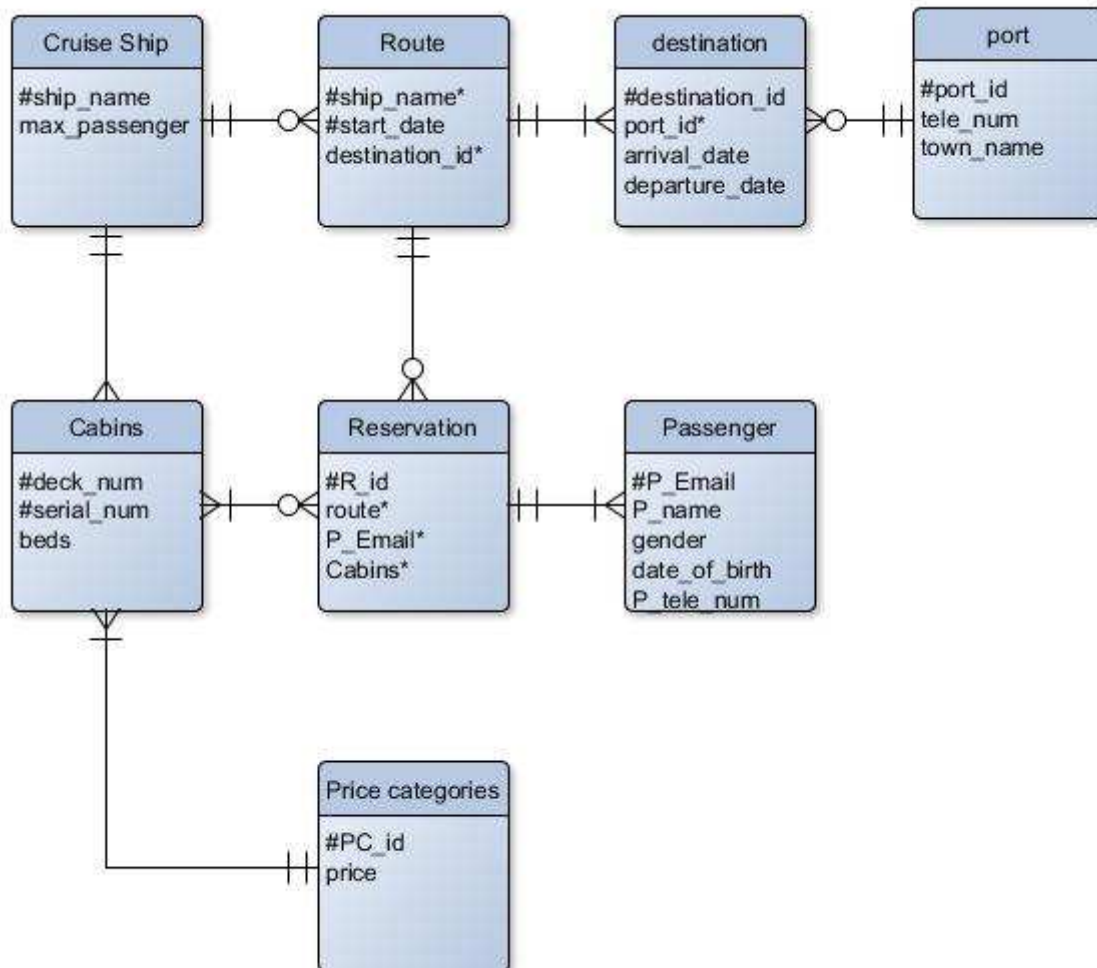
iv)



In this relation graph you can clearly see `gene_id`, `gene_symbol` and `official_name` is the prime attributes here, and `start_coordinate`, `stop_coordinate` and `chromosomes_id` is non-prime attributes. There is no relation between non-prime attributes, which means the 3NF is

obtained. Which means this table is Boyce Codd Normal Form. We did this for every table in the database, and found everything should be Boyce Codd Normal Form.

Task 4.



Task 5)

- i) The first problem is the Truck table have no primary key at all. Second problem is each assignment can have multiple trucks and each trucks can have multiple assignments. This makes a many to many relation between the table, which can be solved by making a new linktable. Thrid "problem" will be the model and maximum_weight in the truck table. Both these columns are non-key columns and

maximum_weight is depending on model, this will be a problem for our table to reach the Third Normal Form.

ii) Functional dependencies:

- a. Register_number \rightarrow register_year
- b. Register_number \rightarrow model
- c. Register_number \rightarrow maximum_weight
- d. Model \rightarrow Maximum_weight

iii) Candidate key in Truck table is Registration_number.

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

Container (#Container_number, Type_id*)

Customer (#Telephone_number, Address)

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

assignmentTruck (#Registration_number*, #Assignment_number*)

Truck (#Registration_number, Registration_year, Model*)

Maximum_weight(#Model, Maximum_weight)