Compulsory assignment 2 INF115   Fredrik Holsten
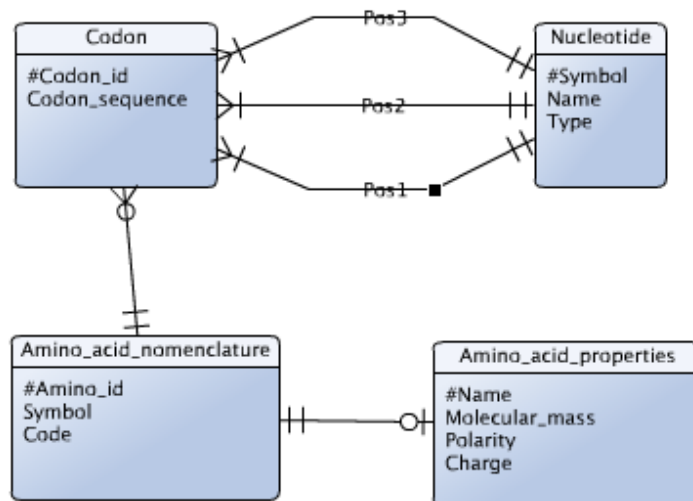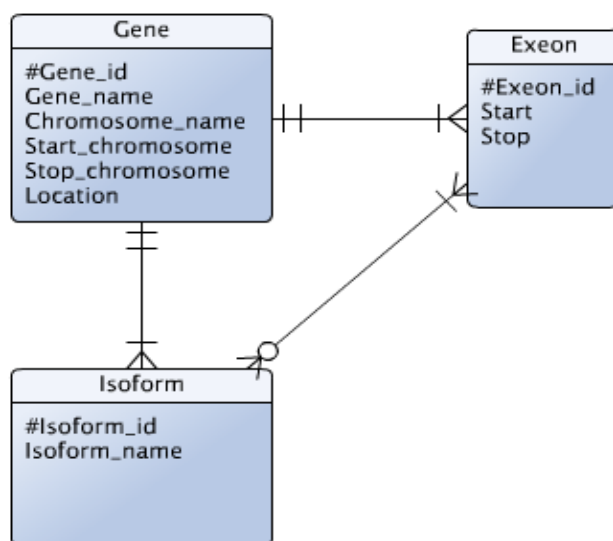
Task 1

Task 2i

The entities are Gene, Exon and Isoform. One can eventually add Chromosome as an entity, as we might be handling several Chromosomes.

Task 2ii
E/R diagram

Task 2iii
Convert to 3NF

Gene(#Gene_id, Gene_name, Chromosome_name, Start_chrom, Stop_chrom, Location)
Isoform(#Isoform_id, Isoform_name, Gene_id*)
Exeon(#Exon_id, Start, Stop, Gene_id*)
IsoExo(#Exon_id, #Isoform_id)

Dependencies:
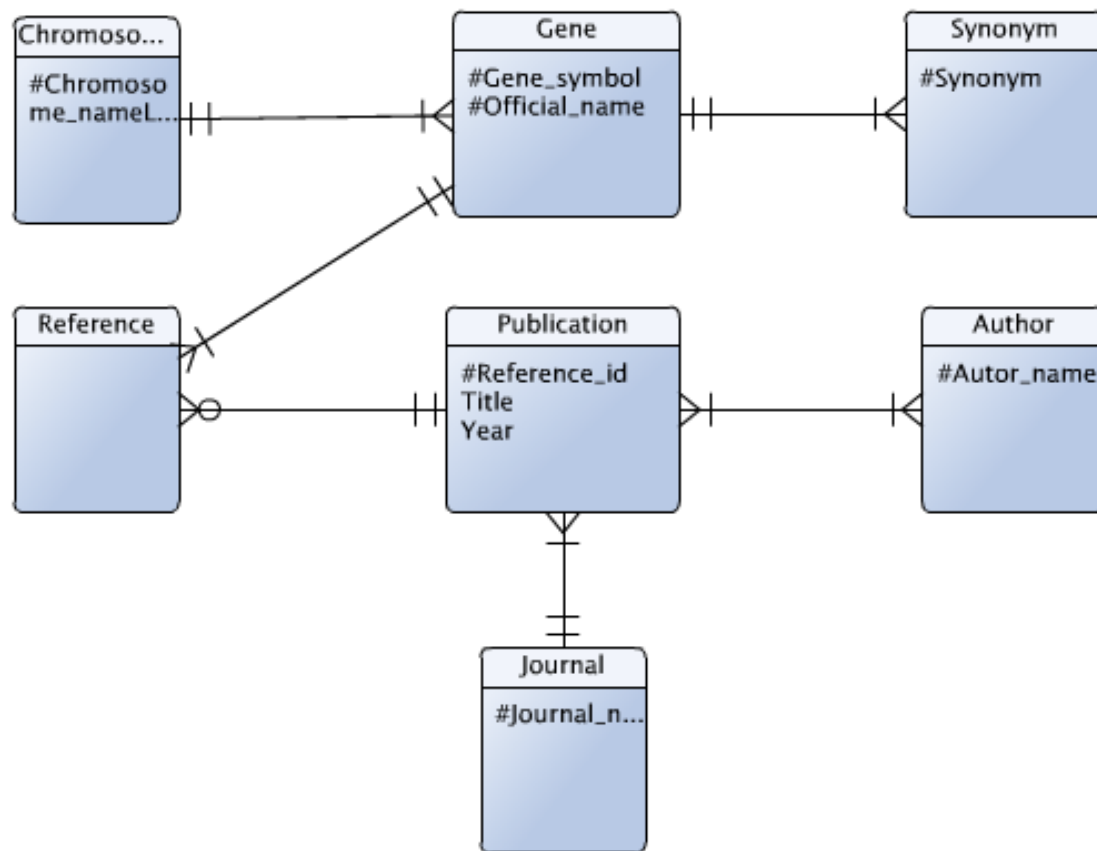Gene: Gene_name → Chromosome_name

Rewriting to 3NF:

Chromosome (#Gene_name, Chromosome_name)
Gene(#Gene_id, Gene_name*, Start_chrom, Stop_chrom, Location)
Isoform(#Isoform_id, Isoform_name, Gene_id*)
Exeon(#Exeon_id, Start, Stop, Gene_id*)
IsoExe(#Exeon_id, #Isoform_id)

Task 3i

Chromosome
Gene
Synonym
Publication
Author
Journal

I have interpreted Journal as a collection of publications.

Task 3ii

Task 3iii
In order to convert this to 1NF, one can put publication year in the references entity, (marked in red):

Chromosome(#Chromosome_name, Length)

Gene(#Gene_symbol, #Official_name, Chromosome_name*)

Synonym(#Syn_name, Official_name*, Gene_symbol*)

Reference(#Official_name, #Gene_symbol, #Reference_id, Year)

Publication(#Reference_id, Title, Journal_name*)

Author(#Reference_id*, #Author_name)

Journal(#Journal_name)

Refererence now contains the dependency Reference_id → Year, where Reference_id is a true subset of the primary key "Official_name + Gene_symbol + Reference_id". This is a violation of the 2NF and is therefore of 1NF.
Here, assumed that neither Gene_symbol nor Official_name is necessarily unique.
Please note that References is a weak entity.

Task 3iv

Converting to BCNF

Chromosome(#Chromosome_name, Length)
Gene(#Gene_symbol, #Official_name, Chromosome_name*)
Synonym(#Syn_name, Official_name*, Gene_symbol*)
Reference(#Official_name, #Gene_symbol, #Reference_id)
Publication(#Reference_id, Title, Journal_name*)
Author(#Author_name, #Reference_id*)
Journal(#Journal_name, year)

Task 4
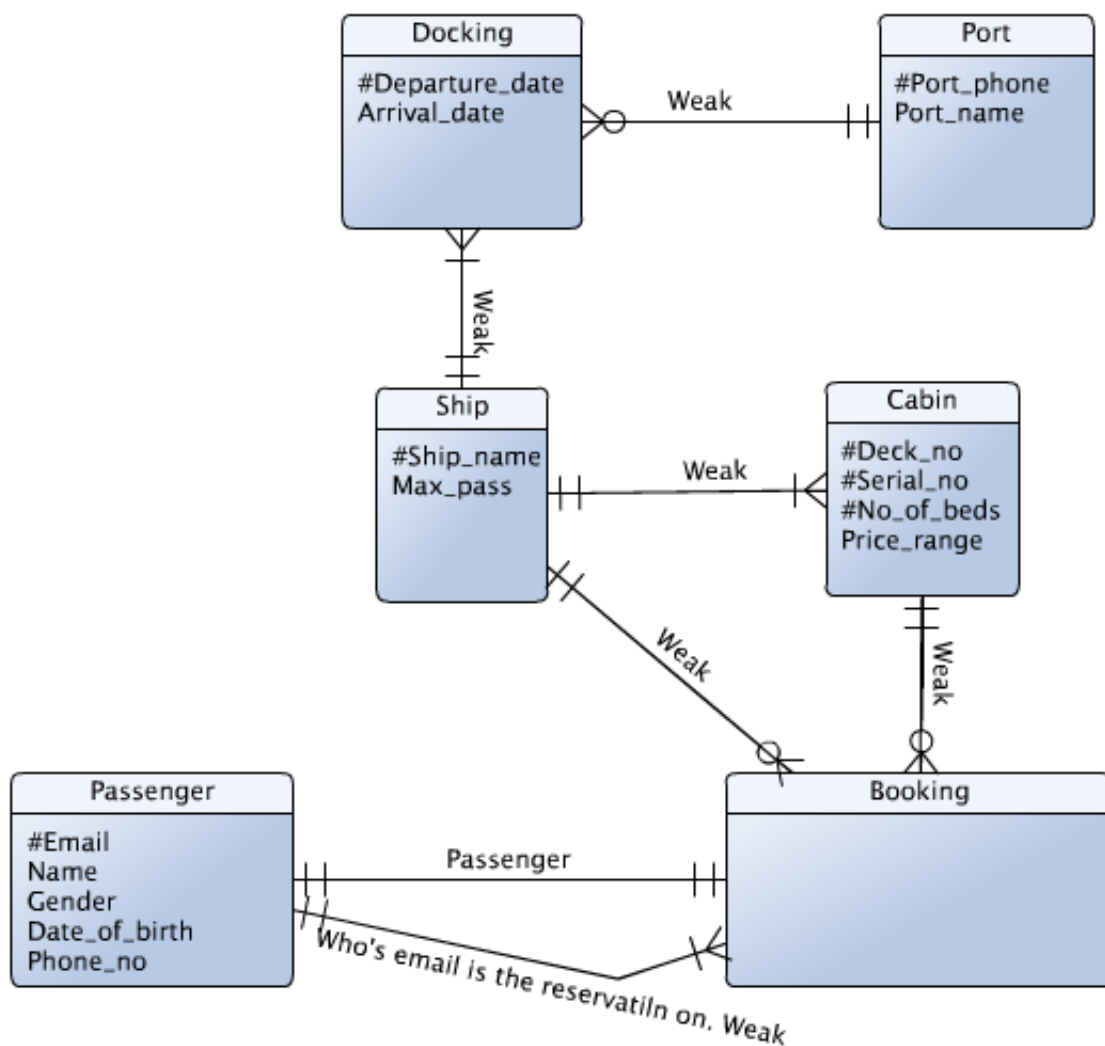
Task 5i

Truck(#Registration_number, Registration_year, Model, Maximum_weight, Assignment_number*)

Truck should be used to keep track of which truck is on which assignment, and not to store permanent information about the truck. We can say that the registration number decides registration year, model and maximum weight. If we put this information in another table, we wont have to repeat all of those attributes every time we mention a specific truck.

Task 5ii

Regisration_number → Registration_year
Regisration_number → Model
Regisration_number → Maximum_weight
Registration_number→ Assignmet_number
Model+Registration_year → Maximum_weight

Task 5iii

The candidate key for Truck is Registration_number, as Registration_number decides all the other attributes and is a minimal super key.

Task 5iv

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)

Truck_assignment(#Registration_number*, Assignment number*)

Truck(#Registration_number, Registration_year*, Model*)

Truck_load(#Registration_year, #Model, Maximum_weight)