CPE 366- 01
Lab 3 Document

PolyDunkers

Team Members :
Aaron Soo
Chun Hui Pek
Elush Shirazpour
Ziyang Tan

Redesign Document:

Entities:

Experiment
- ID (Primary Key)
- date
- cell type
- species (foreign key referencing gene)
- gene abbreviation (foreign key referencing gene)
- from state
- end state

**Gene**
- gene name
- gene abbreviation (Primary Key)
- species (Primary Key)
- begin site
- end site
- chromosome
- regulation

**Promoter_Sequence**
- sequence
- gene abbreviation (Foreign Key referencing gene)
- species (Foreign Key referencing gene)

**Regulatory elements**
- gene abbreviation (Foreign Key referencing gene)
- experiment ID (Foreign Key referencing Experiment)
- beg (Primary Key)
- seq (Primary Key)
- sns (Primary Key)
- mid
- mac

**Comparison**
- mid (Foreign Key referencing regulatory element)
- mac (Foreign Key referencing regulatory element)
- la
- la2
- lq
- ld
- lqv
- sc
- sm
- spv
- ppv

**Model**
- mid (Foreign Key referencing regulatory element)
- mac (Foreign Key referencing regulatory element)
- fac
- fid
- pvalue
- n
- rate

**Weak Entity Sets:**
- Regulatory Element
- Model
- Comparison
- Promoter Sequence

**Hierarchies:**
- Experiment → Gene → Regulatory Element
- Gene → Promoter Sequence
- Experiment → Model → Comparison

**Relationship Sets:**
- Experiment *experiments on* Gene
- Experiment and Gene is *determines* Regulatory Element
- Gene *has* Promoter Sequence

Experiment is *recorded as* Model
- Model is *used for* Comparison

**ER Diagram:**
Change Log:

After speaking with Professor Dekhtyar about the Lab 2 ER diagram, the PolyDunkers group completely changed their ER diagram. The changes are all listed below:
1. **Before:** There was an entity for both species and gene  
   **After:** There is one entity Gene that covers all their attributes from before

2. **Before:** There was an entity for Job Parameters which contained an attribute for comparison and experiments in the past.  
   **After:** Experiment and Comparison our its own entities. Experiment contains attributes of previous experiments and Comparison has the comparison values La, La/, etc.

3. **Before:** There was an entity for promoter.  
   **After:** The promoter entity no longer exists and there now is an entity for Model, which has attributes such as Mid, Mac, etc.

4. **Before:** There was an entity Regulatory Element, which had the comparison attributes.  
   **After:** The Regulatory Element entity still stands, but now the comparison values (La, La/, etc) has been moved to the Comparison entity.

5. **Before:** Promoter sequence was at attribute under the entity Gene.  
   **After:** Promoter Sequence is now a weak entity for Gene.

6. Since so many entities have changed, the relations among them have completely changed. The biggest change to the relations is that Experiment and Gene share a relationship with Regulatory element.

**Database Constraints :**

- A regulatory sequence string must match the substring of the promoter sequence for the same gene that starts at the same position indicated, and goes in the direction specified by the sense attribute
- A regulatory sequence cannot start at a position if the position value is greater than the length of the promoter sequence.
- The sum of the position and the regulatory sequence cannot be greater than the length of the promoter sequence
- The gene species must match the species the experiment was performed on
- Chromosome value must be valid
- FACs are only attribute that can be NULL

**Logical Database Design:**

**Experiment**  
- ID (Primary Key)  
- date  
- cell type  
- species (foreign key referencing gene)  
- gene abbreviation (foreign key referencing gene)  
- from state  
- end state
**Gene**
- gene name
- gene abbreviation (Primary Key)
- species (Primary Key)
- begin site
- end site
- chromosome
- regulation

**Promoter_Sequence**
- sequence
- gene abbreviation (Foreign Key referencing gene)
- species (Foreign Key referencing gene)

**Regulatory elements**
- gene abbreviation (Foreign Key referencing gene)
- experiment ID (Foreign Key referencing Experiment)
- beg (Primary Key)
- seq (Primary Key)
- sns (Primary Key)
- mid
- mac

**Comparison**
- mid (Foreign Key referencing regulatory element)
- mac (Foreign Key referencing regulatory element)
- la
- la2
- lq
- ld
- lqv
- sc
- sm
- spv
- ppv

**Model**
- mid (Foreign Key referencing regulatory element)
- mac (Foreign Key referencing regulatory element)
- fac
- fid
- pvalue
- n
- rate

**Creating the Database:**

remark PolyDunkers
CREATE TABLE gene (  
gene_abbr VARCHAR2(10) UNIQUE,  
name VARCHAR2(50),  
species VARCHAR2(15) UNIQUE,  
Begin_site int,  
End_site int,  
chromosome int,  
regulation VARCHAR2(20),  
PRIMARY KEY (gene_abbr, species)  );

CREATE TABLE promoter_sequence (  
promoter_sequence VARCHAR2(2000) PRIMARY KEY,  
species VARCHAR2(15),  
gene_abbr VARCHAR2(10),  
constraint fk1 FOREIGN KEY (species) REFERENCES gene(species),  
constraint fk2 FOREIGN KEY (gene_abbr) REFERENCES gene(gene_abbr)  );

CREATE TABLE experiment (  
id VARCHAR2(20) PRIMARY KEY,  
e_date DATE,  
cell_type VARCHAR2(20),  
species VARCHAR2(15),  
gene_abbr VARCHAR2(10),  
from_state VARCHAR2(25),  
to_state VARCHAR2(25),  
constraint sid FOREIGN KEY (species) REFERENCES gene(species),  
constraint gbd FOREIGN KEY (gene_abbr) REFERENCES gene(gene_abbr)  );

CREATE TABLE regulatory_elements (  
g_abbr VARCHAR2(5) REFERENCES gene(gene_abbr),  
ex_id VARCHAR2(20) REFERENCES experiment(id),  
beg int,  
len int,  
seq VARCHAR2(50),  
sns VARCHAR2(1),  
mid VARCHAR2(15) UNIQUE,  
mac VARCHAR2(15) UNIQUE,  
PRIMARY KEY(beg, len, seq, sns)  );

CREATE TABLE comparison (  
mid VARCHAR2(15),  
mac VARCHAR2(15),  
l_a FLOAT,  
l_a2 FLOAT,  
l_q FLOAT,  
l_d FLOAT,  
l_qv FLOAT,  
...
sc FLOAT,
sm FLOAT,
spv FLOAT,
ppv FLOAT,
PRIMARY KEY(la, la2, lq, ld, lqv, sc, sm, spv, ppv),
constraint mymid FOREIGN KEY (mid) REFERENCES regulatory_elements(mid),
constraint mymad FOREIGN KEY (mac) REFERENCES regulatory_elements(mac)
);

CREATE TABLE model (  
    mid VARCHAR2(15),
    mac VARCHAR2(15),
    fac VARCHAR2(15),
    fid VARCHAR2(10),
    pvalue FLOAT,
    n FLOAT,
    rate VARCHAR2(15),
    constraint modle_mid FOREIGN KEY (mid) REFERENCES regulatory_elements(mid),
    constraint modle_mad FOREIGN KEY (mac) REFERENCES regulatory_elements(mac)
    );