CREAZIONE TABLES

/\* primo blocco di entità

CREATE TABLE Researcher(

researcher\_id INT IDENTITY(1,1) PRIMARY KEY,

researcher\_name VARCHAR(20),

researcher\_surname VARCHAR(20),

field\_of\_studies VARCHAR(30)

)

alter table Researcher

add is\_boss BIT default 0

CREATE TABLE Institute(

institute\_id INT PRIMARY KEY,

location VARCHAR(30),

boss\_id INT

)

CREATE TABLE Researchers\_Institutes(

researcher\_id INT,

institute\_id INT

)

\*/

/\* relazioni del primo blocco

ALTER TABLE Researchers\_Institutes

ADD FOREIGN KEY (researcher\_id) REFERENCES Researcher(researcher\_id)

ALTER TABLE Researchers\_Institutes

ADD FOREIGN KEY (institute\_id) REFERENCES Institute(institute\_id)

\*/

/\* secondo blocco di entità

CREATE TABLE Research(

research\_id INT PRIMARY KEY,

organism\_id INT,

budget FLOAT,

objective VARCHAR(100)

)

CREATE TABLE Institute\_Researches(

research\_id INT,

institute\_id INT

)

\*/

/\* relazioni secondo blocco

ALTER TABLE Institute\_Researches

ADD FOREIGN KEY (institute\_id) REFERENCES Institute(institute\_id)

ALTER TABLE Institute\_Researches

ADD FOREIGN KEY (research\_id) REFERENCES Research(research\_id)

\*/

/\* terzo blocco di entità

CREATE TABLE Organism(

organism\_id INT PRIMARY KEY,

scientific\_name VARCHAR(50),

dominium VARCHAR(20)

)

CREATE TABLE Genome(

organism\_id INT,

sequence\_ VARCHAR(100)

)

\*/

/\* relazioni terzo blocco

ALTER TABLE Genome

ADD FOREIGN KEY (organism\_id) REFERENCES Organism(organism\_id)

ALTER TABLE Research

ADD FOREIGN KEY (organism\_id) REFERENCES Organism(organism\_id)

\*/

/\* quarto blocco entità

CREATE TABLE SequenceHighlight(

sequence\_ VARCHAR(100) PRIMARY KEY,

highlight\_prot\_id INT

)

CREATE TABLE Protein(

protein\_id INT PRIMARY KEY,

functionality VARCHAR(100),

protein\_name VARCHAR(50)

)

\*/

/\* relazioni quarto blocco

ALTER TABLE Genome

ADD FOREIGN KEY (sequence\_) REFERENCES SequenceHighlight(sequence\_)

ALTER TABLE SequenceHighlight

ADD FOREIGN KEY (highlight\_prot\_id) REFERENCES Protein(protein\_id)

\*/

CREAZIONE STORED PROCEDURES

/\*

CREATE PROCEDURE insertResearcher

@name VARCHAR(20),

@surname VARCHAR(20),

@field VARCHAR(30)

AS

BEGIN

BEGIN TRY

INSERT INTO Researcher (researcher\_name, researcher\_surname, field\_of\_studies)

VALUES (@name, @surname,@field)

END TRY

BEGIN CATCH

SELECT ERROR\_MESSAGE() as 'An error occurred'

END CATCH

END

\*/

/\*

CREATE PROCEDURE insertInstitute

@id INT,

@location Varchar(20),

@boss INT

AS

BEGIN

INSERT INTO [dbo].[Institute] (institute\_id, location, boss\_id)

VALUES (@id, @location, @boss)

END

\*/

/\*

CREATE PROCEDURE insertResearch

@id INT,

@orgid INT,

@budget FLOAT,

@objective VARCHAR(100)

AS

BEGIN

INSERT INTO [dbo].[Research](research\_id, organism\_id, budget,objective)

VALUES (@id, @orgid, @budget,@objective)

END

\*/

/\*

CREATE PROCEDURE insertOrganism

@id INT,

@name VARCHAR(50),

@dominium VARCHAR(20)

AS

BEGIN

INSERT INTO Organism(organism\_id, scientific\_name, dominium)

VALUES (@id, @name, @dominium)

END

\*/

/\*

CREATE PROCEDURE insertGenome

@id INT,

@sequence VARCHAR(100)

AS

BEGIN

INSERT INTO [dbo].[Genome](organism\_id, sequence\_)

VALUES (@id, @sequence)

END

\*/

/\*

CREATE PROCEDURE insertSeqHigh

@sequence VARCHAR(100),

@highProt INT

AS

BEGIN

INSERT INTO [dbo].[SequenceHighlight](sequence\_, highlight\_prot\_id)

VALUES (@sequence, @highProt)

END

\*/

/\*

CREATE PROCEDURE insertProtein

@id INT,

@functionality VARCHAR(100),

@name VARCHAR(50)

AS

BEGIN

INSERT INTO [dbo].[Protein](protein\_id,functionality,protein\_name)

VALUES (@id, @functionality,@name)

END

\*/

--inserire prima tutte le Proteine (si viene a creare così una libreria proteica), poi si inserisce la sequenza di highlight che

-- contiene la codifica di tale proteina, in fine il genoma rappresentato da tale sequenza (che risulterà associato ad un certo organismo)

--stesso concetto viene applicato dalla relazione Ricerca-Organismo : prima viene realizzata la libreria degli organismi, poi si crea una entità

--Ricerca scegliendo appunto un organismo da tale libreria

CREAZIONE TRIGGHER

/\*

CREATE TRIGGER NewBoss

ON Institute

AFTER INSERT, UPDATE

AS

BEGIN

update Researcher

set is\_boss=1

where researcher\_id=(

select boss\_id

from Institute

join Researchers\_Insitutes

on Institute.institute\_id=Researchers\_Insitutes.institute\_id

join Researcher

on Researcher.researcher\_id=Researchers\_Insitutes.researcher\_id

where Institute.boss\_id=Researcher.researcher\_id and Researcher.is\_boss=0)

END

\*/

CREAZIONE FUNCTION

/\*

CREATE FUNCTION avg\_seq\_length ()

RETURNS FLOAT AS

BEGIN

DECLARE @avg\_length float;

set @avg\_length = (SELECT AVG(LEN(sequence\_)) from Genome)

RETURN @avg\_length;

END;

\*/

INSERT VALUES

/\*

exec insertResearcher 'stephen', 'okazaky' , 'genetist'

exec insertResearcher 'mike','zane', 'computer scientist'

exec insertResearcher 'bob' , 'curry' , 'molecular biologist'

exec insertResearcher 'martha' , 'hillberg', 'software engineer'

exec insertResearcher 'diego' , 'rossi' , 'chemist'

exec insertResearcher 'sofia' , 'giannoni' , 'biophisicist'

exec insertInstitute 1,'california', 1

exec insertInstitute 2,'milano', 2

exec insertInstitute 3, 'Warsaw' , 3

insert into Researchers\_Institutes

values (1,1)

insert into Researchers\_Institutes

values (2,2)

insert into Researchers\_Institutes

values (3,3)

insert into Researchers\_Institutes

values (4,1)

insert into Researchers\_Institutes

values (5,2)

insert into Researchers\_Institutes

values (6,3)

exec insertOrganism 1 , 'saccharomyces cerevisiae', 'aukaryota'

exec insertOrganism 2, 'Escherichia coli','prokatyota'

exec insertOrganism 3, 'Streptococcus pneumoniae', 'prokaryota'

exec insertOrganism 4 , 'Lactobacillus acidophilus' , 'prokaryota'

exec insertOrganism 5, 'Clostridium difficile' , 'prokaryota'

exec insertOrganism 6, 'Homo sapiens','eukaryota'

exec insertOrganism 7, 'Saccharomyces cerevisiae' , 'eukaryota'

exec insertOrganism 8, 'Drosophila melanogaster', 'eukaryota'

exec insertOrganism 9, 'Paramecium aurelia' , 'eukaryota'

exec insertOrganism 10 , 'Amoeba proteus' , 'eukaryota'

exec insertResearch 100, 2, 5500.22, 'make it unable to replicate'

exec insertResearch 101,5,10000,'squence the full human genome'

exec insertResearch 102,1,8000000,'engeneer the fungi to produce more CH3CH2OH with less sugars'

insert into Institute\_Researches

values (100,1)

insert into Institute\_Researches

values (101,2)

insert into Institute\_Researches

values (102,3)

insert into Institute\_Researches

values (101,1)

exec insertProtein 66,'carries oxygen in red blood cells','hemoglobin'

exec insertProtein 67,'provides structural support in connective tissue','collagen'

exec insertProtein 68 , 'regulates blood sugar levels', 'insulin'

exec insertProtein 69, 'protein involved in bacterial cell division' , 'FtsZ'

exec insertProtein 70 , 'protein that assists in the folding of other proteins' , 'chaperonin'

exec insertProtein 71 , 'protein involved in DNA repair and recombination', 'RecA'

exec insertSeqHigh 'AUGCCCUGA',69

exec insertSeqHigh 'AUGGGGUGA', 70

exec insertSeqHigh 'AUGTTTUGA', 68

exec insertGenome 1 , 'AUGGGGUGA'

exec insertGenome 2, 'AUGCCCUGA'

exec insertGenome 6 , 'AUGTTTUGA'

\*/

QUARIES

--i ricercatori che sono situati in un centro di ricerca in California

/\*

select researcher\_name, researcher\_surname

from Researcher

join Researchers\_Institutes

on Researchers\_Institutes.researcher\_id=Researcher.researcher\_id

join Institute

on Institute.institute\_id=Researchers\_Institutes.institute\_id

where Institute.location='california'

\*/

--la lunghezza media delle catene associate a proteine di highlight

/\*

select dbo.avg\_seq\_length() as AVG\_CHAIN\_LENGTH

\*/

--la ricerca col maggior budget

/\*

select top 1 \*

from Research

order by budget

\*/

--gli organismi attualmente soggetti ad una ricerca che hanno una proteina di highlight associata presente nella tabella di tutte le proteine del database

/\*

select Organism.organism\_id, Organism.scientific\_name

from Organism

join Genome

on Organism.organism\_id=Genome.organism\_id

join SequenceHighlight

on SequenceHighlight.sequence\_=Genome.sequence\_

join Protein

on Protein.protein\_id=highlight\_prot\_id

where Protein.protein\_name in (select distinct Protein.protein\_name from Protein)

group by Organism.organism\_id, Organism.scientific\_name

having Organism.organism\_id in (select Research.organism\_id from Research)

\*/

/\* seleziona il boss di ogni ricerca con il corrispondente obbiettivo della ricerca (per sbaglio ho masso due Boss di instituto a lavorare sulla stessa

ricerca allora ci sono due boss per il sequenziamento del genoma umano

select Researcher.researcher\_name, Researcher.researcher\_surname, table\_a.objective

from Researcher

join(select distinct Research.research\_id, Research.objective, Institute.boss\_id

from Research

join Institute\_Researches

on Institute\_Researches.research\_id=Research.research\_id

join Institute

on Institute.institute\_id=Institute\_Researches.institute\_id) as table\_a

on table\_a.boss\_id=Researcher.researcher\_id

where table\_a.research\_id in (100,101,102)

\*/

/\* gli organismi nel database che hanno una proteina nel database

select Organism.scientific\_name

from Organism

join(select Protein.protein\_id, Genome.organism\_id from Protein

join SequenceHighlight

on SequenceHighlight.highlight\_prot\_id=Protein.protein\_id

join Genome

on Genome.sequence\_=SequenceHighlight.sequence\_) as table\_b

on table\_b.organism\_id=Organism.organism\_id

where table\_b.protein\_id in (select Protein.protein\_id from Protein)

\*/