

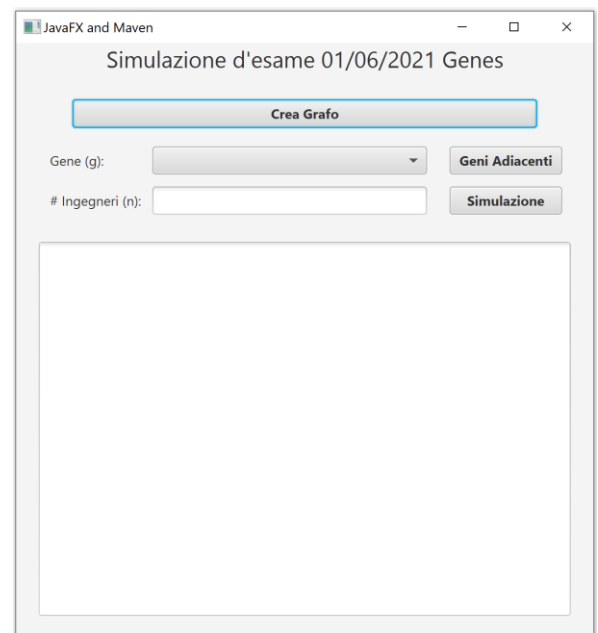
Simulazione d'esame 01/06/2021

Si consideri il database **genes_small.sql**, estratto da un database del genoma umano originariamente creato per una challenge internazionale di biological data mining. I geni codificano le proteine che si localizzano in diverse parti della cellula e interagiscono fra di loro per eseguire le funzioni vitali della cellula stessa. Il dataset contiene informazioni sulla localizzazione (tabella *classification*), sulle interazioni (tabella *interactions*) e sulle funzioni dei geni (tabella *genes*), ed è strutturato secondo il diagramma ER illustrato nella pagina seguente.

Si intende costruire un'applicazione JavaFX che permetta di svolgere le seguenti funzioni:

PUNTO 1

- Alla pressione del bottone "Crea Grafo" si crei un grafo semplice, pesato e non orientato i cui vertici siano tutti i geni essenziali (tabella *genes*, campo *Essential*).
- Un arco collega due geni diversi solo se tale coppia (indipendentemente dall'ordine) appare nella tabella *interactions*. Si ignorino gli archi di tipo "cappio", cioè le connessioni di un gene con se stesso.
- Per definire il peso dell'arco si parta dalla correlazione fra i geni (tabella *interactions*, campo *Expression_Corr*) e si consideri il peso
 - pari al valore assoluto di tale correlazione se i due geni non appartengono allo stesso cromosoma (campo *Chromosome* della tabella *genes*)
 - pari al doppio del valore assoluto di tale correlazione se i due geni appartengono allo stesso cromosoma.
- Permettere all'utente di selezionare dal menù a tendina uno dei geni (g) presenti nel grafo.
- Alla pressione del bottone "Geni Adiacenti" stampare l'elenco dei geni adiacenti a quello selezionato ed il relativo peso, in ordine decrescente di peso.



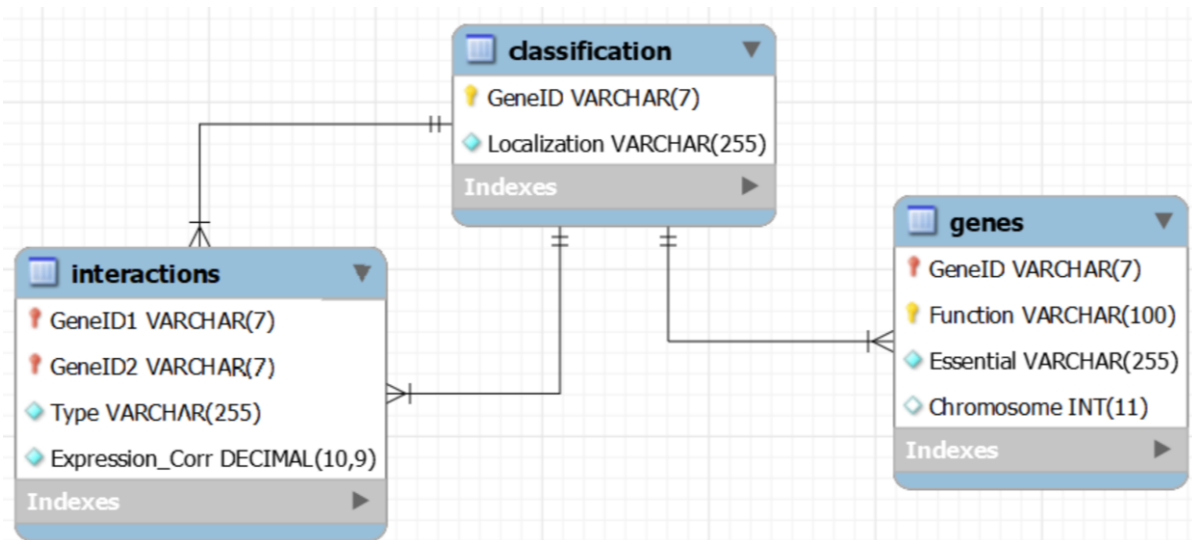
PUNTO 2

- A partire dal grafo calcolato al punto precedente, si avvia una simulazione di un progetto di ricerca di tre anni che coinvolge n bioingegneri per lo studio dei geni. Il parametro n è scelto dall'utente nell'apposita casella di testo.
- Tutti gli n ingegneri partono analizzando le caratteristiche del gene selezionato dall'utente al punto 1.d.
- All'inizio di ogni mese, ciascun ingegnere può decidere :
 - con il 30% di probabilità di continuare a studiare lo stesso gene;
 - con il 70% di probabilità di focalizzarsi sullo studio di uno dei geni adiacenti. Più precisamente, la probabilità di scegliere un certo gene dipende dal peso dell'arco che lo connette con il gene precedente. (Esempio: se il gene precedente g_1 è connesso a g_2 , g_3 e g_4 , la probabilità di spostarsi da g_1 a g_2 sarà data dal peso dell'arco g_1 - g_2 diviso per la somma dei pesi degli archi g_1 - g_2 , g_1 - g_3 , g_1 - g_4).
- Alla fine della simulazione identificare quali geni sono in corso di studio ed il rispettivo numero di ingegneri ad essi associati.

Nella realizzazione del codice, si lavora a partire dalle classi (Bean e DAO, FXML) e dal database contenuti nel progetto di base. È ovviamente permesso aggiungere o modificare classi e metodi.

Tutti i possibili errori di immissione, validazione dati, accesso al database, ed algoritmici devono essere gestiti, non sono ammesse eccezioni generate dal programma. Nelle pagine seguenti, sono disponibili due esempi di risultati per controllare la propria soluzione.

Le tabelle **classification**, **interactions** e **genes** contengono rispettivamente informazioni sulla localizzazione dei geni (e delle rispettive proteine), sul tipo e intensità del rapporto di interazione, e sulla funzione da loro espressa. Tutte le tabelle sono collegate tra loro tramite la chiave primaria che identifica i geni (GeneID).



ESEMPI DI RISULTATI PER CONTROLLARE LA PROPRIA SOLUZIONE:

Simulazione d'esame 01/06/2021 Genes

Crea Grafo

Gene (g):

Ingegneri (n):

Grafo creato con 286 vertici e 213 archi

Geni adiacenti a: G234074
G234735 0.630195718

Geni adiacenti a: G234103
G235336 0.507155643
G235080 0.204707401

Geni adiacenti a: G234141
G234341 0.833400501
G234533 0.564706878

Geni adiacenti a: G234194
G235229 1.679959588
G239123 0.640012496
G235129 0.331438725