

## Σχεδίαση και Χρήση Βάσεων Δεδομένων - Εαρινό Εξάμηνο 2024

### 3<sup>η</sup> Άσκηση

Παράδοση Άσκησης: Δευτέρα 10 Ιουνίου, Ώρα: 23:55

Σας δίνεται μια εφαρμογή τριών επιπέδων. Αποτελείται από τη διεπαφή χρήστη, που είναι web based, τη λογική της εφαρμογής, που είναι σε Python, και τη βάση δεδομένων, που είναι σε MySQL (το σχήμα της βάσης και το dataset είναι τα ίδια με αυτά που χρησιμοποιήσατε στην Άσκηση 2).

### Οδηγίες

Για να την τρέξετε, θα πρέπει να κάνετε τα εξής:

- να μελετήσετε τις διαφάνειες του φροντιστηρίου για την άσκηση αυτή (εκεί θα βρείτε οδηγίες εγκατάστασης για την python 3.xx, καθώς και βασικά παραδείγματα εξοικείωσης με μια web-based εφαρμογή που χρησιμοποιεί ένα web framework που για αυτή την εργασία θα χρησιμοποιηθεί το bottle).
- να κάνετε unzip το application.zip.
- να αλλάξετε το αρχείο settings.py και να βάλετε τα στοιχεία σας για να μπορείτε να συνδεθείτε στη βάση MySQL.
- Στη συνέχεια να τρέξετε το website.py με την python.
- Τέλος, να ανοίξετε κάποιον browser και να βάλετε τη διεύθυνση “<http://localhost:8080>”

Το παρακάτω είναι η αρχική σελίδα που πρέπει να δείτε:

<b>Update Rank</b> Rank1: <input type="text"/> Rank2: <input type="text"/> Movie title: <input type="text"/> <input type="button" value="Update"/>	<b>Colleagues of Colleagues</b> Actor Id 1: <input type="text"/> Actor Id 2: <input type="text"/> <input type="button" value="Search"/>
<b>Actor Pairs</b> Actor Id: <input type="text"/> <input type="button" value="Search"/>	<b>Top N Actors</b> N: <input type="text"/> <input type="button" value="Search"/>
<b>Trace Actor Influence</b> Actor Id: <input type="text"/> <input type="button" value="Search"/>	

Αυτό που καλείστε να κάνετε είναι να αλλάξετε τη λογική της εφαρμογής η οποία βρίσκεται στο **app.py** έτσι ώστε να εκτελεί τα παρακάτω ζητούμενα. Όλες οι συναρτήσεις μέσα στο αρχείο app.py επιστρέφουν μια **λίστα από πλειάδες (tuples)** όπου πάντα η πρώτη πλειάδα είναι η κεφαλίδα με τα ονόματα των πεδίων και οι υπόλοιπες είναι τα αποτελέσματα.

Για παράδειγμα: [(“Name”, “Id”), (“Jim”, 7), (“Tom”, 13,)]

## Περιγραφή των συναρτήσεων

1. **updateRank:** Η συνάρτηση αυτή παίρνει από το χρήστη ως όρισμα δύο αριθμητικές τιμές (Rank1 & Rank2 - double από 0 μέχρι και 10) και το όνομα μιας ταινίας (Movie title) και ενημερώνει το πεδίο rank της ταινίας αυτής με το μέσο όρο που προκύπτει από την τρέχουσα τιμή της στο πεδίο αυτό (αν υπάρχει) και των δύο τιμών που έδωσε ο χρήστης. Αν η τιμή rank απουσιάζει τότε η τιμή θα ενημερώνεται με τον μέσο όρο των δύο τιμών που έδωσε ο χρήστης. Σε περίπτωση επιτυχίας, επιστρέφει **“ok”**. Αντίθετα, αν δεν υπάρχει ταινία με αυτό το όνομα, ή αν υπάρχουν περισσότερες από μια ταινίες με το ίδιο όνομα, ή αν οι αριθμητικές τιμές που εισάγει ο χρήστης είναι εκτός ορίων, επιστρέφει **“error”**.
2. **colleaguesOfColleagues:** Η συνάρτηση αυτή παίρνει ως όρισμα ένα ζευγάρι (a,b) από κωδικούς ηθοποιών (Actor Id 1, Actor Id 2). Βρίσκει τις ταινίες που παίζουν μαζί δύο άλλοι ηθοποιοί (c,d) με την προϋπόθεση ότι ο c έχει παίξει σε ταινία μαζί με τον a και ο d έχει παίξει σε ταινία μαζί με τον b. Επιστρέφει τα εξής:
  - Τον τίτλο της κάθε ταινίας των (c,d).
  - Τον κωδικό του ηθοποιού c.
  - Τον κωδικό του ηθοποιού d.
  - Τον κωδικό του ηθοποιού a.
  - Τον κωδικό του ηθοποιού b.
3. **actorPairs:** Η συνάρτηση αυτή δέχεται ως όρισμα τον κωδικό ενός ηθοποιού. Βρίσκει τους ηθοποιούς ο καθένας από τους οποίους έχει παίξει **μόνο** σε διαφορετικά είδη ταινιών από αυτόν αλλά που μαζί μ' αυτόν έχει παίξει σε τουλάχιστον επτά διαφορετικά είδη. Επιστρέφει το εξής:
  - Τους κωδικούς των ηθοποιών που πληρούν την παραπάνω συνθήκη.
4. **selectTopNactors:** Η συνάρτηση αυτή δέχεται ως όρισμα έναν ακέραιο N. Για κάθε είδος ταινίας, βρίσκει τους N ηθοποιούς που έχουν παίξει στις περισσότερες ταινίες του είδους. Επιστρέφει τα εξής:
  - Το όνομα του είδους.
  - Τον κωδικό κάθε ηθοποιού.
  - Τον αριθμό των ταινιών του είδους στις οποίες έχει παίξει ο κάθε ηθοποιός.
5. **traceActorInfluence (bonus 10%):** Η συνάρτηση αυτή δέχεται ως όρισμα τον κωδικό ενός ηθοποιού. Στην συνέχεια, υπολογίζει το μεταβατικό εγκλεισμό της επιρροής του ηθοποιού αυτού σε άλλους ηθοποιούς ως προς το είδος ταινιών και επιστρέφει τους κωδικούς των ηθοποιών που επηρεάζει.  
*Hint: Πώς επηρεάζεται ένας ηθοποιός από έναν άλλον;*

Ένας ηθοποιός a επηρεάζει έναν ηθοποιό b όταν έχουν παίξει σε μια ταινία μαζί (ανεξάρτητα από το πόσοι άλλοι ηθοποιοί παίζουν) και ο b παίζει σε μια άλλη μεταγενέστερης χρονιάς ταινία ίδιου είδους. Η επιρροή αυτή δημιουργεί ένα γράφο όπου οι κόμβοι του είναι ηθοποιοί και οι

κατευθυνόμενες ακμές του αναπαριστούν επιρροές (υπάρχει κατευθυνόμενη ακμή ανάμεσα σε δύο κόμβους a και b αν ο a επηρέασε τον b). Αυτό που θέλουμε είναι να υπολογίσουμε το μεταβατικό εγκλεισμό αυτού του γράφου [1]. Το αποτέλεσμα περιέχει τους ηθοποιούς που επηρεάζει ο δεδομένος ηθοποιός.

Για παράδειγμα, αν ο a έχει επηρεάσει τον b και ο b έχει επηρεάσει τον c, τότε και ο a έχει επηρεάσει τον c (έμμεσα). Δηλαδή, αν  $a \rightarrow b$  και  $b \rightarrow c$ , τότε στο αποτέλεσμα πρέπει να εμφανίζονται τα εξής:

(b,)

(c,) // Λόγω μεταβατικού εγκλεισμού

Σημείωση: Υποθέτουμε ότι το σύνολο των δεδομένων είναι μεγάλο και δεν μπορεί να χωρέσει ολόκληρο στη μνήμη. Δεδομένου αυτού, θα πρέπει να προσπελάσετε το υποσύνολο των δεδομένων που χρειάζεται να εμφανιστούν στο αποτέλεσμα (θα πρέπει να εκτελέσετε παραπάνω από ένα ερώτημα στην βάση).

1. [http://en.wikipedia.org/wiki/Transitive\\_closure](http://en.wikipedia.org/wiki/Transitive_closure)

### Γενικά ζητήματα που πρέπει να προσέξετε

- Είναι σημαντικό να μελετήσετε/τρέξετε τα παραδείγματα του φροντιστηρίου πριν ξεκινήσετε την υλοποίηση της εργασίας.
- Δεν επιτρέπεται καμία εντολή που δεν έχετε διδαχθεί στο μάθημα. Ενδεικτικά παραδείγματα που **δεν επιτρέπονται** είναι τα παρακάτω:
  - Τα SELECT μέσα στο SELECT και το FROM clause δεν επιτρέπονται.
  - Το JOIN δεν επιτρέπεται καθόλου ως εντολή (δηλαδή δεν επιτρέπονται LEFT JOIN, CROSS JOIN, INNER JOIN κτλ ... ).
  - Δεν επιτρέπεται η χρήση της LIMIT.
- Μπορείτε να χρησιμοποιήσετε περισσότερα από ένα SQL ερωτήματα. Τα ερωτήματα μπορείτε φυσικά να τα δοκιμάζετε και στο Workbench όμως στο τελικό παραδοτέο πρέπει να τρέχουν **μόνο** μέσω των αρχείων που σας δίνουμε **σε python**, διαφορετικά δεν βαθμολογούνται.
- Οτιδήποτε άλλο αφορά την python επιτρέπεται (εκτός από την βιβλιοθήκη pandas).
- Στο αρχείο app.py έχετε έναν σκελετό για την εργασία και θα χρειαστεί να τροποποιήσετε τις συναρτήσεις που δίνονται ώστε να υλοποιηθούν τα ζητούμενα. Αντιθέτως τα υπόλοιπα αρχεία στον φάκελο application.zip δεν χρειάζεται να τροποποιηθούν.
- Θα παραδώσετε την εφαρμογή σας, δηλαδή το αρχείο **app.py**, σε ένα αρχείο της μορφής sdixxxx\_app.zip (τα υπόλοιπα αρχεία που περιέχονται στο φάκελο application.zip, δεν χρειάζεται να συμπεριληφθούν στο παραδοτέο της εργασίας). Η εργασία πρέπει να εμφανίζει τα αποτελέσματα στον browser, στη διεύθυνση "<http://localhost:8080>".

Καλή Επιτυχία!