Αναφορά Εξαμηνιαίας Εργασίας Βάσεων Δεδομένων

Κριθαρίδης Κωνσταντίνος

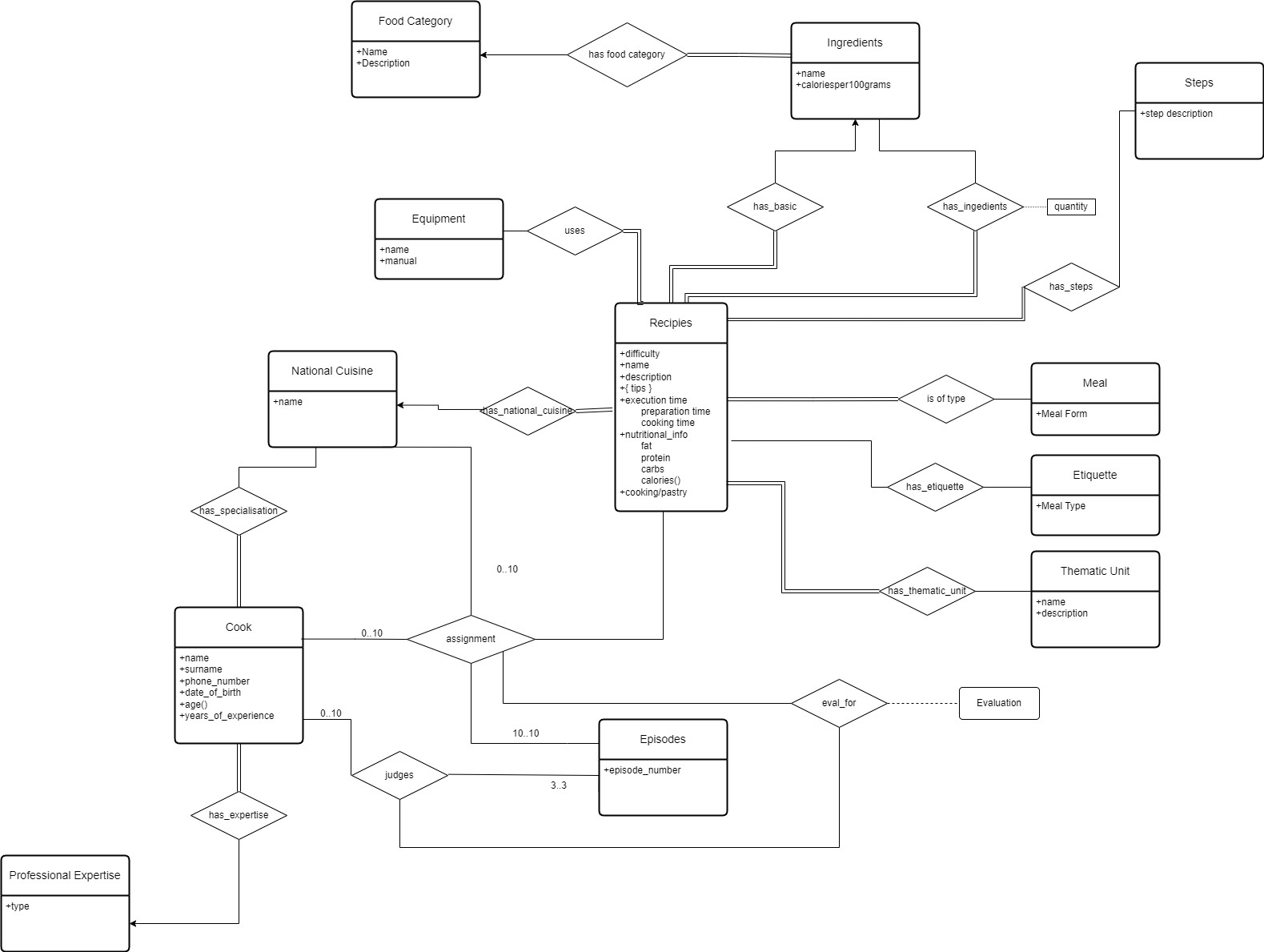
Κάκος Σωτήριος

Μηνάγιας Δημήτριος

Ζητούμενο της εργασίας είναι η σχεδίαση και η υλοποίηση κατάλληλης βάσης δεδομένων, η οποία θα αξιοποιηθεί από δημοφιλή διαγωνισμό μαγειρικής για την αποθήκευση και την διαχείριση σχετικών πληροφοριών.

Χρησιμοποιήσαμε το εργαλείο MySQLWorkbench για την ανάπτυξη του κώδικα SQL της εφαρμογής σε MySQL, όπως επίσης Python για να γεμίσουμε την βάση με τα δεδομένα. Τα δεδομένα παράχθηκαν με χρήση LLMs, με τα οποία διασφαλίσαμε την εγκυρότητά τους και ότι αποτελούν αληθινές συνταγές.

Προκειμένου η βάση δεδομένων να ανταποκρίνεται στις ζητούμενες προδιαγραφές, αρχικά αναπτύξαμε το διάγραμμα οντοτήτων-συσχετίσεων (E-R diagram), το οποίο επισυνάπτεται παρακάτω:



Ο διαγωνισμός διαδραματίζεται σε επεισόδια σε καθένα από τα οποία επιλέγονται με τυχαίο τρόπο 10 εθνικές κουζίνες, 3 κριτές και 10 συνταγές. Στην υλοποίησή μας χρησιμοποιούμε δύο συσχετίσεις (relations) εκ των οποίων η μία συνδέει τους μάγειρες με τα επεισόδια για την αποθήκευση των κριτών ανά επεισόδιο και η δεύτερη συσχετίζει τα επεισόδια με τους μάγειρες και τις εθνικές συνταγές. Ένας περιορισμός του E-R μοντέλου είναι ότι δεν μπορεί να εκφράσει συσχετίσεις μεταξύ συσχετίσεων. Μία τέτοια δομή είναι αναγκαίο στην περίπτωσή μας προκειμένου να διατηρούμε την πληροφορία της βαθμολογίας του κάθε κριτή σε κάθε μάγειρα ανά επεισόδιο. Ο καλύτερος τρόπος μοντελοποίησης αυτής της κατάστασης είναι η χρήση συνάθροισης (συσχέτιση eval\_for).

Αφού ολοκληρώσαμε τον εννοιολογικό σχεδιασμό, προχωρήσαμε στον λογικό σχεδιασμό, το αποτέλεσμα του οποίου είναι το παρακάτω σχεσιακό διάγραμμα (relational schema).

A computer diagram of a computer program

Description automatically generated with medium confidence

Στο σχεσιακό διάγραμμα συναντάμε κάποια tables τα οποία διαδραματίζουν καθοριστικό ρόλο στην υλοποίηση της βάσης. Καταρχάς, το table Cooks έχει αποθηκευμένες όλες τις σχετικές πληροφορίες για τους μάγειρες του διαγωνισμού, είτε κριτές είτε διαγωνιζόμενοι. Επίσης, το table Recipes έχει αποθηκευμένες πληροφορίες σχετικά με την κάθε συνταγή, οι οποίες θα ανατίθενται, σύμφωνα με τις προδιαγραφές, στους μάγειρες, μετά από κλήρωση. Η προαναφερθείσα κλήρωση προκύπτει στο table Assignment, στο οποίο καταχωρείται η κάθε ανάθεση συνταγής σε κάποιον μάγειρα. Οι μάγειρες αξιολογούνται για κάθε assignment από 3 κριτές (η κλήρωση των κριτών αποθηκεύεται στον πίνακα Judges), και το κάθε tuple που αποτελείται από τα στοιχεία διαγωνιζόμενου, στοιχεία κριτή- στοιχεία συνταγής και τελική βαθμολόγηση καταχωρείται στον πίνακα eval\_for, που αποτελεί και τον ακρογωνιαίο λίθο της βάσης δεδομένων που αναπτύξαμε.

Τα ευρετήρια (indexes) που υλοποιήσαμε, είναι τα εξής:

* Στον πίνακα Recipies, έχουμε το name, δηλαδή το όνομα της συνταγής στην οποία αναφερόμαστε (Primary Key).
* Στον ίδιο πίνακα υπάρχει το attribute basic\_ingredient που είναι foreign key στον πίνακα Ingredients και έχουμε ένα attribute National\_Cuisine\_name που είναι foreign key στον πίνακα National\_Cuisine.
* Ο πίνακας Ingredients έχει primary key το name του.
* Ο πίνακας που περιέχει τα υλικά της κάθε συνταγής (Recipies\_has\_Ingredients) έχει foreign keys στον πίνακα των συνταγών και των υλικών.
* Ο πίνακας Equipment έχει primary key το όνομα name του εξοπλισμού και ο Recipies\_has\_Equipment έχει foreign keys στα πίνακα του εξοπλισμού και των συνταγών.
* Ο πίνακας των συμβουλών Tips έχει foreign key στον πίνακα των συνταγών.
* Ο πίνακας Recipes\_has\_Steps στον οποίον αποθηκεύουμε τα βήματα κάθε συνταγής έχει primary key το ζεύγος Recipies\_name και step\_number και foreign keys στον πίνακα των συνταγών των βημάτων Steps.
* Ο πίνακας των ετικετών etiquette έχει primary key το όνομα της ετικέτας Meal\_Type.
* Ο πίνακας Recipies\_has\_Etiquette έχει primary key το ζεύγος Recipies\_name και Etiquette\_Meal\_Type και foreign key προς τον πίνακα των συνταγών και των ετικετών αντίστοιχα.
* Ο πίνακας των μορφών γευμάτων Meal έχει primary key το Meal\_Form και ο πίνακας Meal\_has\_Recipies έχει primary key το ζεύγος Recipies\_name, Meal\_meal\_form και foreign keys προς τον πίνακα των συνταγών και των γευμάτων αντίστοιχα.
* Ο πίνακας των θεματικών ενοτήτων Thematic\_Unit έχει primary key το όνομα name και ο πίνακας Recipies\_has\_Thematic\_Unit έχει primary key το ζεύγος Recipies\_name, Thematic\_Unit\_name και foreign keys στον πίνακα των συνταγών και των θεματικών ενοτήτων αντίστοιχα.
* Ο πίνακας των εθνικών κουζίνων National\_Cuisine έχει primary key το όνομα name.
* Ο πίνακας της εξειδίκευσης των μαγείρων Has\_specialisation έχει primary key την τριάδα National\_Cuisine\_name, Cook\_name, Cook\_surname και foreign keys προς τον πίνακα των Εθνικών κουζινών και των μαγείρων αντίστοιχα.
* Ο πίνακας των μαγείρων Cook έχει primary key το ζεύγος name, surname και foreign key στον πίνακα της επαγγελματικής κατάρτισης Professional\_Expertise.
* Ο πίνακας της επαγγελματικής κατάρτισης Professional\_Expertise έχει primary key το είδος type.
* Ο πίνακας των επεισοδίων έχει primary key τον αριθμό των επεισοδίων episode\_number.
* Ο πίνακας Assignment έχει primary key το σύνολο Αριθμός Επεισοδίου Episode\_episode\_number, το όνομα εθνικής κουζίνας National\_Cuisine\_name, το όνομα της συνταγής Recipies\_name, το ονοματεπώνυμο του μάγειρα, Cook\_name, Cook\_surname με foreign keys προς τους αντίστοιχους πίνακες.
* Ο κύριος πίνακας της βάσης eval\_for έχει primary key τον συνδυασμό εθνικής κουζίνας, όνομα συνταγής, μάγειρα, κριτή και αριθμό επεισοδίου με foreign keys προς τα αντίστοιχα tables.

Πέρα από αυτά τα ευρετήρια που δημιουργούνται αυτόματα λόγων των primary και foreign κλειδιών, υλοποιήσαμε κάποια ευρετήρια για την βελτιστοποίηση της απόδοσης των ερωτημάτων επί της βάσης. Πιο συγκεκριμένα έχουμε:

* Την ηλικία του μάγειρα age
* Στον αριθμό των γραμμαρίων υδατανθράκων του πίνακα συνταγών

Για την εγκατάσταση της εφαρμογής απαραίτητο είναι το MySQLWorkbench, στο οποίο αρχικά θα τρέξετε το αρχείο ddl\_masterchef\_final.sql, για να δημιουργηθούν τα tables. Μετά έχουμε φροντίσει να δημιουργήσουμε το κατάλληλο dml, που είναι το αρχείο dml\_masterchef\_actual\_final\_hex\_2.sql, το οποίο θα τρέξετε μετά το ddl ώστε να γεμίσουν τα tables με δεδομένα. To dml παράχθηκε χρησιμοποιώντας το εργαλείο Export Data του MySQLWorkbench, αφού είχαμε ήδη γεμίσει τα tables, τόσο με το Python script database\_fill.py, στο οποίο γίνεται χρήση των βιβλιοθηκών pymysql, csv, pandas, όπως και του stored procedure populate\_all, με το οποίο γίνεται η κλήρωση του διαγωνισμού. Το populate\_all φροντίζει να καλύπτει όλους τους περιορισμούς που περιγράφονται στην εκφώνηση, μέσα από μια σειρά μικρότερων stored procedures.

Το link του GitHub Repository της εφαρμογής μας, όπου μπορείτε να βρείτε το DDL και το DML είναι το παρακάτω. Να σημειωθεί ότι δεν επισυνάψαμε τον κώδικα του DDL ή του DML, καθώς πρόκειται για αρχεία των 1000+ γραμμών και των 10000+ γραμμών αντίστοιχα.