

### 3 Το Σχεσιακό Μοντέλο και η Σχεσιακή Άλγεβρα

**Συγγραφέας:** Ευαγγελία Πιτουρά

#### *Σκοπός*

Το κεφάλαιο αυτό καλύπτει τις βασικές έννοιες του σχεσιακού μοντέλου που είναι απαραίτητες για την αναπαράσταση των δεδομένων καθώς και το σύνολο τελεστών της σχεσιακής άλγεβρας που επιτρέπουν την ανάκτηση δεδομένων. Ο σκοπός αυτού του κεφαλαίου είναι διττός. Πρώτος σκοπός είναι η κατανόηση της αναπαράστασης των δεδομένων και των περιορισμών ακεραιότητας μιας βάσης δεδομένων με τη χρήση του σχεσιακού μοντέλου. Δεύτερος σκοπός είναι η κατανόηση των βασικών τελεστών της σχεσιακής άλγεβρας και η δυνατότητα συνδυασμού τους για τη διατύπωση πολύπλοκων ερωτημάτων για την ανάκτηση δεδομένων από μια σχεσιακή βάση δεδομένων.

#### *Προσδοκώμενα Αποτελέσματα*

Μετά τη μελέτη αυτού του κεφαλαίου θα μπορείτε:

- με βάση τη περιγραφή των απαιτήσεων για μια βάση δεδομένων, να σχεδιάσετε ένα σχεσιακό σχήμα για την αναπαράσταση των δεδομένων και των περιορισμών ακεραιότητας αυτής της βάσης δεδομένων,
- δεδομένης μιας περιγραφής του εννοιολογικού σχήματος μιας βάσης δεδομένων με χρήση του μοντέλου ΟΣ, να σχεδιάσετε ένα σχεσιακό σχήμα κατάλληλο για αυτή τη βάση δεδομένων,
- να διατυπώνετε πολύπλοκα ερωτήματα για την ανάκτηση δεδομένων από μια σχεσιακή βάση δεδομένων, συνδυάζοντας τους βασικούς τελεστές της σχεσιακής άλγεβρας.

#### *Έννοιες Κλειδιά*

- Σχεσιακό Μοντέλο
- Σχέση

- Κλειδί
- Ξένο Κλειδί
- Σχεσιακή Άλγεβρα
- Επιλογή
- Προβολή
- Συνένωση
- Πράξεις Συνόλου

### *Εισαγωγικές Παρατηρήσεις*

Στο προηγούμενο κεφάλαιο, είδαμε τα βασικά βήματα του σχεδιασμού μιας βάσης δεδομένων και μελετήσαμε σε βάθος θέματα που αφορούν στον εννοιολογικό σχεδιασμό με χρήση του μοντέλου ΟΣ και των επεκτάσεων του. Στο επόμενο βήμα, που καλείται και λογικός σχεδιασμός, το εννοιολογικό σχήμα μετασχηματίζεται στο μοντέλο δεδομένων που υποστηρίζει το ΣΔΒΔ, στο οποίο θα υλοποιηθεί το σύστημα βάσεων δεδομένων. Η πλειοψηφία των ΣΔΒΔ υποστηρίζει το σχεσιακό μοντέλο, το οποίο θα παρουσιάσουμε αναλυτικά στο παρών κεφάλαιο.

Στην Ενότητα 3.1, θα εισάγουμε τις βασικές έννοιες και τους περιορισμούς ακεραιότητας που υποστηρίζει το σχεσιακό μοντέλο. Στην Ενότητα 3.2, θα περιγράψουμε πως μπορούμε να μετατρέψουμε το σχήμα μιας βάσης δεδομένων από το μοντέλο ΟΣ σε σχεσιακό μοντέλο. Το σχεσιακό μοντέλο υποστηρίζει μια τυπική γλώσσα διατύπωσης ερωτήσεων που καλείται σχεσιακή άλγεβρα. Η σχεσιακή παρουσιάζεται διεξοδικά στην Ενότητα 3.3.

## **3.1 Το Σχεσιακό Μοντέλο**

### **3.1.1 Βασικές Έννοιες**

Το βασικό δομικό στοιχείο του σχεσιακού μοντέλου είναι η *σχέση* (*relation*) ή *πίνακας* (*table*). Η πληροφορία που θέλουμε να αποθηκεύσουμε στη βάση δεδομένων δομείται σε μια συλλογή πινάκων. Το *σχήμα μιας σχέσης* προσδιορίζεται από το όνομα της σχέσης και από μια λίστα από γνωρίσματα. Μπορείτε να θεωρείτε το όνομα της σχέσης ως το όνομα του πίνακα, και τα γνωρίσματα ως τα ονόματα των

στηλών του πίνακα. Κάθε γραμμή του πίνακα καλείται *πλειάδα (tuple)* ή *εγγραφή (record)* του πίνακα (σχέσης) και συνήθως αντιστοιχεί σε μια οντότητα ή συσχέτιση του φυσικού κόσμου. Ένα *στιγμιότυπο μιας σχέσης* ή απλά *σχέση* είναι ένα σύνολο από πλειάδες, δηλαδή είναι το σύνολο των γραμμών του πίνακα.

Ας δούμε ένα παράδειγμα. Στην Εικόνα 3.1, βλέπετε τρία σχήματα σχέσεων για τη βάση δεδομένων με κινηματογραφικές ταινίες που είδαμε στο Κεφάλαιο 2 (προς το παρόν, αγνοείστε τα βέλη) και στην Εικόνα 3.2 ένα στιγμιότυπο σχέσης για καθένα από αυτά. Για ευκολία, θεωρήσαμε ότι έχουμε μόνο απλά και μονότιμα γνωρίσματα. Θα δούμε αργότερα πως μπορούμε να χειριστούμε άλλα είδη γνωρισμάτων. Για παράδειγμα, όπως βλέπετε στην Εικόνα 3.1, το *σχήμα* της σχέσης για τους ηθοποιούς είναι ACTOR(Name, Day, Month, Year, Sex Nationality) και το *στιγμιότυπο* της σχέσης ACTOR στην Εικόνα 3.2 περιέχει πέντε πλειάδες που η καθεμία από αυτές αντιστοιχεί σε μια οντότητα ηθοποιού. Για παράδειγμα, μια πλειάδα της σχέσης ACTOR είναι η (George Clooney, 6, May, 1961, Male, American).

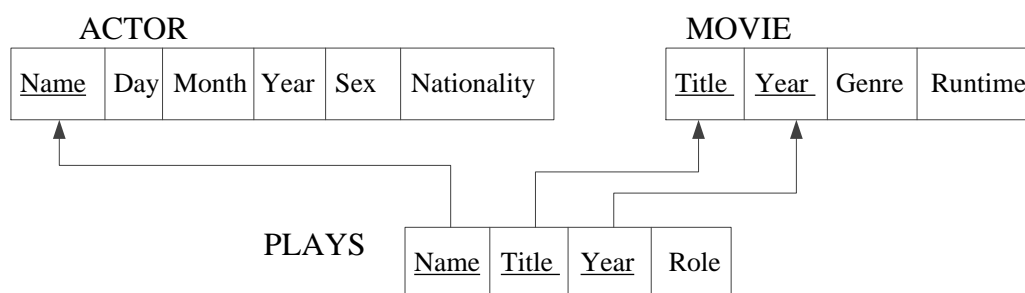
Ας ορίσουμε τώρα πιο τυπικά τις παραπάνω έννοιες. Έστω ένα σχήμα σχέσης  $R(A_1, A_2, \dots, A_n)$  όπου  $X = \{A_1, A_2, \dots, A_n\}$  είναι το σύνολο γνωρισμάτων της  $R$ . Κάθε γνώρισμα  $A_i \in X$  σχετίζεται με ένα σύνολο που ονομάζεται πεδίο ορισμού του  $A_i$ , το οποίο θα συμβολίζουμε με  $dom(A_i)$ . Μια πλειάδα  $t$  της  $R$  είναι μια διατεταγμένη λίστα από  $n$  τιμές  $\langle v_1, v_2, \dots, v_n \rangle$ , όπου κάθε τιμή  $v_i$  είναι ένα στοιχείο του  $dom(A_i)$  ή η ειδική τιμή null. Η τιμή null έχει την ίδια σημασία όπως στο μοντέλο ΟΣ. Το γεγονός ότι κάθε γνώρισμα παίρνει τιμές από ένα πεδίο ορισμού καλείται *περιορισμός πεδίου ορισμού*. Αντίθετα από το μοντέλο ΟΣ που επιτρέπει σύνθετα και πλειότιμα γνωρίσματα, στο σχεσιακό μοντέλο κάθε γνώρισμα παίρνει *ατομικές* τιμές, δηλαδή μια απλή τιμή. Ένα στιγμιότυπο  $r(R)$  της  $R$  είναι σύνολο από πλειάδες, συγκεκριμένα ισχύει:  $r(R) \subseteq dom(A_1) \times dom(A_2) \times \dots \times dom(A_n)$ .

Σημειώνουμε ότι ένα στιγμιότυπο μιας σχέσης είναι ένα σύνολο από πλειάδες. Η χρήση του συνόλου ως μοντέλου αναπαράστασης σημαίνει ότι σε οποιοδήποτε στιγμιότυπο δεν μπορούμε να έχουμε *διπλότυπα (duplicates)*, δηλαδή, δύο πλειάδες που να έχουν τις ίδιες τιμές σε όλα τα γνωρίσματά τους. Επίσης, η σειρά εμφάνισης των πλειάδων σε ένα στιγμιότυπο δεν έχει σημασία. Για παράδειγμα, στην Εικόνα 3.2, η σειρά των γραμμών στους πίνακες είναι τυχαία. Αντίθετα, θεωρούμε ότι έχει σημασία η διάταξη των τιμών των γνωρισμάτων σε μια πλειάδα και κάθε τιμή αντιστοιχεί στη θέση του αντίστοιχου γνωρίσματος στον ορισμό του σχήματος της

σχέσης. Για παράδειγμα, στην Εικόνα 3.2, η σειρά των στηλών και αντίστοιχα οι τιμές τους στους πίνακες ακολουθεί τη σειρά των γνωρισμάτων στα αντίστοιχα σχήματα στην Εικόνα 3.1.

Ονομάζουμε *βαθμό* (*degree* ή *arity*) της σχέσης το πλήθος των γνωρισμάτων της και *πληθικότητα* (*cardinality*) της σχέσης το πλήθος των πλειάδων. Για παράδειγμα, η σχέση ACTOR της Εικόνας 3.1 έχει βαθμό 6 και η πληθικότητα του στιγμιότυπού της στην Εικόνα 3.2 είναι 5.

Επίσης, στα επόμενα θα χρησιμοποιούμε για μια πλειάδα  $t \in R(X)$  τον συμβολισμό  $t[X']$  ή  $t.X'$  για να δηλώσουμε τις τιμές της  $t$  στα γνωρίσματα  $X' \subseteq X$ . Για παράδειγμα, αν  $t$  είναι η τρίτη πλειάδα του πίνακα ACTOR στην Εικόνα 3.2, το  $t[\text{Name}]$  είναι <Sandra Bullock>, ενώ το  $t[\text{Sex}, \text{Nationality}]$  είναι <Female, American>.



**Εικόνα 3.1** Παράδειγμα σχήματος μιας σχεσιακής βάσης δεδομένων για κινηματογραφικές ταινίες

Τέλος, το *σχήμα μιας σχεσιακής βάσης δεδομένων* είναι μια συλλογή από σχήματα σχέσεων. Ένα *στιγμιότυπο μιας σχεσιακής βάσης δεδομένων* είναι μια συλλογή από στιγμιότυπα, ένα για κάθε ένα από τα σχήματα σχέσεων που αποτελούν το σχήμα της. Το σχήμα της σχέσης αναφέρεται στην πρόθεση της βάσης δεδομένων, ενώ το στιγμιότυπο της σχέσης, το οποίο μερικές φορές καλείται και *κατάσταση* (*state*) αναφέρεται στην έκταση της σχέσης. Κατά τη διάρκεια του σχεδιασμού καθορίζουμε το σχήμα το οποίο τροποποιείται σπάνια. Αντίθετα, το στιγμιότυπο μεταβάλλεται, καθώς πλειάδες εισάγονται στις σχέσεις, τροποποιούνται ή διαγράφονται.

### 3.1.2 Περιορισμοί Ακεραιότητας στο Σχεσιακό Μοντέλο

Όπως και στο μοντέλο ΟΣ, ο ορισμός ενός σχεσιακού σχήματος πέρα από τον ορισμό των δομικών στοιχείων της βάσης δεδομένων περιλαμβάνει και τον ορισμό

περιορισμών ακεραιότητας. Ένας βασικός περιορισμός ακεραιότητας είναι ο περιορισμός του κλειδιού. *Υπερκλειδί* ενός σχεσιακού σχήματος  $R(X)$  είναι ένα υποσύνολο  $K \subseteq X$  των γνωρισμάτων του, τέτοιο ώστε σε κάθε πιθανό στιγμιότυπο της σχέσης, κανένα ζευγάρι πλειάδων δε μπορεί να έχει τον ίδιο συνδυασμό τιμών για τα γνωρίσματα αυτά, δηλαδή, δεν μπορεί να υπάρχουν δύο διαφορετικές πλειάδες  $t_1$  και  $t_2$  σε οποιοδήποτε στιγμιότυπο της  $R$  για τις οποίες να ισχύει  $t_1[K] = t_2[K]$ . *Υποψήφιο κλειδί* είναι το μικρότερο υπερκλειδί, δηλαδή, ένα υπερκλειδί που αν του αφαιρέσουμε οποιοδήποτε από τα γνωρίσματα του παύει να είναι υπερκλειδί. Όταν υπάρχουν περισσότερα από ένα υποψήφια κλειδιά επιλέγουμε ένα από αυτά ως το *πρωτεύον κλειδί* για τη σχέση. Δείτε και τους ανάλογους ορισμούς για το μοντέλο ΟΣ στην Ενότητα 2.2.3.

#### ACTOR

Name	Day	Month	Year	Sex	Nationality
George Clooney	6	May	1961	Male	American
Emmanuelle Riva	24	Feb	1927	Female	French
Sandra Bullock	26	Jul	1964	Female	American
Brad Pitt	18	Dec	1963	Male	American
Frank Sinatra	12	Dec	1915	Male	American

#### MOVIE

Title	Year	Genre	Runtime
Amour	2012	drama	117
Gravity	2013	science-fiction	91
Ocean's Eleven	1960	crime	127
Frozen	2013	comedy	102
Ocean's Eleven	2001	crime	116

#### PLAYS

Name	Title	Year	Role
George Clooney	Gravity	2013	Mark Kowalski
George Clooney	Ocean's Eleven	2001	Danny Ocean
Sandra Bullock	Gravity	2013	Ryan Stone
Emmanuelle Riva	Amour	2012	Anne
Frank Sinatra	Ocean's Eleven	1960	Danny Ocean
Brad Pitt	Ocean's Eleven	2001	Rusty Ryan

**Εικόνα 3.2** Ένα στιγμιότυπο του σχήματος της σχεσιακής βάσης δεδομένων για τις κινηματογραφικές ταινίες

Στο σχήμα κάθε σχέσης, τα αντίστοιχα πρωτεύοντα κλειδιά εμφανίζονται υπογραμμισμένα. Για παράδειγμα, για την PLAYS στην Εικόνα 3.1, πρωτεύον κλειδί είναι το σύνολο γνωρισμάτων  $K = \{\text{Name, Title, Year}\}$ . Αυτό σημαίνει ότι δεν μπορεί να υπάρχουν δύο πλειάδες σε οποιοδήποτε στιγμιότυπο της PLAYS με τις ίδιες τιμές και στα τρία αυτά γνωρίσματα. Το σύνολο  $K$  είναι ελάχιστο με την έννοια ότι αν του αφαιρέσουμε οποιοδήποτε γνώρισμα παύει να είναι υπερκλειδί. Για παράδειγμα, αν αφαιρέσουμε το γνώρισμα Title από το  $K$ , το σύνολο  $\{\text{Name, Year}\}$  που προκύπτει δεν είναι υπερκλειδί. Αυτό ισχύει γιατί αν το  $\{\text{Name, Year}\}$  ήταν υπερκλειδί, αυτό θα σήμαινε ότι δεν μπορεί να υπάρχουν πλειάδες με τις ίδιες τιμές στα δύο αυτά

γνωρίσματα που αντιστοιχούν στο όνομα ηθοποιού και τη χρονιά, δηλαδή ένας ηθοποιός δε θα μπορούσε να παίζει σε δυο ταινίες την ίδια χρονιά.

Τα γνωρίσματα των υποψήφιων κλειδιών δε μπορεί να πάρουν την τιμή null. Αυτός ο περιορισμός ονομάζεται και *περιορισμός ακεραιότητας οντοτήτων (entity integrity constraint)*.

Μέχρι τώρα, είδαμε περιορισμούς που αφορούν στο σχήμα μίας μόνο σχέσης. Συχνά πληροφορία που καταχωρείται σε μια σχέση συνδέεται με πληροφορία που καταχωρείται σε κάποια άλλη σχέση. Ένας τέτοιος περιορισμός μεταξύ δύο σχημάτων σχέσεων  $R_1(X)$  και  $R_2(Y)$  είναι ο περιορισμός *αναφορικής ακεραιότητας (referential integrity constraint)* ή *ξένου κλειδιού (foreign key constraint)* που με απλά λόγια μας λέει ότι οι πλειάδες της  $R_1$  πρέπει να αφορούν πλειάδες που υπάρχουν στην  $R_2$ . Τυπικά, έστω δύο σχήματα σχέσεων  $R_1(X)$  και  $R_2(Y)$ , ένα σύνολο γνωρισμάτων  $F \subseteq X$  της  $R_1$  είναι *ξένο κλειδί* της  $R_1$  που *αναφέρεται* στην  $R_2$  αν (1) το σύνολο  $F$  αποτελείται από το ίδιο πλήθος και με το ίδιο πεδίο ορισμού γνωρισμάτων όπως και το πρωτεύον κλειδί  $K \subseteq Y$  της  $R_2$  και (2) σε οποιοδήποτε στιγμιότυπο της βάσης δεδομένων, για μια πλειάδα  $t_1$  της  $R_1$  ισχύει ότι είτε (α) όλα τα γνωρίσματα της  $t_1$  έχουν την τιμή null είτε (β) στο ίδιο στιγμιότυπο, υπάρχει μια πλειάδα  $t_2$  της  $R_2$ , τέτοια ώστε  $t_1[F] = t_2[K]$ . Στη δεύτερη περίπτωση λέμε ότι η πλειάδα  $t_1$  της  $R_1$  αναφέρεται στην πλειάδα  $t_2$  της  $R_2$ . Η  $R_2$  καλείται *αναφερόμενη* σχέση και η  $R_1$  *αναφέρουσα* σχέση.

Διαγραμματικά, το ξένο κλειδί αναπαρίσταται με ένα κατευθυνόμενο βέλος από τα ξένα κλειδιά προς τα γνωρίσματα στα οποία αυτά αναφέρονται. Στην Εικόνα 3.1, η PLAYS έχει δύο ξένα κλειδιά: το ξένο κλειδί {Name} που αναφέρεται στην ACTOR και το ξένο κλειδί {Title, Year} που αναφέρεται στη MOVIE. Με απλά λόγια, για παράδειγμα, το ξένο κλειδί στην ACTOR μας λέει ότι σε κάθε στιγμιότυπο της βάσης δεδομένων (όπως αυτό στην Εικόνα 3.2), στην PLAYS πρέπει να εμφανίζονται μόνο ηθοποιοί που υπάρχουν στην ACTOR. Αντίστοιχα, το ξένο κλειδί στη MOVIE μας λέει ότι σε κάθε στιγμιότυπο, στην PLAYS πρέπει να εμφανίζονται μόνο ταινίες που υπάρχουν στην MOVIE. Στην Εικόνα 3.2, για παράδειγμα η τέταρτη πλειάδα αναφέρεται στην δεύτερη πλειάδα της ACTOR μέσω του ξένου κλειδιού στην ACTOR και στην πρώτη πλειάδα της MOVIE μέσω του ξένου κλειδιού στη MOVIE.

Ξένα κλειδιά εμφανίζονται συνήθως σε σχέσεις που μοντελοποιούν συσχετίσεις και αναφέρονται στα πρωτεύοντα κλειδιά των συσχετιζόμενων οντοτήτων. Ένα ξένο

κλειδί μιας σχέσης  $R$  μπορεί να αναφέρεται στο πρωτεύον κλειδί της ίδιας της σχέσης  $R$ . Τέτοιου είδους κλειδιά προκύπτουν συνήθως από αναδρομικές συσχετίσεις.

Όπως έχουμε ήδη αναφέρει, οι περιορισμοί ακεραιότητας, όπως αυτοί του κλειδιού και του ξένου κλειδιού, ορίζονται στο σχήμα της βάσης δεδομένων και πρέπει να ικανοποιούνται από οποιοδήποτε στιγμιότυπο της βάσης για να είναι αυτό το στιγμιότυπο αποδεκτό. Οι περιορισμοί προκύπτουν από την περιγραφή των απαιτήσεων του μικρόκοσμου που μοντελοποιεί η βάση δεδομένων.

Πέρα από τους παραπάνω περιορισμούς, συχνά σε ένα σχήμα ισχύουν και γενικού τύπου ή σημασιολογικοί περιορισμοί. Για παράδειγμα, σε μια βάση δεδομένων για κάποιο πανεπιστήμιο, ένας γενικός περιορισμός θα μπορούσε να αφορά στο μέγιστο αριθμό μαθημάτων που μπορεί να παρακολουθεί ένας φοιτητής. Τέτοιοι περιορισμοί δε μπορούν να εκφραστούν στο σχεσιακό μοντέλο. Στο Κεφάλαιο 4 θα δούμε πως υποστηρίζονται από σχεσιακά ΣΔΒΔ.

### Άσκηση Αυτοαξιολόγησης 3.1: Περιορισμός Κλειδιού

(α) Έστω μια σχέση  $R(A, B, C)$  και το στιγμιότυπο της  $r(R)$  του Πίνακα 3.1. Για καθένα από τα παρακάτω, πείτε αν ισχύουν, δεν ισχύουν ή δε μπορείτε να πείτε αν ισχύουν:

- (i) Το  $\{A\}$  είναι υποψήφιο κλειδί,
- (ii) Το  $\{C, D\}$  είναι υποψήφιο κλειδί,
- (iii) Το  $\{A, B\}$  είναι υποψήφιο κλειδί,
- (iv) Αν το  $\{A, B\}$  είναι υποψήφιο κλειδί, τότε το  $\{A, B, D\}$  δεν είναι υποψήφιο κλειδί.

(β) Θεωρείστε το σχήμα σχέσης  $R(A, B, C)$ . Οι πληθάρημοι (cardinality) του πεδίου ορισμού των γνωρισμάτων  $A$ ,  $B$  και  $C$  είναι αντίστοιχα 500, 100 και 300. Ποιος είναι ο μέγιστος αριθμός πλειάδων που μπορεί να έχει ένα οποιοδήποτε αποδεκτό στιγμιότυπο της σχέσης  $R$  αν:

- (i) το μοναδικό υποψήφιο κλειδί της  $R$  είναι το  $\{A\}$ ,
- (ii) τα μοναδικά υποψήφια κλειδιά της  $R$  είναι τα  $\{A, B\}$  και  $\{B, C\}$ .

**Πίνακας 3.1** Στιγμιότυπο σχέσης για την Άσκηση Αυτοαξιολόγησης 3.1.

A	B	C	D
1	2	3	1
5	3	5	2
5	4	3	1
1	3	3	2

(γ) Έστω το σχήμα σχέσης  $R(A_1, A_2, \dots, A_n)$  βαθμού  $n$ . Πόσα είναι τα πιθανά υπερκλειδιά αν

- (i) το μοναδικό υποψήφιο κλειδί της  $R$  είναι το  $\{A_1\}$  και
- (ii) τα μοναδικά υποψήφια κλειδιά της  $R$  είναι τα  $\{A_1\}$  και  $\{A_2\}$ .

**Άσκηση Αυτοαξιολόγησης 3.2:** Περιορισμός Ξένου Κλειδιού

(α) Δώστε μια πλειάδα της οποίας η εισαγωγή στη σχέση PLAYS της Εικόνας 3.2 θα προκαλούσε παραβίαση του ξένου κλειδιού που αναφέρεται στη σχέση ACTOR και μια πλειάδα της οποίας η εισαγωγή στη σχέση PLAYS θα προκαλούσε παραβίαση του ξένου κλειδιού που αναφέρεται στη σχέση MOVIE.

(β) Επαναλάβετε το ερώτημα (α) δίνοντας τώρα όμως από ένα παράδειγμα διαγραφής μιας πλειάδας από κάποια από τις σχέσεις της Εικόνας 3.2.

### 3.2 Μετατροπή Μοντέλου Οντοτήτων-Συσχετίσεων σε Σχεσιακό Μοντέλο

Στην ενότητα αυτή, θα θεωρήσουμε ότι κατά τη διάρκεια της σχεδίασης μιας βάσης δεδομένων μας δίνεται το εννοιολογικό σχήμα της βάσης σε διάγραμμα ΟΣ και θέλουμε να σχεδιάσουμε ένα σχεσιακό σχήμα που να το προσεγγίζει στο μεγαλύτερο δυνατό βαθμό.

#### 3.2.1 Τύποι Οντοτήτων και Συσχετίσεων με Απλά Γνωρίσματα

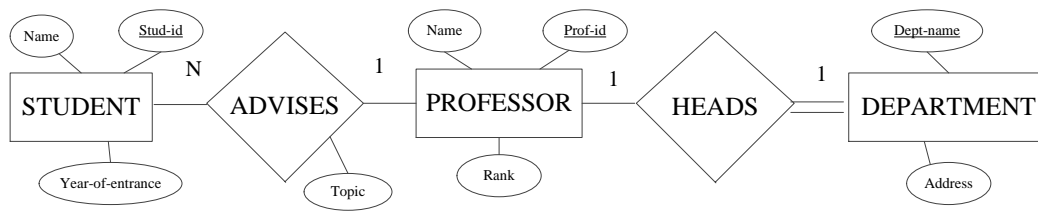
Αρχικά θα θεωρήσουμε ότι έχουμε μόνο ισχυρούς τύπους οντοτήτων και συσχετίσεις με απλά, μονότιμα γνωρίσματα. Για κάθε τύπο οντοτήτων δημιουργούμε έναν αντίστοιχο σχήμα σχέσης με τα ίδια γνωρίσματα και με το ίδιο πρωτεύον κλειδί.



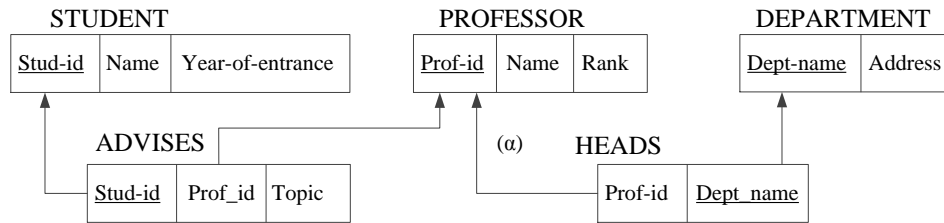
Έστω τώρα μια συσχέτιση  $R$  μεταξύ  $n$  τύπων οντοτήτων που τις έχουμε αντιστοιχήσει στα σχήματα σχέσεων  $S_1, S_2, \dots$  και  $S_n$ . Για τη συσχέτιση  $R$  δημιουργούμε ένα σχήμα σχέσης με γνωρίσματα: (1) τα γνωρίσματα του πρωτεύοντος κλειδιού κάθε συμμετέχουσας σχέσης  $S_i$ , τα οποία ορίζονται ως ξένα κλειδιά προς την  $S_i$  και (2) τα γνωρίσματα της αρχικής συσχέτισης  $R$ , αν υπάρχουν. Το κλειδί για τη σχέση  $R$  που δημιουργήσαμε καθορίζεται από την πληθικότητα της συσχέτισης.

Για δυαδικές συσχετίσεις, αν η συσχέτιση είναι N-M, τότε το πρωτεύον κλειδί της  $R$  πρέπει να περιέχει τα γνωρίσματα των πρωτευόντων κλειδιών και των δύο σχετιζόμενων σχέσεων. Ας πάρουμε ως παράδειγμα το διάγραμμα ΟΣ για τη βάση δεδομένων με τις ταινίες της Εικόνας 2.2. Το αντίστοιχο σχεσιακό μοντέλο φαίνεται στην Εικόνα 3.1. Οι τύποι οντοτήτων ACTOR και MOVIE και η συσχέτιση PLAYS έχουν μοντελοποιηθεί με τα σχήματα σχέσεις με τα ίδια ονόματα. Το σχήμα της PLAYS περιλαμβάνει ως γνωρίσματα (1) το γνώρισμα Name του πρωτεύοντος κλειδιού της ACTORS, τα γνωρίσματα Title και Year του πρωτεύοντος κλειδιού της MOVIE, τα οποία έχουν προσδιοριστεί και ως ξένα κλειδιά που αναφέρονται σε αυτές και (2) το γνώρισμα Role της αντίστοιχης συσχέτισης. Πρωτεύον κλειδί έχει οριστεί το {Name, Title, Year}.

Ας δούμε τώρα πως ορίζουμε το πρωτεύον κλειδί στην περίπτωση που η συσχέτιση που απεικονίζεται σε μια σχέση είναι 1-N ή 1-1. Θα χρησιμοποιήσουμε ως παράδειγμα το διάγραμμα ΟΣ της Εικόνας 3.3(α) που περιγράφει μια βάση δεδομένων για πανεπιστήμια και καταγράφει πληροφορία για καθηγητές που επιβλέπουν φοιτητές και είναι πρόεδροι τμημάτων. Οι φοιτητές, οι καθηγητές και τα τμήματα έχουν μοντελοποιηθεί με τους τύπους οντοτήτων STUDENT, PROFESSOR και DEPARTMENT αντίστοιχα. Το γεγονός ότι ένας καθηγητής επιτηρεί φοιτητές έχει μοντελοποιηθεί με τη συσχέτιση ADVISES που έχει προσδιοριστεί ως 1-N που σημαίνει ότι ένας καθηγητής μπορεί να έχει υπό την επίβλεψή του πολλούς φοιτητές, αλλά ένας φοιτητής έχει το πολύ έναν επιβλέποντα καθηγητή. Το γεγονός ότι ένας καθηγητής μπορεί να είναι πρόεδρος ενός τμήματος έχει μοντελοποιηθεί με τη συσχέτιση HEADS που έχει προσδιοριστεί ως 1-1 που σημαίνει ότι ένας καθηγητής μπορεί να είναι πρόεδρος του πολύ ενός τμήματος και ότι ένα τμήμα έχει το πολύ έναν πρόεδρο.



(α)



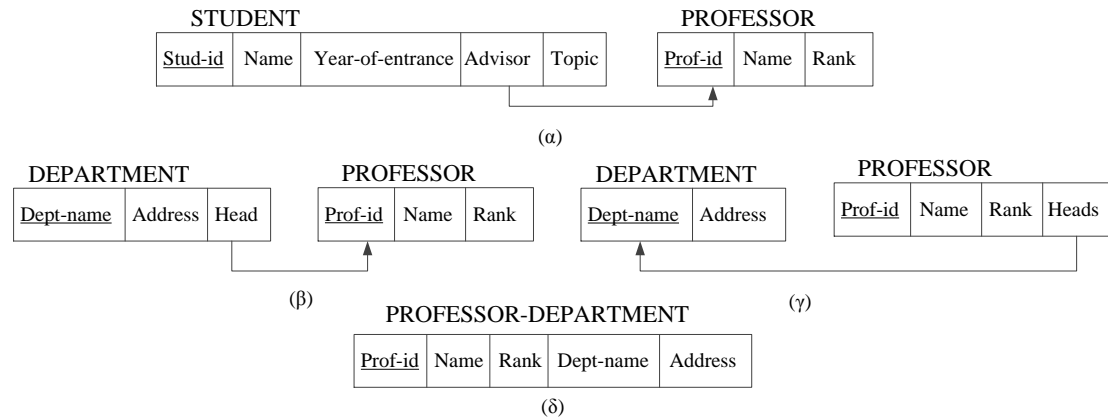
(β)

**Εικόνα 3.3** Παράδειγμα μετατροπής διαγράμματος ΟΣ σε σχεσιακό μοντέλο: (α) σχήμα μιας βάσης δεδομένων σε διάγραμμα ΟΣ, (β) το αντίστοιχο σχεσιακό σχήμα της βάσης δεδομένων

Στην Εικόνα 3.3(β) βλέπετε ένα σχεσιακό σχήμα για αυτή τη βάση δεδομένων. Όπως ήδη συζητήσαμε, έχουμε τρεις σχέσεις που αντιστοιχούν στους τρεις τύπους οντοτήτων και δυο σχέσεις που αντιστοιχούν στους δύο τύπους συσχετίσεων. Οι δύο αυτές σχέσεις έχουν ως γνωρίσματα τα πρωτεύοντα κλειδιά των οντοτήτων που συσχετίζουν, τα οποία είναι επίσης και ξένα κλειδιά και στην περίπτωση της ADVISES και το γνώρισμα Topic της αντίστοιχης συσχέτισης. Ποια γνωρίσματα όμως πρέπει να ορίσουμε ως υποψήφια κλειδιά της ADVISES και της HEADS; Το γεγονός ότι η ADVISES είναι 1-N σημαίνει ότι ένας φοιτητής έχει μόνο έναν επιβλέποντα καθηγητή και άρα ένας φοιτητής εμφανίζεται μόνο μια φορά στην ADVISES, γιατί αν εμφανιζόταν και δεύτερη φορά αυτό θα σήμαινε ότι ο φοιτητής συσχετίζεται και με δεύτερο επιβλέποντα καθηγητή. Αυτό σημαίνει ότι το γνώρισμα Stud-id αποτελεί κλειδί για την ADVISES. Η συσχέτιση HEADS είναι 1-1 και με παρόμοιο σκεπτικό κάθε καθηγητής αλλά και κάθε τμήμα εμφανίζεται μόνο μια φορά στη σχέση HEADS. Άρα η σχέση HEADS έχει δύο υποψήφια κλειδιά, το γνώρισμα Dept-name και το γνώρισμα Prof-id. Θα μπορούσαμε να επιλέξουμε οποιοδήποτε από αυτά τα δύο υποψήφια κλειδιά ως πρωτεύον κλειδί της HEADS. Στην Εικόνα 3.3(β), επιλέξαμε το Dept-name.

Γενικεύοντας, με βάση την ανάλυση αυτού του παραδείγματος, βλέπουμε ότι αν η πληθικότητα μιας συσχέτισης είναι 1-N ή 1-1, τότε το πρωτεύον κλειδί από την

αντίθετη πλευρά του 1 αποτελεί υποψήφιο κλειδί της σχέσης που μοντελοποιεί τη συσχέτιση.



**Εικόνα 3.4** Εναλλακτική μοντελοποίηση (α) της ADVISES, (β) και (γ) της HEADS, και (δ) ενοποίηση οντοτήτων της Εικόνας 3.3(β)

Στις περιπτώσεις που η συσχέτιση είναι 1-N ή 1-1, υπάρχουν εναλλακτικά σχεσιακά σχήματα. Γενικά σε αυτήν την περίπτωση δεν είναι απαραίτητο να δημιουργήσουμε ένα ξεχωριστό σχήμα σχέσης για την συσχέτιση. Συγκεκριμένα, αν η συσχέτιση είναι 1-N από έναν τύπο οντοτήτων  $S_1$  σε έναν τύπο οντοτήτων  $S_2$ , μπορούμε να ενσωματώσουμε την πληροφορία της συσχέτισης στο σχήμα της σχέσης που μοντελοποιεί την οντότητα από την άλλη πλευρά του 1, δηλαδή στην  $S_2$ , προσθέτοντας σε αυτήν ως γνωρίσματα (1) τα γνωρίσματα του πρωτεύοντος κλειδιού της  $S_1$  ως ξένα κλειδιά που αναφέρονται στην  $S_1$  και (2) οποιαδήποτε γνωρίσματα της συσχέτισης. Αυτός ο σχεδιασμός για την 1-N συσχέτιση ADVISES που αποφεύγει τη δημιουργία της σχέσης ADVISES δίνεται στην Εικόνα 3.4(α). Με αυτό το σχεδιασμό, έχουμε ένα λιγότερο σχήμα σχέσης. Το αρνητικό είναι ότι μπορεί να έχουμε πολλές τιμές null αν η συμμετοχή της  $S_2$  δεν είναι ολική. Για παράδειγμα, τα γνωρίσματα Advisor και Topic της STUDENT θα έχουν την τιμή null για όλους τους φοιτητές που δε συμμετέχουν στη συσχέτιση ADVISES, δηλαδή δεν έχουν συσχετιστεί με κάποιον καθηγητή και θέμα.

Αντίστοιχα, μπορούμε να ενσωματώσουμε την πληροφορία μιας συσχέτισης 1-1 από έναν τύπο οντοτήτων  $S_1$  σε έναν τύπο οντοτήτων  $S_2$  είτε στο σχήμα σχέσης της  $S_1$  είτε στο σχήμα σχέσης της  $S_2$ . Οι δύο αυτοί εναλλακτικοί σχεδιασμοί για την 1-1 συσχέτιση HEADS που αποφεύγουν τη δημιουργία της σχέσης HEADS δίνονται στην Εικόνα 3.4(β) όπου η πληροφορία ενσωματώνεται στο DEPARTMENT και στην Εικόνα 3.4(γ) όπου η πληροφορία ενσωματώνεται στο PROFESSOR. Από αυτές

τις δύο επιλογές, η ενσωμάτωση στο DEPARTMENT είναι προτιμότερη γιατί αποφεύγει τις τιμές null, επειδή η συμμετοχή του DEPARTMENT είναι ολική και όλα τα τμήματα θα έχουν κάποια μη null τιμή στο γνώρισμα Head.

Τέλος, όταν η συσχέτιση είναι 1-1 από έναν τύπο οντοτήτων  $S_1$  σε έναν τύπο οντοτήτων  $S_2$ , θα μπορούσαμε να ενοποιήσουμε το σχήμα της σχέσης για το  $S_1$ , το σχήμα της σχέσης για το  $S_2$  και το σχήμα της σχέσης για τη συσχέτιση σε ένα ενιαίο σχήμα. Υποψήφια κλειδιά είναι το πρωτεύον κλειδί της  $S_1$  και το πρωτεύον κλειδί της  $S_2$ . Αυτός ο σχεδιασμός που αντικαθιστά τις σχέσεις PROFESSOR, DEPARTMENT και HEADS με μια ενιαία σχέση PROFESSOR-DEPARTMENT φαίνεται στην 3.4(δ). Αξίζει να σημειώσουμε τα παρακάτω. Πρώτον, ο σχεδιασμός αυτός δεν είναι πολύ φυσικός γιατί η σημασιολογία μιας πλειάδας της σχέσης δεν είναι εμφανής, αφού μια πλειάδα δεν αντιστοιχεί ούτε σε ένα τμήμα, ούτε σε έναν καθηγητή. Δεύτερον, παρατηρείστε ότι επιλέξαμε ως πρωτεύον κλειδί το Prof-id. Η επιλογή ως πρωτεύοντος κλειδιού το Dept-name (που είναι επίσης υποψήφιο κλειδί) θα οδηγούσε στο παρακάτω πρόβλημα. Δε θα μπορούσαμε να αποθηκεύσουμε στη βάση δεδομένων καθηγητές που δεν ήταν πρόεδροι κάποιου τμήματος, ενώ τέτοιοι καθηγητές μπορεί να υπάρχουν αφού η συμμετοχή του PROFESSOR στο HEADS είναι μερική. Τρίτον, τα γνωρίσματα Dept-name και Address θα παίρνουν null τιμές για τους καθηγητές που δεν είναι πρόεδροι τμημάτων.

Ως τελευταία παρατήρηση, σημειώνουμε ότι το σχεσιακό μοντέλο δεν παρέχει άμεσο τρόπο για τη διατύπωση περιορισμών συμμετοχής.

### **Δραστηριότητα 3.1:** Μετατροπή αναδρομικής συσχέτισης διαγράμματος ΟΣ σε σχεσιακό σχήμα

(α) Μετατρέψτε το διάγραμμα ΟΣ της Εικόνας 2.5(α) σε σχεσιακό μοντέλο, χρησιμοποιώντας για την μετατροπή της 1-1 αναδρομικής συσχέτισης SEQUEL τις διαφορετικές τεχνικές που συζητήσαμε σε αυτήν την ενότητα. Θεωρείστε το γνώρισμα Genre ως μονότιμο.

(β) Δώστε το περιεχόμενο του στιγμιότυπου της SEQUEL για το αντίστοιχο στιγμιότυπο της Εικόνας 2.5(β). Για ευκολία, αγνοείστε τις πραγματικές τιμές των γνωρισμάτων Genre και Runtime, και θεωρείστε ότι είναι ίσες αντίστοιχα με drama και 100 για όλες τις ταινίες του στιγμιότυπου.

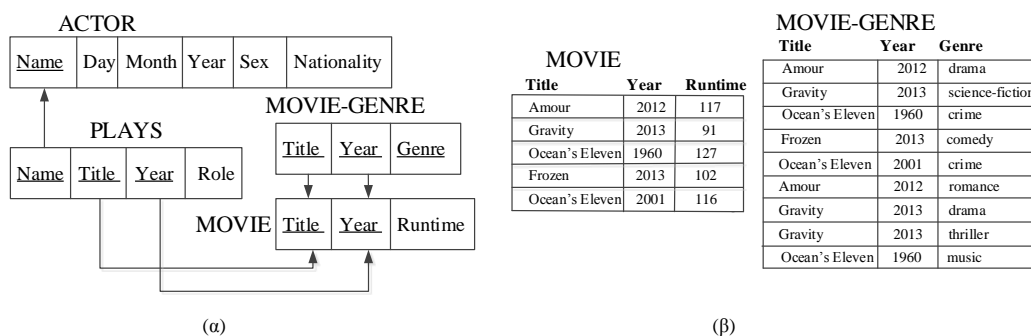
*Εκτιμώμενος χρόνος ολοκλήρωσης: 30 λεπτά*

Στόχος: Η κατανόηση της αντιστοιχίας μεταξύ του μοντέλου ΟΣ και του σχεσιακού μοντέλου

### 3.2.2 Άλλα Είδη Γνωρισμάτων

Ας δούμε τώρα πως μπορούμε να χειριστούμε μη μονότιμα γνωρίσματα του μοντέλου ΟΣ. Για ένα σύνθετο γνώρισμα ενός τύπου οντοτήτων ή συσχετίσεων, δημιουργούμε ένα γνώρισμα στο σχήμα της σχέσης για καθένα από τα απλά γνωρίσματα από τα οποία αυτό απαρτίζεται. Για παράδειγμα, για το σύνθετο γνώρισμα Date-of-birth του τύπου οντοτήτων ACTOR του διαγράμματος ΟΣ της Εικόνας 2.2, δημιουργήσαμε τρία γνωρίσματα, Day, Month και Year στο σχήμα σχέσης ACTOR της Εικόνας 3.1.

Για κάθε πλειότιμο γνώρισμα  $A$  ενός τύπου οντοτήτων ή συσχετίσεων δημιουργούμε ένα νέο σχήμα σχέσης με γνωρίσματα: (1) το γνώρισμα  $A$  και (2) τα γνωρίσματα του πρωτεύοντος κλειδιού της σχέσης που αντιστοιχεί στον τύπο οντοτήτων ή συσχετίσεων του οποίου γνώρισμα είναι το  $A$  ως ξένο κλειδί που αναφέρεται σε αυτήν τη σχέση. Στο πρωτεύον κλειδί της νέας σχέσης συμμετέχουν όλα τα γνωρίσματα της νέας σχέσης.



**Εικόνα 3.5** Μοντελοποίηση πλειότιμου γνωρίσματος (α) το σχεσιακό σχήμα της Εικόνας 3.1 τροποποιημένο ώστε να υποστηρίζει το πλειότιμο γνώρισμα Genre και (β) στιγμιότυπα των σχέσεων MOVIE και MOVIE-GENRE

Ας πάρουμε ως παράδειγμα το πλειότιμο γνώρισμα Genre του τύπου οντοτήτων MOVIES του διαγράμματος ΟΣ της Εικόνας 2.2. Στην Εικόνα 3.1, θεωρήσαμε ότι είναι μονότιμο και το μοντελοποιήσαμε με το γνώρισμα Genre. Αυτό έχει ως αποτέλεσμα, το γνώρισμα αυτό να μπορεί να πάρει μόνο μια τιμή για κάθε ταινία. Στην Εικόνα 3.5(α), μοντελοποιούμε το γνώρισμα Genre με βάση τον τρόπο που

περιγράψαμε, ο οποίος επιτρέπει στο γνώρισμα αυτό να πάρει παραπάνω από μια τιμή. Το σχήμα MOVIE έχει αντικατασταθεί από το νέο σχήμα MOVIE χωρίς το πλειότιμο γνώρισμα και από το σχήμα MOVIE-GENRE που περιέχει το πρωτεύον κλειδί της MOVIE ως ξένο κλειδί και το πλειότιμο γνώρισμα Genre. Παρατηρείστε ότι το πρωτεύον κλειδί της MOVIE-GENRE είναι το {Title, Year, Genre}. Στην Εικόνα 3.5(β), βλέπουμε το στιγμιότυπο της MOVIE και MOVIE-GENRE. Μια ταινία εμφανίζεται τόσο φορές στο στιγμιότυπο της MOVIE-GENRE όσες και οι διαφορετικές τιμές του γνωρίσματος Genre.

Τέλος, για τα γνωρίσματα που έχουν οριστεί στο διάγραμμα ΟΣ ως *παραγόμενα*, έχουμε την επιλογή είτε να τα αποθηκεύσουμε ως γνωρίσματα, αν με βάση τις λειτουργικές απαιτήσεις αυτά χρησιμοποιούνται συχνά οπότε είναι χρήσιμο να αποθηκεύσουμε την τιμή τους, είτε στην αντίθετη περίπτωση, να τα αγνοήσουμε.

### **Δραστηριότητα 3.2:** Μετατροπή διαγράμματος ΟΣ σε σχεσιακό σχήμα

Δώστε ένα κατάλληλο σχεσιακό σχήμα για τη βάση δεδομένων που περιγράφεται από το διάγραμμα ΟΣ της Εικόνας 2.11.

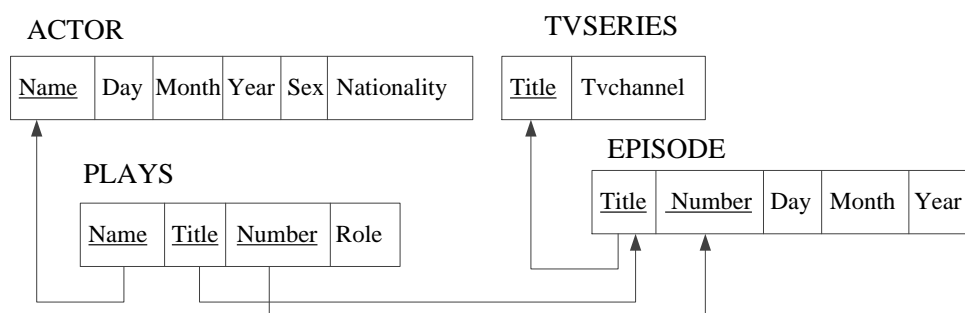
*Εκτιμώμενος χρόνος ολοκλήρωσης:* 30 λεπτά

*Στόχος:* Η κατανόηση της αντιστοιχίας μεταξύ του μοντέλου ΟΣ και του σχεσιακού μοντέλου

### **3.2.3 Ασθενείς Τύποι Οντοτήτων**

Έως τώρα έχουμε εξετάσει μόνο ισχυρούς τύπους οντοτήτων. Στο σχεσιακό μοντέλο, δεν υπάρχει η έννοια της «ασθενούς» σχέσης. Θυμηθείτε ότι δεν επιτρέπονται διπλότυπα σε οποιοδήποτε στιγμιότυπο μιας σχέσης και άρα πάντα το σύνολο όλων των γνωρισμάτων ενός σχήματος σχέσης αποτελεί υπερκλειδί για αυτήν. Για να μοντελοποιήσουμε έναν ασθενή τύπο οντοτήτων  $O_1$  που εξαρτάται από τον ισχυρό τύπο οντοτήτων  $O_2$  μέσω μιας προσδιορίζουσα συσχέτισης  $R$  δημιουργούμε ένα σχήμα σχέσης με γνωρίσματα: τα γνωρίσματα του  $O_1$ , τα γνωρίσματα του πρωτεύοντος κλειδιού του  $O_2$  ως ξένο κλειδί και οποιαδήποτε γνωρίσματα της συσχέτισης αν αυτά υπάρχουν. Το πρωτεύον κλειδί για τη σχέση αποτελείται από το μερικό κλειδί της  $O_1$  και το πρωτεύον κλειδί της  $O_2$ . Προσέξτε ότι η προσδιορίζουσα συσχέτιση έχει ήδη μοντελοποιηθεί μέσω του νέου σχήματος σχέσης και δε χρειάζεται ειδικό σχήμα σχέσης για αυτήν.

Ας δούμε ένα παράδειγμα. Έστω το διάγραμμα ΟΣ της Εικόνας 2.7 για τις τηλεοπτικές σειρές. Οι (ισχυροί) τύποι οντοτήτων ACTOR και EPISODE μοντελοποιούνται από τα ομώνυμα σχεσιακά σχήματα. Ο ασθενής τύπος οντοτήτων EPISODE προσδιορίζεται μέσω της συσχέτισης HAS από τον τύπο οντοτήτων TVSERIES. Στο σχεσιακό μοντέλο, το EPISODE αναπαρίσταται με το σχήμα σχέσης EPISODE που έχει ως γνωρίσματα: τα γνωρίσματα του ασθενούς τύπου και το γνώρισμα Title που είναι το πρωτεύον κλειδί του προσδιορίζοντα τύπου TVSERIES, το οποίο ορίζεται και ως ξένο κλειδί. Πρωτεύον κλειδί είναι το {Title, Number}, δηλαδή το πρωτεύον κλειδί Title της TVSERIES και το μερικό κλειδί Number του EPISODE. Όπως ήδη έχουμε συζητήσει, η N-M συσχέτιση PLAYS μεταξύ ACTOR και EPISODE μοντελοποιείται ως σχήμα σχέσης με γνωρίσματα: τα πρωτεύοντα κλειδιά των δύο τύπων οντοτήτων που συσχετίζει, που είναι {Name} για τον ACTOR και {Title, Number} για το EPISODE ως ξένα κλειδιά και το δικό της γνώρισμα Role.



**Εικόνα 3.6** Το σχεσιακό σχήμα για τη βάση δεδομένων για τις τηλεοπτικές σειρές της οποίας το διάγραμμα ΟΣ δίνεται στην Εικόνα 2.7

### **Δραστηριότητα 3.3:** Μετατροπή διαγράμματος ΟΣ σε σχεσιακό σχήμα

Δώστε ένα κατάλληλο σχεσιακό σχήμα για τη βάση δεδομένων που περιγράφεται από το διάγραμμα ΟΣ της Εικόνας 2.12.

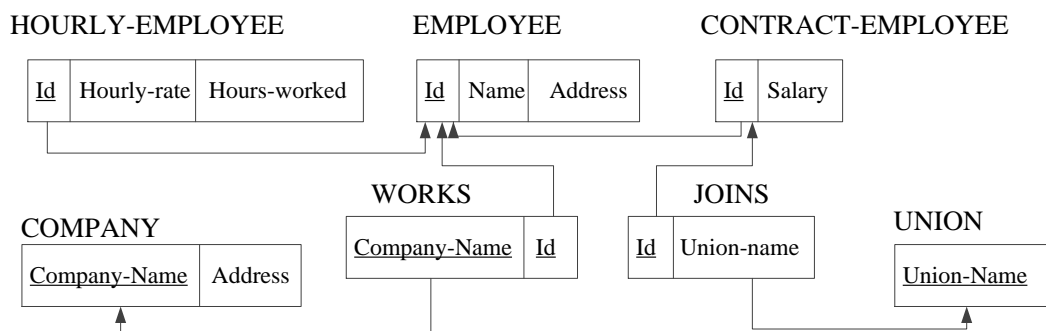
*Εκτιμώμενος χρόνος ολοκλήρωσης:* 30 λεπτά

*Στόχος:* Η κατανόηση της αντιστοιχίας μεταξύ του μοντέλου ΟΣ και του σχεσιακού μοντέλου και ιδιαίτερα των ασθενών οντοτήτων

### 3.2.4 Συσχετίσεις Κλάσης/Υποκλάσης

Το απλό σχεσιακό μοντέλο δεν υποστηρίζει συσχετίσεις κλάσης/υποκλάσης και κληρονομικότητας. Ένας γενικός τρόπος για να απεικονίσουμε μια υπερκλάση και τις υποκλάσεις της είναι να ορίσουμε ξεχωριστούς τύπους σχέσεων για την υπερκλάση και για κάθε μια από τις υποκλάσεις της. Σε αυτήν την περίπτωση, τα σχήματα σχέσεων που μοντελοποιούν υποκλάσεις έχουν ως γνωρίσματα τα τοπικά γνωρίσματα των υποκλάσεων και το πρωτεύον κλειδί της υπερκλάσης τους, το οποίο αποτελεί για αυτές πρωτεύον κλειδί και ξένο κλειδί που αναφέρεται στην υπερκλάση τους.

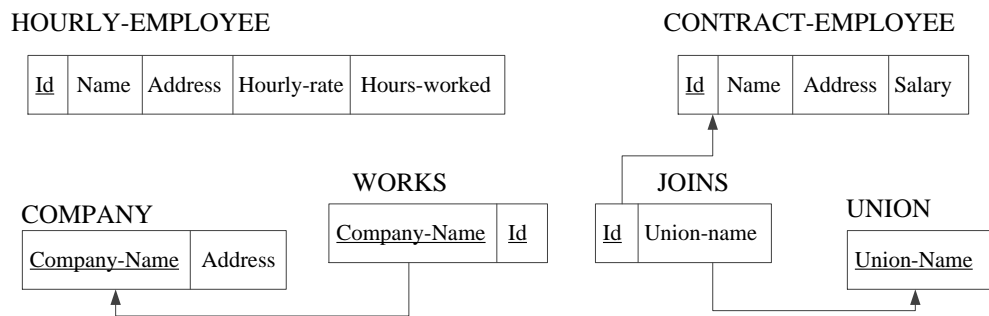
Ένα σχετικό παράδειγμα που ακολουθεί αυτόν τον γενικό τρόπο βλέπετε στην Εικόνα 3.7 για τη βάση δεδομένων με τους εργαζόμενους του διαγράμματος ΕΟΣ της Εικόνας 2.8.



**Εικόνα 3.7** Σχεσιακό μοντέλο που ακολουθεί το γενικό τρόπο σχεδιασμού για τη βάση δεδομένων για τους εργαζόμενους που περιγράφεται από το διάγραμμα ΕΟΣ της Εικόνας 2.8

Εναλλακτικά, μπορούμε να δημιουργήσουμε μόνο σχήματα σχέσεων για τις υποκλάσεις. Σε αυτήν την περίπτωση, κάθε σχήμα σχέσης που αντιστοιχεί σε μια υποκλάση έχει ως γνωρίσματα τα τοπικά γνωρίσματα της υποκλάσης και όλα τα γνωρίσματα της υπερκλάσης. Αυτή η επιλογή είναι δυνατή μόνο αν η εξειδίκευση της υπερκλάσης είναι ολική, γιατί δε δίνει τη δυνατότητα να αποθηκευτούν οντότητες που δεν ανήκουν σε κάποια από τις υποκλάσεις. Επίσης, στην περίπτωση επικάλυψης, αυτός ο σχεδιασμός εισάγει επανάληψη, γιατί για τις οντότητες που ανήκουν σε παραπάνω από μια υποκλάσεις επαναλαμβάνονται οι τιμές των κοινών τους γνωρισμάτων, δηλαδή, αυτών που προέρχονται από την υπερκλάση. Τέλος, αν η υπερκλάση συμμετέχει σε κάποια συσχέτιση, χάνουμε τη δυνατότητα προσδιορισμού ξένου κλειδιού για τον τύπο σχέσης που μοντελοποιεί τη συσχέτιση, όπως θα εξηγήσουμε στο παράδειγμα που ακολουθεί.

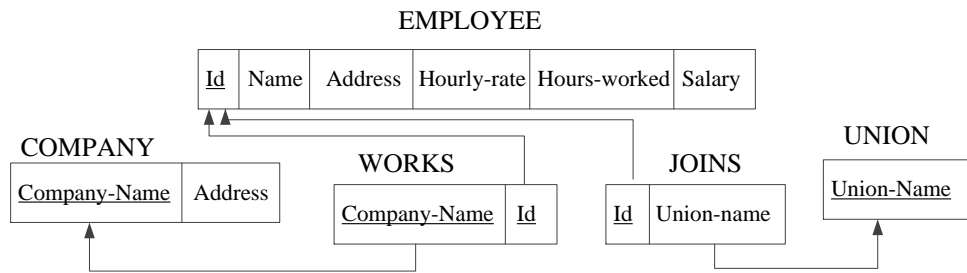




**Εικόνα 3.8** Σχεσιακό μοντέλο χωρίς σχήμα σχέσης για την υπερκλάση για τη βάση δεδομένων για τους εργαζόμενους που περιγράφεται από το διάγραμμα ΕΟΣ της Εικόνας 2.8

Στην Εικόνα 3.8 βλέπετε την εφαρμογή αυτής της προσέγγισης για τη μοντελοποίηση της βάσης δεδομένων για τους εργαζόμενους. Αυτή η επιλογή είναι δυνατή γιατί η εξειδίκευση του EMPLOYEE είναι ολική. Επίσης, δεν υπάρχει επανάληψη των τιμών των κοινών γνωρισμάτων (δηλαδή, των Name, Address), γιατί η εξειδίκευση είναι μη επικαλυπτόμενη και κάθε εργαζόμενος ανήκει μόνο σε μία από τις δύο υποκλάσεις. Ωστόσο, στη συσχέτιση WORKS δε μπορούμε να εκφράσουμε τον περιορισμό ξένου κλειδιού για το γνώρισμα Id, επειδή θα έπρεπε το Id της WORKS να αναφέρεται ως ξένο κλειδί σε δύο σχέσεις (στην HOURLY-EMPLOYEE και στην CONTRACT-EMPLOYEE), που δεν είναι δυνατόν.

Μια άλλη επιλογή θα ήταν να έχουμε ένα ενιαίο σχήμα για την υπερκλάση που να περιλαμβάνει τα δικά της γνωρίσματα και τα τοπικά γνωρίσματα όλων των υποκλάσεων. Η εφαρμογή αυτής της επιλογής για τη βάση δεδομένων με τους εργαζόμενους οδηγεί στο σχεσιακό σχήμα της Εικόνας 3.9. Η προσέγγιση αυτή έχει δυο βασικά μειονεκτήματα. Ένα μειονέκτημα είναι ότι τα τοπικά γνωρίσματα παίρνουν τιμές null για όλες τις οντότητες που δεν ανήκουν στην αντίστοιχη υποκλάση. Για παράδειγμα, για τη βάση δεδομένων με τους εργαζόμενους, τα γνωρίσματα Hourly-rate και Hours-worked θα πάρουν τιμές null για κάθε CONTRACT-EMPLOYEE. Το άλλο μειονέκτημα είναι ότι δεν μπορούμε να επιβάλουμε συμμετοχή σε συσχέτιση μόνο συγκεκριμένων οντοτήτων. Για παράδειγμα, ενώ η συσχέτιση JOINS αφορά μόνο οντότητες τύπου CONTRACT-EMPLOYEE, αυτό δε μπορεί να εκφραστεί στο σχεσιακό σχήμα της Εικόνας 3.9.



**Εικόνα 3. 9** Σχεσιακό μοντέλο με χρήση μόνο μιας σχέσης για τη συσχέτιση κλάσης/υποκλάσης για τη βάση δεδομένων για τους εργαζόμενους που περιγράφεται από το διάγραμμα ΕΟΣ της Εικόνας 2.8

Γενικά, η κατάλληλη απεικόνιση ενός διαγράμματος ΕΟΣ με συσχετίσεις κλάσης/υποκλάσης εξαρτάται από τις απαιτήσεις του προβλήματος.

#### **Δραστηριότητα 3.4:** Μετατροπή διαγράμματος ΟΣ σε σχεσιακό σχήμα

Δώστε ένα κατάλληλο σχεσιακό σχήμα για τη βάση δεδομένων που περιγράφεται από το διάγραμμα ΟΣ της Εικόνας 2.13.

*Εκτιμώμενος χρόνος ολοκλήρωσης:* 30 λεπτά

*Στόχος:* Η κατανόηση της αντιστοιχίας μεταξύ του μοντέλου ΕΟΣ και του σχεσιακού μοντέλου

### **3.3 Σχεσιακή Άλγεβρα**

#### **3.3.1 Εισαγωγή**

Το σχεσιακό μοντέλο, εκτός από τις έννοιες που αναφέρονται στη δομή της βάσης δεδομένων και στους περιορισμούς ακεραιότητας περιλαμβάνει επίσης ένα σύνολο πράξεων ή τελεστών (*operators*) για τη διαχείριση των δεδομένων. Το βασικό σύνολο τελεστών του σχεσιακού μοντέλου αποτελεί τη *σχεσιακή άλγεβρα*. Οι τελεστές της σχεσιακής άλγεβρας εφαρμόζονται πάνω σε σχέσεις και έχουν ως αποτελέσματα νέες σχέσεις. Μια ακολουθία από τελεστές της σχεσιακής άλγεβρας καλείται και *έκφραση της σχεσιακής άλγεβρας* (*relational algebra expression*). Οι εκφράσεις της σχεσιακής άλγεβρας παρέχουν έναν τρόπο για την ανάκτηση πληροφορίας από μια σχεσιακή βάση δεδομένων.

Οι τελεστές της σχεσιακής άλγεβρας μπορεί να χωριστούν σε δύο βασικές κατηγορίες. Η πρώτη κατηγορία περιλαμβάνει *τελεστές συνόλων* από τη μαθηματική θεωρία συνόλων όπως η *ένωση* (*union*), *τομή* (*intersection*), *διαφορά* (*difference*) και το *καρτεσιανό γινόμενο* (*Cartesian product*). Οι τελεστές αυτοί μπορεί να εφαρμοστούν σε σχέσεις γιατί μια σχέση έχει οριστεί ως ένα σύνολο πλειάδων. Η δεύτερη κατηγορία περιλαμβάνει τελεστές που έχουν οριστεί ειδικά για τη διαχείριση σχεσιακών βάσεων δεδομένων με βασικούς τους μοναδιαίους τελεστές της *επιλογής* (*select*) και της *προβολής* (*project*) που εφαρμόζονται σε μία μόνο σχέση καθώς και το *δυναμικό τελεστή της συνένωσης* (*join*) που εφαρμόζεται σε δύο σχέσεις. Στη συνέχεια αυτής της ενότητας, θα μελετήσουμε αυτούς τους τελεστές αναλυτικά.

Η σχεσιακή άλγεβρα είναι ιδιαίτερα σημαντική για πολλούς λόγους. Εκτός της θεωρητικής της σημασίας ως το τυπικό μοντέλο διατύπωσης ερωτημάτων στο σχεσιακό μοντέλο, η σχεσιακή άλγεβρα αποτελεί τη βάση για την υλοποίηση και τη βελτιστοποίηση ερωτήσεων στα σχεσιακά ΣΔΒΔ. Επίσης, πολλές από τις έννοιες της έχουν ενσωματωθεί στην γλώσσα SQL, την τυπική γλώσσα ερωτήσεων σε σχεσιακά ΣΔΒΔ.

### 3.3.2 Οι Τελεστές της Επιλογής και της Προβολής

Οι τελεστές της επιλογής και της προβολής είναι οι δύο βασικοί μοναδιαίοι τελεστές της σχεσιακής άλγεβρας. Ο τελεστής της *επιλογής* χρησιμοποιείται για την επιλογή ενός υποσυνόλου των πλειάδων μιας σχέσης που ικανοποιούν μια συνθήκη, ενώ ο τελεστής της *προβολής* χρησιμοποιείται για την προβολή συγκεκριμένων γνωρισμάτων της σχέσης. Με απλά λόγια, με την επιλογή διαλέγουμε συγκεκριμένες γραμμές ενός πίνακα, ενώ με την προβολή συγκεκριμένες στήλες τους.

**Επιλογή.** Η επιλογή συμβολίζεται με  $\sigma_P(R)$  όπου  $\sigma$  είναι το σύμβολο του τελεστή της επιλογής, ο δείκτης  $P$  είναι μια συνθήκη επιλογής και  $R$  μια σχέση. Αποτέλεσμα της πράξης είναι μια σχέση που έχει το ίδιο σχήμα με τη σχέση  $R$  και περιέχει τις πλειάδες της  $R$  που ικανοποιούν την συνθήκη επιλογής  $P$ . Συγκεκριμένα, η συνθήκη επιλογής είναι μια λογική έκφραση που αποτελείται από λογικές προτάσεις που συνδέονται με τους λογικούς τελεστές AND, OR, NOT και κάθε λογική πρόταση έχει τη μορφή  $A \text{ op } B$  όπου  $A$  είναι ένα γνώρισμα της  $R$ ,  $\text{op}$  είναι ένας από τους τελεστές σύγκρισης  $\{=, \leq, <, >, \geq, \neq\}$  και  $B$  είναι ένα γνώρισμα της  $R$  ή μια τιμή από το πεδίο ορισμού ενός γνωρίσματος της  $R$ . Η συνθήκη της επιλογής εφαρμόζεται

σε κάθε μία από τις πλειάδες της  $R$  ξεχωριστά, και αν η συνθήκη αποτιμάται ως αληθής για αυτήν την πλειάδα, η συγκεκριμένη πλειάδα εμφανίζεται στο αποτέλεσμα.

MOVIE			
Title	Year	Genre	Runtime
Amour	2012	drama	117
Gravity	2013	science-fiction	91
Frozen	2013	comedy	102

(α)

ACTOR					
Name	Day	Month	Year	Sex	Nationality
George Clooney	6	May	1961	Male	American
Brad Pitt	18	Dec	1963	Male	American

(β)

MOVIE			
Title	Year	Genre	Runtime
Amour	2012	drama	117
Ocean's Eleven	1960	crime	127
Ocean's Eleven	2001	crime	116

(γ)

**Εικόνα 3.10** Παραδείγματα αποτελεσμάτων στο στιγμιότυπο της Εικόνας 3.2, των εκφράσεων (α)  $\sigma_{Year > 2010}$  (*MOVIE*), (β)  $\sigma_{Sex = "male" \text{ AND } Year > 1960}$  (*ACTOR*), (γ)  $\sigma_{Genre = "drama" \text{ OR } Genre = "crime"}$  (*MOVIE*)

Ας δούμε μερικά παραδείγματα χρησιμοποιώντας τη βάση δεδομένων για τις κινηματογραφικές ταινίες της Εικόνας 3.1. Έστω η έκφραση  $\sigma_{Year > 2010}$  (*MOVIE*). Η συνθήκη  $Year > 2010$  αποτιμάται για κάθε μια πλειάδα της *MOVIE* και μια πλειάδα ανήκει στο αποτέλεσμα αν για αυτήν, η συνθήκη ισχύει, δηλαδή το γνώρισμα  $Year$  έχει τιμή μεγαλύτερη του 2010. Με ίδιο τρόπο η έκφραση  $\sigma_{Sex = "male" \text{ AND } Year > 1960}$  (*ACTOR*) θα έχει ως αποτέλεσμα τους άντρες ηθοποιούς που γεννήθηκαν μετά το 1960 και η έκφραση  $\sigma_{Genre = "drama" \text{ OR } Genre = "crime"}$  (*MOVIE*) τις ταινίες που το είδος τους έχει την τιμή *drama* ή την τιμή *crime*. Το αποτέλεσμα αυτών των ερωτήσεων στο στιγμιότυπο της Εικόνας 3.2 δίνεται στην Εικόνα 3.10.

Ο βαθμός της σχέσης που προκύπτει ως αποτέλεσμα της επιλογής είναι ίσος με το βαθμό της αρχικής σχέσης, ενώ η *πληθικότητα* της σχέσης που προκύπτει, δηλαδή, το πλήθος πλειάδων της, είναι μικρότερη ή ίση της πληθικότητας της αρχικής σχέσης. Το ποσοστό των πλειάδων που επιλέγονται καλείται και *επιλεκτικότητα* (*selectivity*) της επιλογής.

Το αποτέλεσμα της επιλογής είναι και αυτό μια σχέση στην οποία επίσης μπορούμε να εφαρμόσουμε την πράξη της επιλογής. Για τον τελεστή της επιλογής ισχύει η αντιμεταθετική ιδιότητα, δηλαδή ισχύει  $\sigma_{p1}(\sigma_{p2}(R)) = \sigma_{p2}(\sigma_{p1}(R))$ , δηλαδή δεν έχει σημασία η σειρά εφαρμογής των συνθηκών επιλογής. Ισχύει επίσης ότι  $\sigma_{p1}(\sigma_{p2}(R)) = \sigma_{p1 \text{ AND } p2}(R)$ .

**Προβολή.** Η προβολή συμβολίζεται με  $\pi_L(R)$  όπου  $\pi$  είναι το σύμβολο του τελεστή της προβολής και ο δείκτης  $L$  είναι μια λίστα γνωρισμάτων της σχέσης  $R$ . Αποτέλεσμα της προβολής είναι μια σχέση που έχει ως γνωρίσματα τα γνωρίσματα

που εμφανίζονται στη λίστα  $L$  με τη ίδια διάταξη που αυτά εμφανίζονται στη λίστα και περιέχει τις τιμές των πλειάδων της  $R$  σε αυτά τα γνωρίσματα.

<b>Genre</b>	<b>Name      Sex      Nationality</b>	<b>Genre</b>
drama	George Clooney      Male      American	drama
science-fiction	Emmanuelle Riva      Female      French	science-fiction
crime	Sandra Bullock      Female      American	comedy
comedy	Brad Pitt      Male      American	
	Frank Sinatra      Male      American	
(α)	(β)	(γ)

**Εικόνα 3.11** Παραδείγματα αποτελεσμάτων στο στιγμιότυπο της Εικόνας 3.2, των εκφράσεων (α)  $\pi_{Genre}(MOVIE)$ , (β)  $\pi_{Name, Sex, Nationality}(ACTOR)$ , (γ)  $\pi_{Genre}(\sigma_{Year > 2000}(MOVIE))$

Προσέξτε ότι επειδή το αποτέλεσμα της προβολής είναι μια σχέση, δεν επιτρέπεται να υπάρχουν σε αυτήν διπλότυπα, κατά συνέπεια πλειάδες της  $R$  με τις ίδιες τιμές στα γνωρίσματα της  $L$  εμφανίζονται μόνο μια φορά στο αποτέλεσμα της προβολής. Αυτή η διαδικασία ονομάζεται *απαλοιφή διπλότυπων* (*duplicate elimination*).

Ας δούμε μερικά παραδείγματα χρησιμοποιώντας πάλι τη βάση δεδομένων για τις ταινίες. Η έκφραση  $\pi_{Genre}(MOVIE)$  έχει ως αποτέλεσμα μια σχέση που έχει όλα τα (διαφορετικά) είδη ταινιών, ενώ η έκφραση  $\pi_{Name, Sex, Nationality}(ACTOR)$  μας δίνει μια σχέση που έχει όλα τα γνωρίσματα για τους ηθοποιούς εκτός αυτών που αφορούν την ημερομηνία γέννησής τους.

Τέλος, μπορούμε να συνδυάσουμε πράξεις επιλογής και προβολής για να διατυπώσουμε πιο περίπλοκα ερωτήματα. Για παράδειγμα, η έκφραση  $\pi_{Genre}(\sigma_{Year > 2000}(MOVIE))$  έχει ως αποτέλεσμα τα είδη των ταινιών που γυρίστηκαν μετά το 2000. Το αποτέλεσμα των παραπάνω ερωτήσεων στο στιγμιότυπο της Εικόνας 3.2 δίνεται στην Εικόνα 3.11.

Ο βαθμός της σχέσης που προκύπτει ως αποτέλεσμα της προβολής είναι ίσος με το πλήθος των γνωρισμάτων τη λίστας  $L$ , ενώ η *πληθικότητα* της είναι μικρότερη ή ίση με την πληθικότητα της αρχικής σχέσης. Συγκεκριμένα, το πλήθος των πλειάδων της σχέσης που προκύπτει είναι ίσο με το πλήθος των πλειάδων της αρχικής σχέσης αν δε γίνεται απαλοιφή διπλότυπων, το οποίο συμβαίνει πάντα αν η λίστα  $L$  περιέχει τα γνωρίσματα ενός κλειδιού.

Η προβολή δεν είναι αντιμεταθετική. Συγκεκριμένα, η έκφραση  $\pi_{L1}(\pi_{L2}(R)) = \pi_{L1}(R)$ , αν  $L1 \subseteq L2$ , αλλιώς η έκφραση δεν ορίζεται.

### Άσκηση Αυτοαξιολόγησης 3.3: Τελεστές επιλογής και προβολής

Έστω το σχεσιακό σχήμα για μια βάση δεδομένων που περιέχει πληροφορίες για πίτσες και αποτελείται από τα παρακάτω τρία σχήματα σχέσεων: PIZZA(Name, Ingredient), που περιέχει τα συστατικά κάθε πίτσας, LIKES(Student, Ingredient) που περιέχει πληροφορία για το ποια συστατικά αρέσουν σε ποιον φοιτητή και SERVES(Place, Pizza-name) που περιέχει πληροφορία για το ποια μαγαζιά σερβίρουν ποιες πίτσες. Στην Εικόνα 3.12, βλέπετε ένα στιγμιότυπο αυτής της βάσης δεδομένων.

(α) Δώστε εκφράσεις της σχεσιακής άλγεβρας που να έχουν ως αποτέλεσμα:

- (i) τα ονόματα των πιτσών που περιέχουν μανιτάρι (mushroom),
- (ii) τα συστατικά της πίτσας με όνομα Special,
- (iii) τα συστατικά που αρέσουν τουλάχιστον σε έναν φοιτητή.

PIZZA		LIKES		SERVES		
Name	Ingredient	Student	Ingredient	Place	Pizza-name	Price
Vegetarian	olive	Dimitris	mushroom	Mama's	Vegetarian	8
Vegetarian	mushroom	Costas	ham	Mama's	Special	11
Vegetarian	onion	Maria	olive	Tony-pizza	Vegetarian	7
Hawaiian	pineapple	Katerina	mushroom	Tony-pizza	Greek	9
Hawaiian	ham	Maria	ham	Pizza-place	Hawaiian	8
Special	ham	Dimitris	bacon	Pizza-place	Special	9
Special	bacon	Maria	pineapple	Pizza-place	Greek	8
Special	mushroom			Napoli	Special	10
Greek	olive					
Greek	feta-cheese					

**Εικόνα 3.12** Ένα στιγμιότυπο του σχήματος της σχεσιακής βάσης δεδομένων για τις πίτσες

(β) Υπολογίστε τα αποτελέσματα των παραπάνω εκφράσεων στο στιγμιότυπο της Εικόνας 3.12.

### 3.3.3 Συνδυασμός Τελεστών και Μετονομασία

Η σχέση που προκύπτει ως αποτέλεσμα της εφαρμογής των τελεστών της σχεσιακής άλγεβρας δεν έχει όνομα. Σε ορισμένες περιπτώσεις διευκολύνει να δώσουμε κάποιο όνομα στο αποτέλεσμα μιας έκφρασης της σχεσιακής άλγεβρας, ιδιαίτερα αν θέλουμε

να το χρησιμοποιήσουμε ως ενδιάμεσο για το σχηματισμό πιο πολύπλοκων εκφράσεων.

Για παράδειγμα, θεωρείστε την έκφραση  $\pi_{\text{Genre}}(\sigma_{\text{Year} > 2000}(\text{MOVIE}))$ . Με το συμβολισμό,  $\text{NEW-MOVIE} \leftarrow \sigma_{\text{Year} > 2000}(\text{MOVIE})$  ονομάζουμε  $\text{NEW-MOVIE}$  τη σχέση που προκύπτει από την εφαρμογή της επιλογής. Η αρχική έκφραση μπορεί τώρα να γραφεί χρησιμοποιώντας τη  $\text{NEW-MOVIE}$  ως  $\pi_{\text{Genre}}(\text{NEW-MOVIE})$ .

Επίσης, μπορούμε να μετονομάσουμε τα γνωρίσματα μιας σχέσης. Με το συμβολισμό  $\text{NEW-MOVIE}(\text{Title}, \text{Year}, \text{New-genre}, \text{Runtime}) \leftarrow \sigma_{\text{Year} > 2000}(\text{MOVIE})$  ονομάζουμε πάλι τη σχέση που προκύπτει  $\text{NEW-MOVIE}$  αλλά παράλληλα μετονομάζουμε το γνώρισμα της από  $\text{Genre}$  σε  $\text{New-genre}$ . Η αρχική έκφραση μπορεί τώρα να γραφεί ως  $\pi_{\text{New-genre}}(\text{NEW-MOVIE})$ .

Τέλος, εναλλακτικά συχνά χρησιμοποιείται και ένας *τελεστής μετονομασίας* (*rename operator*) που συμβολίζεται με  $\rho$  όπου για μια σχέση  $R(A_1, A_2, \dots, A_n)$  η έκφραση  $\rho_{S(B_1, B_2, \dots, B_n)}(R)$  είναι ισοδύναμη του συμβολισμού  $S(B_1, B_2, \dots, B_n) \leftarrow R(A_1, A_2, \dots, A_n)$ . Τόσο η σχέση  $S$ , όσο και η λίστα γνωρισμάτων  $B_1, B_2, \dots, B_n$  είναι προαιρετικές και όταν λείπουν αντίστοιχα δε δίνουμε όνομα στη νέα σχέση ή η νέα σχέση διατηρεί τα ίδια ονόματα γνωρισμάτων με την  $R$ .

### 3.3.4 Τελεστές Συνόλων

Ο ορισμός της σχέσης ως συνόλου πλειάδων επιτρέπει την εφαρμογή των τυπικών σύνολο-θεωρητικών τελεστών της ένωσης, τομής, διαφοράς και καρτεσιανού γινομένου.

**Ένωση, τομή και διαφορά.** Για να εφαρμόσουμε έναν από τους τελεστές της ένωσης, τομής και διαφοράς σε δύο σχέσεις  $R$  και  $S$  πρέπει αυτές οι δύο σχέσεις να είναι συμβατές, δηλαδή να περιέχουν πλειάδες του ίδιου τύπου. Συγκεκριμένα, λέμε ότι δυο σχέσεις  $R(A_1, A_2, \dots, A_n)$  και  $S(B_1, B_2, \dots, B_n)$  είναι *συμβατές ως προς την ένωση* (*union compatible*) αν έχουν τον ίδιο βαθμό  $n$  και επίσης  $\text{dom}(A_i) = \text{dom}(B_i)$  για  $1 \leq i \leq n$ . Το αποτέλεσμα της εφαρμογής της ένωσης, τομής και διαφοράς είναι μια νέα σχέση η οποία, κατά σύμβαση, θεωρούμε ότι έχει τα ίδια γνωρίσματα με την πρώτη από τις δύο σχέσεις. Η ένωση  $R \cup S$  περιέχει τις πλειάδες που ανήκουν στην  $R$ , στην  $S$  ή και στις δύο, η τομή  $R \cap S$  περιέχει τις πλειάδες που ανήκουν και στην  $R$  και στην  $S$  και τέλος η διαφορά  $R - S$  περιέχει τις πλειάδες που ανήκουν στην  $R$  αλλά δεν ανήκουν στην  $S$ .

R			
Title	Year	Genre	Runtime
Gravity	2013	science-fiction	91
Frozen	2013	comedy	102

(α)

S			
Title	Year	Genre	Runtime
Amour	2012	drama	117
Ocean's Eleven	1960	crime	127
Frozen	2013	comedy	102
Ocean's Eleven	2001	crime	116

(β)

R ∪ S			
Title	Year	Genre	Runtime
Amour	2012	drama	117
Gravity	2013	science-fiction	91
Ocean's Eleven	1960	crime	127
Frozen	2013	comedy	102
Ocean's Eleven	2001	crime	116

(γ)

R ∩ S			
Title	Year	Genre	Runtime
Frozen	2013	comedy	102

(δ)

R - S			
Title	Year	Genre	Runtime
Gravity	2013	science-fiction	91

(ε)

**Εικόνα 3.13** Παραδείγματα αποτελεσμάτων πράξεων συνόλου στο στιγμιότυπο της Εικόνας 3.2: (α)  $R$ , (β)  $S$ , (γ)  $R \cup S$ , (δ)  $R \cap S$  και (ε)  $R - S$ .

Ας δούμε μερικά παραδείγματα χρησιμοποιώντας πάλι τη βάση δεδομένων για τις ταινίες. Ας ορίσουμε ως  $R \leftarrow \sigma_{\text{Runtime} \leq 110}(\text{MOVIE})$  και  $S \leftarrow \sigma_{\text{Runtime} \geq 100}(\text{MOVIE})$ . Οι  $R$  και  $S$  είναι συμβατές ως προς την ένωση. Στην Εικόνα 3.13, βλέπετε τις σχέσεις  $R \cup S$ ,  $R \cap S$  και  $R - S$ . Προσέξτε ότι στην περίπτωση της ένωσης, γίνεται και απαλοιφή διπλότυπων, αφού το αποτέλεσμα μιας πράξης της σχεσιακής άλγεβρας, ορίζεται ως σχέση, δηλαδή ως σύνολο.

Για τους τελεστές της ένωσης και της τομής ισχύει η αντιμεταθετική ιδιότητα, δηλαδή ισχύει αντίστοιχα  $R \cup S = S \cup R$  και  $R \cap S = S \cap R$ . Επίσης, ισχύει και η προσεταιριστική ιδιότητα, δηλαδή ισχύει αντίστοιχα  $R \cup (S \cup T) = (R \cup S) \cup T$  και  $R \cap (S \cap T) = (R \cap S) \cap T$ . Για τον τελεστή της διαφοράς δεν ισχύει η μεταθετική ιδιότητα και γενικά  $S - R \neq R - S$ . Για παράδειγμα, δείτε στην Εικόνα 3.13 ότι  $S - R \neq R - S$ , αφού το  $S - R$  περιλαμβάνει όλες τις πλειάδες της  $S$  εκτός από την τρίτη.

**Καρτεσιανό γινόμενο.** Ο τελεστής του καρτεσιανού γινομένου ή χιαστί γινομένου (*cross product*) ή χιαστί συνένωση (*cross join*) χρησιμοποιείται για να συνδυάσουμε πληροφορίες από δύο σχέσεις. Το αποτέλεσμα του καρτεσιανού γινομένου μεταξύ δυο σχέσεων  $R(A_1, A_2, \dots, A_n)$  και  $S(B_1, B_2, \dots, B_m)$  συμβολίζεται με  $R \times S$  και είναι μια σχέση, έστω  $Q$ , με  $n + m$  γνωρίσματα: τα γνωρίσματα της  $R$  με τη σειρά που εμφανίζονται στην  $R$  ακολουθούμενα από τα γνωρίσματα της  $S$  με τη σειρά που εμφανίζονται στην  $S$ . Η  $Q$  αποτελείται από τις πλειάδες που προκύπτουν αν συνδυάσουμε κάθε μια πλειάδα της  $R$  με κάθε μια πλειάδα της  $S$ . Άρα, αν η  $R$  έχει  $n_R$  πλειάδες και η  $S$  έχει  $n_S$ , τότε η  $Q$  έχει  $n_R * n_S$  πλειάδες.



Θα κάνουμε τη σύμβαση ότι αν ένα γνώρισμα της  $R$  έχει το ίδιο όνομα με ένα γνώρισμα της  $S$  αυτό εμφανίζεται ανώνυμο στην  $Q$ . Σε αυτές τις περιπτώσεις, ενδείκνυται να μετονομάσουμε τα γνωρίσματα ώστε να αποφύγουμε τέτοιες συγκρούσεις στην ονοματολογία.

<b>R</b>	
<b>A</b>	<b>B</b>
1	2
5	2
8	5
2	7

<b>S</b>	
<b>B</b>	<b>C</b>
2	3
3	4
5	9

<b>A</b>	<b>R-B</b>	<b>S-B</b>	<b>C</b>
1	2	2	3
1	2	3	4
1	2	5	9
5	2	2	3
5	2	3	4
5	2	5	9
8	5	2	3
8	5	3	4
8	5	5	9
2	7	2	3
2	7	3	4
2	7	5	9

(α)
(β)
(γ)

**Εικόνα 3.14** Παράδειγμα καρτεσιανού γινομένου

Ας δούμε πρώτα ένα απλό παράδειγμα. Έστω οι σχέσεις  $R$  και  $S$  που δίνονται στην Εικόνα 3.14(α) και 3.14(β) αντίστοιχα. Το καρτεσιανό τους γινόμενο  $\rho_{(A, R-B)}(R) \times \rho_{(S-B, C)}(S)$ , όπου μετονομάσαμε τα γνωρίσματα με το ίδιο όνομα, δίνεται στην Εικόνα 3.14(γ). Το καρτεσιανό γινόμενο έχει 4 γνωρίσματα (το σύνολο των γνωρισμάτων των δύο σχέσεων) και 12 πλειάδες (που προκύπτουν από τους συνδυασμούς των 4 πλειάδων της  $R$  με τις 3 πλειάδες της  $S$ ).

Ας δούμε ακόμα ένα παράδειγμα καρτεσιανού γινομένου χρησιμοποιώντας τη βάση δεδομένων για τις ταινίες. Υποθέστε ότι θέλετε να δείτε σε ποια είδη ταινιών παίζει κάποιος ηθοποιός. Η πληροφορία αυτή υπάρχει σε δύο διαφορετικές σχέσεις, την PLAYS και τη σχέση MOVIE, τις οποίες θα συνενώσουμε με τη χρήση καρτεσιανού γινομένου. Ας μετονομάσουμε πρώτα τα κοινά γνωρίσματα Title και Year των

PLAYS και MOVIE ως εξής:  $PLAYS1(Name, Plays-Title, Plays-Year, Role) \leftarrow PLAYS(Name, Title, Year, Role)$  και  $MOVIE1(Movie-Title, Movie-Year, Genre, Runtime) \leftarrow MOVIE(Title, Year, Genre, Runtime)$ . Στην Εικόνα 3.15, βλέπετε το αποτέλεσμα του καρτεσιανού γινομένου  $Q \leftarrow PLAYS1 \times MOVIE1$  στο στιγμιότυπο της Εικόνας 3.2. Εναλλακτικά, το  $Q$  θα μπορούσαμε να το διατυπώσουμε χρησιμοποιώντας τον τελεστή μετονομασίας ως  $Q \leftarrow \rho(Name, Plays-Title, Plays-Year, Role)(PLAYS) \times \rho(Movie-Title, Movie-Year, Genre, Runtime)(MOVIE)$ .

Name	Play-Title	Play-Year	Role	Movie-Title	Movie-Year	Genre	Runtime
George Clooney	Gravity	2013	Mark Kowalski	Amour	2012	drama	117
George Clooney	Gravity	2013	Mark Kowalski	Gravity	2013	science-fiction	91
			...				
George Clooney	Gravity	2013	Mark Kowalski	Ocean's Eleven	2001	crime	116
			..				
Brad Pitt	Ocean's Eleven	2001	Rusty Ryan	Amour	2012	drama	117
Brad Pitt	Ocean's Eleven	2001	Rusty Ryan	Gravity	2013	science-fiction	91
			...				
Brad Pitt	Ocean's Eleven	2001	Rusty Ryan	Ocean's Eleven	2001	crime	116

**Εικόνα 3.15** Παράδειγμα καρτεσιανού γινομένου στο στιγμιότυπο της Εικόνας 3.2

Το αποτέλεσμα περιέχει τις 30 πλειάδες που προκύπτουν από τους συνδυασμούς των 6 πλειάδων της PLAYS με κάθε μια από τις 5 πλειάδες της MOVIE. Στο αποτέλεσμα έχουμε την πληροφορία και από τις δύο σχέσεις αλλά πολλές από τις πλειάδες δεν αφορούν ταινίες στις οποίες έπαιξε ένας ηθοποιός, π.χ., η πρώτη πλειάδα. Για να πάρουμε πληροφορία για τις ταινίες στις οποίες έπαιξε ένας ηθοποιός πρέπει να τις επιλέξουμε ως εξής  $\sigma_{Plays-Title = Movie-Title \text{ AND } Plays-Year = Movie-Year}(Q)$ , ώστε κάθε πλειάδα να αφορά την ίδια ταινία, δηλαδή να εφαρμόσουμε μια πράξη επιλογής στο αποτέλεσμα του καρτεσιανού γινομένου.

#### **Άσκηση Αυτοαξιολόγησης 3.4:** Ένωση, Τομή, Διαφορά

Θεωρείστε τη βάση δεδομένων με πληροφορίες για πίτσες της Άσκησης Αυτοαξιολόγησης 3.3. Δώστε εκφράσεις της σχεσιακής άλγεβρας που να έχουν ως αποτέλεσμα τις πίτσες (τα ονόματά τους) που:

- (i) δεν περιέχουνμανιτάρι,
- (ii) περιέχουνμανιτάρι ή ζαμπόν (ham),
- (iii) περιέχουνμανιτάρι και ζαμπόν,

(iv) περιέχουν μανιτάρι αλλά όχι ζαμπόν.

### Άσκηση Αυτοαξιολόγησης 3.5: Καρτεσιανό Γινόμενο

Θεωρείστε τη βάση δεδομένων με πληροφορίες για πίτσες της Άσκησης Αυτοαξιολόγησης 3.3. Δώστε μια έκφραση της σχεσιακής άλγεβρας που να έχει ως αποτέλεσμα τις πίτσες (τα ονόματα τους) που περιέχουν (τουλάχιστον) ένα συστατικό που να αρέσει στο φοιτητή Δημήτρη.

#### 3.3.5 Συνένωση

Πολύ συχνά θέλουμε να συνδυάσουμε πληροφορία από δύο σχέσεις και να επιλέξουμε από τις πλειάδες που προκύπτουν αυτές που ικανοποιούν κάποια συνθήκη, όπως είδαμε και στο παράδειγμα με τα είδη των ταινιών που παίζει ένας ηθοποιός. Αυτό εκφράζεται μέσω του τελεστή της *συνένωσης* (*join*). Η συνένωση δύο σχέσεων  $R$  και  $S$  συμβολίζεται με το σύμβολο  $\bowtie$  και ορίζεται ως:  $R \bowtie_P S \equiv \sigma_P (R \times S)$ . Η συνθήκη  $P$  καλείται *συνθήκη συνένωσης* και είναι μια λογική έκφραση που αποτελείται από μία ή περισσότερες λογικές προτάσεις της μορφής  $A \text{ op } B$  οι οποίες συνδέονται με το λογικό τελεστή AND, όπου  $A$  είναι ένα γνώρισμα της  $R$ ,  $op$  ένας τελεστής σύγκρισης και  $B$  ένα γνώρισμα της  $S$ .

Η πιο συνηθισμένη μορφή συνένωσης είναι αυτή στην οποία η συνθήκη συνένωσης περιλαμβάνει μόνο τον τελεστή ισότητας, όπως είδαμε και στο παράδειγμα με τα είδη ταινιών. Αυτή η συνένωση καλείται *συνένωση ισότητας* (*equijoin*), ενώ η συνένωση της οποίας η συνθήκη περιλαμβάνει και άλλους τελεστές σύγκρισης συχνά καλείται και *συνένωση θήτα* (*theta join*).

Ας δούμε μερικά παραδείγματα. Θεωρείστε τις σχέσεις  $R$  και  $S$  της Εικόνας 3.14. Στην Εικόνα 3.16(α), μπορείτε να δείτε το αποτέλεσμα της συνένωσης θήτα  $\rho_{(A, R-B)}(R) \bowtie_{A > C} \rho_{(S-B, C)}(S)$ , όπου από το καρτεσιανό γινόμενο των δύο σχέσεων επιλέγονται μόνο οι πλειάδες για τις οποίες ισχύει η συνθήκη  $A > C$ . Στην Εικόνα 3.16(β), μπορείτε να δείτε το αποτέλεσμα της συνένωσης ισότητας  $\rho_{(A, R-B)}(R) \bowtie_{R-B = C-B} \rho_{(S-B, C)}(S)$ , όπου από το καρτεσιανό γινόμενο των δύο σχέσεων επιλέγονται οι πλειάδες που έχουν την ίδια τιμή στο γνώρισμα  $B$ .

Μια ειδική περίπτωση συνένωσης ισότητας είναι η *φυσική συνένωση* (*natural join*) που συμβολίζεται με το σύμβολο  $*$ . Η φυσική συνένωση  $R * S$  δύο σχέσεων  $R$  και  $S$  επιλέγει από το καρτεσιανό γινόμενο  $R \times S$  τις πλειάδες που έχουν την ίδια τιμή στα γνωρίσματα των  $R$  και  $S$  με το ίδιο όνομα. Στο αποτέλεσμα της φυσικής συνένωσης,

τα γνωρίσματα με το ίδιο όνομα εμφανίζονται μόνο μια φορά, κατά σύμβαση ως γνωρίσματα της πρώτης σχέσης.

A	R-B	S-B	C
5	2	2	3
5	2	3	4
8	5	2	3
8	5	3	4

(α)

A	R-B	S-B	C
1	2	2	3
5	2	2	3
8	5	5	9

(β)

A	B	C
1	2	3
5	2	3
8	5	9

(γ)

**Εικόνα 3.16** Παράδειγμα συνενώσεων των σχέσεων R και S της Εικόνας 3.14: (α)  $\rho_{(A, R-B)}(R) \bowtie_{A > C} \rho_{(S-B, C)}(S)$ , (β)  $\rho_{(A, R-B)}(R) \bowtie_{R-B = C, B} \rho_{(S-B, C)}(S)$ , (γ)  $R * S$

Στην Εικόνα 3.16(γ), βλέπετε το αποτέλεσμα της φυσικής συνένωσης  $R * S$  για τις σχέσεις της Εικόνας 3.14. Παρατηρείστε ότι σε αντίθεση με τη συνένωση ισότητας στο κοινό γνώρισμα που δίνεται στην Εικόνα 3.16(β), το κοινό γνώρισμα στην περίπτωση της φυσικής συνένωσης εμφανίζεται μόνο μια φορά στο αποτέλεσμα.

### Άσκηση Αυτοαξιολόγησης 3.6: Συνενώσεις

Θεωρείστε τη βάση δεδομένων με πληροφορίες για πίτσες της Άσκησης Αυτοαξιολόγησης 3.3.

(α) Δώστε εκφράσεις της σχεσιακής άλγεβρας που έχουν ως αποτέλεσμα τα μαγαζιά που σερβίρουν (τουλάχιστον) μία πίτσα που περιέχει:

- (i) μανιτάρι,
- (ii) (τουλάχιστον) ένα συστατικό που αρέσει στον Δημήτρη.

(β) Δώστε εκφράσεις της σχεσιακής άλγεβρας που έχουν ως αποτέλεσμα τις πίτσες (τα ονόματα τους) που έχουν:

- (i) τουλάχιστον δυο διαφορετικά συστατικά,
- (ii) μόνο ένα συστατικό,
- (iii) ακριβώς δύο συστατικά.

### 3.3.6 Εξωτερική Συνένωση

Στις πράξεις συνένωσης που είδαμε μέχρι τώρα, οι πλειάδες των αρχικών σχέσεων που δεν συνενώνονται με καμία πλειάδα της άλλης σχέσης, δεν εμφανίζονται στο αποτέλεσμα. Οι πράξεις της εξωτερικής συνένωσης επεκτείνουν τη συνένωση, ώστε και οι πλειάδες που δεν συνενώνονται να εμφανίζονται στο αποτέλεσμα.

Συγκεκριμένα, η *αριστερή εξωτερική συνένωση (left outer join)* δύο σχέσεων  $R$  και  $S$  συμβολίζεται με  $R \bowtie_P S$  και έχει ως αποτέλεσμα τη συνένωση των  $R$  και  $S$  με τη διαφορά ότι οι πλειάδες της σχέσης  $R$  που δε συνενώνονται με καμία πλειάδα της  $S$ , εμφανίζονται στο αποτέλεσμα με τιμή null στα γνωρίσματα της  $S$ . Αντίστοιχα, η *δεξιά εξωτερική συνένωση (right outer join)* δύο σχέσεων  $R$  και  $S$ ,  $R \ltimes_P S$ , έχει ως αποτέλεσμα τη συνένωση των  $R$  και  $S$  μόνο που τώρα οι πλειάδες της  $S$  που δε συνενώνονται με καμία πλειάδα της  $R$  εμφανίζονται στο αποτέλεσμα με τιμή null στα γνωρίσματα της  $R$ . Τέλος, στο αποτέλεσμα της *πλήρους εξωτερική συνένωση (full outer join)*  $R \bowtie_{\text{full}} S$ , δύο σχέσεων  $R$  και  $S$  εμφανίζονται και οι πλειάδες της  $R$  και οι πλειάδες της  $S$  που δε συνενώνονται με τιμή null στα άλλα γνωρίσματα.

Ένα παράδειγμα δίνεται στην Εικόνα 3.17. Η διαφορά με το αποτέλεσμα της αντίστοιχης συνένωσης 3.16(β), είναι ότι στην περίπτωση (α) της αριστερής εξωτερικής συνένωσης η πλειάδα (2, 7) της  $R$  που δε συνενώνονταν, τώρα εμφανίζεται στο αποτέλεσμα. Αντίστοιχα στην περίπτωση (β) της δεξιάς εξωτερικής συνένωσης εμφανίζεται η πλειάδα (3, 4) της  $S$  και στην περίπτωση (γ) της πλήρους εξωτερικής συνένωσης εμφανίζονται και οι δύο πλειάδες.

A	R-B	S-B	C
1	2	2	3
5	2	3	4
8	5	5	9
2	7	null	null

(α)

A	R-B	S-B	C
1	2	2	3
5	2	2	3
8	5	5	9
null	null	3	4

(β)

A	R-B	S-B	C
1	2	2	3
5	2	2	3
8	5	5	9
2	7	null	null
null	null	3	4

(γ)

**Εικόνα 3.17** Παράδειγμα εξωτερικής συνένωσης των σχέσεων  $R$  και  $S$  της Εικόνας 3.14: (α)  $\rho_{(A, R-B)}(R) \bowtie_{R-B=C-B} \rho_{(S-B, C)}(S)$ , (β)  $\rho_{(A, R-B)}(R) \ltimes_{R-B=C-B} \rho_{(S-B, C)}(S)$ , (γ)  $\rho_{(A, R-B)}(R) \bowtie_{\text{full}} \rho_{(S-B, C)}(S)$ .

### 3.3.7 Ο τελεστής της Διαίρεσης

Ο τελεστής της διαίρεσης χρησιμοποιείται για να εκφράσει ερωτήσεις ειδικής μορφής, όπως «βρείτε τους ηθοποιούς που έχουν παίξει σε όλες τις ταινίες που έχει παίξει ο ηθοποιός George Clooney». Αρχικά, θα δώσουμε τον τυπικό ορισμό της διαίρεσης και στη συνέχεια θα δείξουμε πως μπορούμε να διατυπώσουμε τη διαίρεση με συνδυασμό άλλων τελεστών της σχεσιακής άλγεβρας

Ο τελεστής της διαίρεση  $R \div S$  ορίζεται μεταξύ δύο σχέσεων  $R(X)$  και  $S(Y)$  όπου  $X \supseteq Y$ , δηλαδή, το σύνολο των γνωρισμάτων του διαιρέτη είναι υπερσύνολο των γνωρισμάτων του διαιρετέου και έχει ως αποτέλεσμα μια σχέση  $Q(Z)$  όπου  $Z = X - Y$ , δηλαδή μια σχέση με γνωρίσματα τα μη κοινά γνωρίσματα των  $R$  και  $S$ . Μια πλειάδα  $t$  ανήκει στην  $Q$ , αν υπάρχει μια πλειάδα  $r$  της  $R(X)$  τέτοια ώστε  $r[Z] = t[Z]$  και  $r[Y] = s[Y]$ , για κάθε πλειάδα  $s$  της  $S$ . Δηλαδή, ως αποτέλεσμα παίρνουμε τις τιμές του διαιρέτη που εμφανίζονται με όλες τις πλειάδες του διαιρετέου.

Στην Εικόνα 3.18, μπορείτε να δείτε μερικά παραδείγματα. Ας δούμε αναλυτικά το πρώτο από αυτά, δηλαδή τη διαίρεση  $R1 \div S1$ . Σε αυτό το παράδειγμα,  $X = \{A, B\}$  και  $Y = \{B\}$ . Άρα, το αποτέλεσμα θα έχει το σύνολο γνωρισμάτων  $Z = X - Y = \{A\}$ . Για να ανήκει μια πλειάδα στο αποτέλεσμα, θα πρέπει η τιμή της να εμφανίζεται στην  $R1$  με όλες τις τιμές του  $B$  στην  $S1$ . Στο παράδειγμα, αυτό συμβαίνει μόνο για την  $a2$ .

<b>R1</b>	<b>A</b>	<b>B</b>	<b>S1</b>	<b>B</b>	<b>R2</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>S2</b>	<b>A</b>	<b>S3</b>	<b>A</b>	<b>B</b>		
a1	a1	a1	b2	a1	a1	a1	a2	a2	a1	a2	a1	a1	b1		
a1	a1	a1	b4	a2	a2	a2	b1	a2	a2	a3	a2	a2	b2		
a2	a2	a2	b2	a3	a3	a3	b2	a3	a3	a3	a3	a3	b1		
a2	a2	a2	b4	a2	a3	a3	b1	a3	a3	a3	a3	a3	b2		
a3	a3	a3	b2	a2	a3	a3	b1	a3	a3	a3	a3	a3	b2		
			<b>R1 ÷ S1</b>							<b>R2 ÷ S2</b>				<b>R2 ÷ S3</b>	
			a2							b1	c1				c1
			a2							b1	c1				c2

**Εικόνα 3.18** Παραδείγματα διαίρεσης

Ας διατυπώσουμε τώρα χρησιμοποιώντας διαίρεση το ερώτημα με τους ηθοποιούς που παίζουν σε όλες τις ταινίες που παίζει ο George Clooney για τη βάση δεδομένων με τις ταινίες της Εικόνας 3.2. Ένας ηθοποιός για να ανήκει στο αποτέλεσμα θα πρέπει να εμφανίζεται στην PLAYS με όλες τις ταινίες που παίζει ο George Clooney.

Οπότε θα πρέπει να διαιρέσουμε τη σχέση PLAYS με τις ταινίες που παίζει ο George Clooney. Ας το δούμε αναλυτικά.

$$\text{MOVIES-CLOONEY} \leftarrow \pi_{\text{Title, Year}} (\sigma_{\text{Name} = \text{"George Clooney"}}(\text{PLAYS}))$$

$$Q \leftarrow \pi_{\text{Name, Title, Year}}(\text{PLAYS}) \div \text{MOVIES-CLOONEY}$$

Προσέξτε ότι πριν διαιρέσουμε την PLAYS πρέπει να προβάλουμε τα γνωρίσματα που μας ενδιαφέρουν, ουσιαστικά να μην κρατήσουμε το Role.

Η διαίρεση μπορεί να διατυπωθεί χρησιμοποιώντας άλλες πράξεις της σχεσιακής άλγεβρας. Η ισοδύναμη έκφραση δεν είναι προφανής. Η βασική ιδέα είναι να βρούμε και να αφαιρέσουμε το σύνολο, έστω  $Q'$ , των πλειάδες που δεν ανήκουν στο αποτέλεσμα, δηλαδή, στο παραπάνω παράδειγμα, τους ηθοποιούς που δεν παίζουν σε όλες τις ταινίες του George Clooney. Ας σκεφτούμε τι συμβαίνει με αυτούς τους ηθοποιούς. Θεωρείστε το συνδυασμό τους με όλες τις ταινίες στη σχέση MOVIES-CLOONEY. Ένας ηθοποιός δεν ανήκει στο αποτέλεσμα αν κάποιος, έστω ένας, από αυτούς τους συνδυασμούς δεν ανήκει στην PLAYS. Δηλαδή, αν αφαιρέσουμε από αυτό το συνδυασμό την PLAYS, αυτός ο ηθοποιός θα ανήκει στο αποτέλεσμα της αφαίρεσης. Το παραπάνω διατυπώνεται ως εξής:  $Q' \leftarrow (\pi_{\text{Name}}(\text{PLAYS}) \times \text{MOVIES-CLOONEY}) - \pi_{\text{Name, Title, Year}}(\text{PLAYS})$

Γενικεύοντας το σύνολο  $Q'$  για τη διαίρεση δυο σχέσεων  $R(X)$  και  $S(Y)$ , μπορεί να υπολογιστεί ως εξής:

$$Q' \leftarrow (\pi_Z(R) \times S) - R$$

Τώρα, αρκεί να αφαιρέσουμε από όλους τους ηθοποιούς αυτούς που εμφανίζονται στο σύνολο  $Q'$ , δηλαδή,  $Q \leftarrow \pi_{\text{Name}}(\text{PLAYS}) - \pi_{\text{Name}}(Q')$  και γενικεύοντας:

$$Q \leftarrow \pi_Z(R) - \pi_Z(Q').$$

Η διαίρεση έχει κάποια ομοιότητα με την ακέραια διαίρεση, όπου το πηλίκο  $A/B$  δύο ακεραίων, είναι ο μεγαλύτερος ακέραιος  $X$  τέτοιος ώστε  $X * B \leq A$ . Κατά αναλογία, για τη διαίρεση στη σχεσιακή άλγεβρα, ισχύει  $Q \times S \subseteq R$ .

### Άσκηση Αυτοαξιολόγησης 3.7: Διαίρεση

Διατυπώστε μια ερώτηση στο παράδειγμα με τις πίτσες της Άσκησης Αυτοαξιολόγησης 3.3 που ο φυσικός τρόπος για να διατυπωθεί στη σχεσιακή άλγεβρα είναι με χρήση του τελεστή της διαίρεσης. Στη συνέχεια, εκφράστε αυτήν την ερώτηση (i) χρησιμοποιώντας διαίρεση και (ii) με τον εναλλακτικό τρόπο που χρησιμοποιεί αφαίρεση.

### 3.3.8 Ελάχιστο Σύνολο Τελεστών

Είδαμε ότι κάποιοι τελεστές της σχεσιακής άλγεβρας, για παράδειγμα η συνένωση και η διαίρεση, μπορούν να εκφραστούν χρησιμοποιώντας άλλους τελεστές της σχεσιακής άλγεβρας. Είναι εύκολο να δείτε ότι το σύνολο των τελεστών της σχεσιακής άλγεβρας που περιλαμβάνει την επιλογή, προβολή, ένωση, διαφορά και καρτεσιανό γινόμενο είναι ένα *πλήρες* (*complete*) σύνολο με την έννοια ότι οι άλλοι τελεστές μπορούν να διατυπωθούν χρησιμοποιώντας τους.

Για παράδειγμα, η πράξη της τομής  $R \cap S$ ,  $R \setminus S$  μεταξύ δύο σχέσεων  $R$  και  $S$  μπορεί να διατυπωθεί ως  $(R \cap S) = ((R - S) \cup (S - R))$ .

### Σύνοψη

Στο κεφάλαιο αυτό, παρουσιάστηκαν οι βασικές έννοιες του σχεσιακού μοντέλου και της σχεσιακής άλγεβρας.

Στο σχεσιακό μοντέλο, τα δεδομένα δομούνται σε σχέσεις ή πίνακες. Το σχήμα μιας σχέσης αποτελείται από το όνομα της σχέσης και μια λίστα από γνωρίσματα. Ένα στιγμιότυπο μιας σχέσης είναι ένα σύνολο πλειάδων, όπου κάθε πλειάδα είναι μια διατεταγμένη λίστα τιμών όπου υπάρχει μία τιμή για κάθε γνώρισμα.

Ένα στιγμιότυπο μιας σχέσης για να είναι αποδεκτό πρέπει να ικανοποιεί τους περιορισμούς ακεραιότητας που καθορίζονται στο σχήμα. Οι βασικοί περιορισμοί ακεραιότητας που υποστηρίζει το σχεσιακό μοντέλο είναι ο περιορισμός πεδίου ορισμού, ο περιορισμός κλειδιού και ο περιορισμός ξένου κλειδιού ή αναφορικής ακεραιότητας.

Συχνά, η περιγραφή του σχήματος μια βάσης δεδομένων γίνεται με χρήση ενός διαγράμματος ΟΣ, το οποίο στη συνέχεια μετατρέπεται σε ένα αντίστοιχο σχεσιακό σχήμα. Στη γενική περίπτωση, στο σχεσιακό μοντέλο, οι τύποι οντοτήτων αναπαρίστανται με σχέσεις με τα ίδια γνωρίσματα και πρωτεύοντα κλειδιά. Μια συσχέτιση αναπαρίσταται από μια σχέση που περιλαμβάνει ως γνωρίσματα τα πρωτεύοντα κλειδιά των οντοτήτων που συσχετίζει. Ο ορισμός του πρωτεύοντος κλειδιού αυτής της σχέσης εξαρτάται από την πληθικότητα της συσχέτισης που μοντελοποιεί. Τέλος, χρειάζεται ιδιαίτερη προσοχή για την αναπαράσταση πλειότιμων γνωρισμάτων, ασθενών οντοτήτων και ιεραρχιών.



Η σχεσιακή άλγεβρα αποτελεί την τυπική γλώσσα για τη διατύπωση ερωτήσεων σε μια σχεσιακή βάση δεδομένων. Οι βασικοί μοναδιαίοι τελεστές της σχεσιακής άλγεβρας είναι η επιλογή που επιλέγει τις πλειάδες μιας σχέσης που ικανοποιούν μια συνθήκη και η προβολή που προβάλλει συγκεκριμένα γνωρίσματα μιας σχέσης. Οι δυαδικοί τελεστές περιλαμβάνουν τους γνωστούς σύνολο-θεωρητικούς τελεστές και τον τελεστή της συνένωσης.

### ***Λίστα Ελέγχου Γνώσεων***

Έχοντας ολοκληρώσει τη μελέτη του κεφαλαίου που εισάγει το σχεσιακό μοντέλο και τη σχεσιακή άλγεβρα, αξιολογείστε το βαθμό κατανόησης των εννοιών που αναπτύχθηκαν, επιστρέφοντας στα αρχικά προσδοκώμενα αποτελέσματα και ελέγχοντας κατά πόσο μπορείτε:

- να αξιολογείστε αν ένα σχεσιακό σχήμα ικανοποιεί ή όχι και σε ποιο βαθμό τις απαιτήσεις δεδομένων μιας εφαρμογής,
- να σχεδιάσετε το σχήμα για μια σχεσιακή βάση δεδομένων χρησιμοποιώντας τα κατάλληλα δομικά στοιχεία και περιορισμούς ακεραιότητας που υποστηρίζει το σχεσιακό μοντέλο,
- να μετατρέψτε το εννοιολογικό σχήμα μιας βάσης δεδομένων από το μοντέλο ΟΣ σε ένα λογικό σχήμα χρησιμοποιώντας το σχεσιακό μοντέλο,
- να χρησιμοποιείτε τις πράξεις της σχεσιακής άλγεβρας (προβολή, επιλογή, συνένωση, πράξεις συνόλων) για τη διατύπωση ερωτήσεων σε μια σχεσιακή βάση δεδομένων.

### ***Ευρετήριο όρων***

βαθμός σχέσης

διαίρεση

διαφορά

ένωση

εξωτερική συνένωση

επιλογή

πίνακας

πλειάδα

πληθικότητα

προβολή

ξένο κλειδί, 48, 54, 55, 56, 57, 80

συνένωση

σχέση

σχεσιακή άλγεβρα, 60, 73, 74

σχεσιακό μοντέλο, 45, 49, 50, 51, 54, 56, 57, 60, 73, 74, 77

τομή

φυσική συνένωση

## Γλωσσάρι

**Βαθμός (degree ή arity) της σχέσης:** το πλήθος των γνωρισμάτων της σχέσης

**Πλειάδα (tuple):** μια γραμμή του πίνακα που συνήθως αντιστοιχεί σε μια οντότητα ή συσχέτιση του φυσικού κόσμου

**Πληθικότητα (cardinality) της σχέσης:** το πλήθος των πλειάδων

**Περιορισμός αναφορικής ακεραιότητας (referential integrity constraint) ή περιορισμός ξένου κλειδιού (foreign key constraint):** μεταξύ δύο σχημάτων σχέσεων  $R_1(X)$  και  $R_2(Y)$  προσδιορίζει ότι οι πλειάδες της  $R_1$  πρέπει να αφορούν πλειάδες που υπάρχουν στην  $R_2$ . Τυπικά, ένα σύνολο γνωρισμάτων  $F \subseteq X$  της  $R_1$  είναι ξένο κλειδί της  $R_1$  που αναφέρεται στην  $R_2$  αν (1) το σύνολο  $F$  αποτελείται από το ίδιο πλήθος και με το ίδιο πεδίο ορισμού γνωρισμάτων όπως και το πρωτεύον κλειδί  $K \subseteq Y$  της  $R_2$  και (2) σε οποιοδήποτε στιγμιότυπο της βάσης δεδομένων, για μια πλειάδα  $t_1$  της  $R_1$  ισχύει ότι είτε (α) όλα τα γνωρίσματα της  $t_1$  έχουν την τιμή null είτε (β) στο ίδιο στιγμιότυπο, υπάρχει μια πλειάδα  $t_2$  της  $R_2$ , τέτοια ώστε  $t_1[F] = t_2[K]$ .

**Σχέση (relation) ή πίνακας (table):** το βασικό δομικό στοιχείο του σχεσιακού μοντέλου

**Σχεσιακή άλγεβρα (relational algebra):** Το βασικό σύνολο τελεστών του σχεσιακού μοντέλου.

## Βιβλιογραφία

Codd E. F. (1970): *A Relational Model of Data for Large Shared Data Banks*. Commun. ACM Vol 13 (No 6), pp. 377-387.

Elmasri, R. & Navathe, S.B. (2003). *Fundamentals of Database Systems* (4th edition). Reading, MA: Addison Wesley.

Ramakrishnan, R. & Gehrke, J. (2002). *Database management systems* (3rd. edition). Boston: McGraw-Hill.

### **Οδηγός για περαιτέρω μελέτη**

Το σχεσιακό μοντέλο είναι το μοντέλο που υποστηρίζουν τα περισσότερα ΣΔΒΔ. Η περιγραφή του καλύπτεται από όλα τα διδακτικά βιβλία των βάσεων δεδομένων.

Αν θέλετε να δείτε περισσότερα παραδείγματα χρήσης του σχεσιακού μοντέλου για την αναπαράσταση δεδομένων μπορείτε να ανατρέξετε για παράδειγμα στο Κεφάλαιο 5 του βιβλίου των Elmasri, R. & Navathe, S.B. (2003) και στο Κεφάλαιο 3 του βιβλίου των Ramakrishnan, R. & Gehrke, J. (2002). Επίσης έχει ενδιαφέρον να μελετήσει κανείς το αρχικό άρθρο του Codd E. F. (1970) στο οποίο εισήγαγε το σχεσιακό μοντέλο.

Τέλος, περισσότερα παραδείγματα ερωτήσεων στη σχεσιακή άλγεβρα, μπορείτε να δείτε στο Κεφάλαιο 6 του βιβλίου των Elmasri, R. & Navathe, S.B. (2003) και στο Κεφάλαιο 4 του βιβλίου των Ramakrishnan, R. & Gehrke, J. (2002).

## **ΑΠΑΝΤΗΣΕΙΣ ΑΣΚΗΣΕΩΝ ΑΥΤΟΑΞΙΟΛΟΓΗΣΗΣ**

### **Απάντηση Άσκησης Αυτοαξιολόγησης 3.1**

(α) Η γενική αρχή είναι ότι από ένα στιγμιότυπο μπορούμε να δείξουμε ότι ένας περιορισμός ακεραιότητας δεν ισχύει αλλά δε μπορούμε να πούμε αν ισχύει. Αυτό συμβαίνει γιατί οι περιορισμοί ορίζονται στο σχήμα μιας βάσης δεδομένων και πρέπει να ισχύουν σε όλα τα στιγμιότυπά της. Το γεγονός ότι ένα περιορισμός δεν παραβιάζεται σε ένα συγκεκριμένο στιγμιότυπο δεν σημαίνει ότι ισχύει σε όλα τα στιγμιότυπα. Άρα (i) δεν ισχύει, γιατί η πρώτη και η τελευταία πλειάδα στον Πίνακα 3.1 έχουν την ίδια τιμή στο  $\{A\}$ , (ii) δεν ισχύει, γιατί η πρώτη και η τέταρτη πλειάδα

στον Πίνακα 3.1 έχουν την ίδια τιμή στο  $\{C, D\}$ , ενώ (iii) δεν μπορούμε να πούμε αν ισχύει.

Το (iv) ισχύει γιατί από τον ορισμό του υποψήφιου κλειδιού, κανένα γνήσιο υπερσύνολο ενός υποψήφιου κλειδιού δεν είναι υποψήφιο κλειδί.

(β) Η βασική ιδέα είναι ότι ο περιορισμός κλειδιού περιορίζει το πλήθος των πλειάδων γιατί δεν μπορούμε να έχουμε πλειάδες με τις ίδιες τιμές στα γνωρίσματα του κλειδιού.

(i) Ένα αποδεκτό στιγμιότυπο μπορεί να έχει το πολύ 500 πλειάδες, γιατί τόσες είναι και οι διαφορετικές τιμές που μπορεί να πάρει το γνώρισμα  $A$ . Αν είχαμε περισσότερες πλειάδες, τότε κάποιες από αυτές θα είχαν την ίδια τιμή στο  $A$ , που δεν είναι δυνατόν από τον περιορισμό του κλειδιού.

(ii) Το γεγονός ότι το  $\{A, B\}$  είναι υποψήφιο κλειδί σημαίνει ότι μπορούμε να έχουμε το πολύ  $500 * 100$  πλειάδες, δηλαδή, όσες και οι διαφορετικοί συνδυασμοί τιμών για τα γνωρίσματα  $A$  και  $B$ . Όμοια, το γεγονός ότι το  $\{B, C\}$  είναι υποψήφιο κλειδί σημαίνει ότι μπορούμε να έχουμε το πολύ  $100 * 300$  πλειάδες. Άρα αφού και τα δύο σύνολα είναι υποψήφια κλειδιά, ισχύουν και οι δύο παραπάνω περιορισμοί και ο μέγιστος αριθμός πλειάδων είναι 30.000, δηλαδή, ο μικρότερος από τους δύο.

(γ) Για να λύσουμε την άσκηση πρέπει να σκεφτούμε ότι υπερκλειδί είναι οποιοδήποτε υπερσύνολο ενός υπερκλειδιού.

(i) Αφού το  $\{A_1\}$  είναι υποψήφιο κλειδί όλα τα υπερσύνολα τους είναι υπερκλειδιά. Αλλά πόσα είναι αυτά; Είναι όλα τα δυνατά υποσύνολα του  $\{A_2, A_3, \dots, A_n\}$ , δηλαδή τα δυναμοσύνολα αυτού του συνόλου. Το πλήθος του είναι  $2^{n-1}$ .

(ii) Αφού το  $\{A_1\}$  και το  $\{A_2\}$  είναι υποψήφια κλειδιά όλα τα υπερσύνολα τους είναι υπερκλειδιά. Τα υπερσύνολα του  $\{A_1\}$  από το ερώτημα (i) είναι  $2^{n-1}$ .

Προφανώς τόσα είναι και τα υπερσύνολα του  $\{A_2\}$ , αλλά από αυτά έχουμε ήδη μετρήσει αυτά στα οποία ανήκει το  $\{A_1\}$ . Άρα τα διαφορετικά υπερκλειδιά του  $\{A_2\}$  που δεν έχουμε ήδη μετρήσει είναι όσα και τα δυνατά υποσύνολα του  $\{A_3, A_4, \dots, A_n\}$ , δηλαδή ίσα με  $2^{n-2}$ .

Συνολικά έχουμε  $2^{n-1} + 2^{n-2} = 3 \times 2^{n-2}$ .

### Απάντηση Άσκησης Αυτοαξιολόγησης 3.2

(α) Παραβίαση στον περιορισμού ξένου κλειδιού που αναφέρεται στην ACTOR προκαλείται με την εισαγωγή μιας πλειάδας που αναφέρεται σε ηθοποιό που δεν εμφανίζεται στον πίνακα (σχέση) ACTOR, για παράδειγμα (Julia Roberts, Ocean's

Eleven, 2001, Tess Ocean). Αντίστοιχα, παραβίαση στον περιορισμό ξένου κλειδιού που αναφέρεται στη MOVIE προκαλείται με την εισαγωγή μιας πλειάδας που αναφέρεται σε ταινία που δεν εμφανίζεται στον πίνακα (σχέση) MOVIE, για παράδειγμα (George Clooney, Birth After Reading, 2008, Harry Pfarrer).

(β) Παραβίαση στον περιορισμού ξένου κλειδιού της PLAYS που αναφέρεται στην ACTOR προκαλείται με τη διαγραφή μιας πλειάδας από την ACTOR στην οποία αναφέρονται κάποιες πλειάδες της σχέσης PLAYS. Για παράδειγμα, αν διαγράψουμε την πρώτη πλειάδα της σχέσης ACTOR, οι δύο πρώτες πλειάδες της σχέσης PLAYS χάνουν την πλειάδα στην οποία αναφέρονται. Αντίστοιχα, παραβίαση του ξένου κλειδιού της PLAYS που αναφέρεται στη MOVIE προκαλεί για παράδειγμα η διαγραφή της πρώτης πλειάδας από τη MOVIE, αφού τώρα η τέταρτη πλειάδα της PLAYS θα αναφέρεται σε μια πλειάδα της MOVIE που δεν υπάρχει.

### Απάντηση Άσκησης Αυτοαξιολόγησης 3.3

(α) Για να απαντήστε αρχικά σκεφτείτε σε ποιόν πίνακα (σχέση) υπάρχει η σχετική πληροφορία και στη συνέχεια ποια ακολουθία πράξεων μας δίνει το αποτέλεσμα.

(i) Για παράδειγμα, η πληροφορία για τα ονόματα των πιτσών και τα συστατικά τους υπάρχει στον πίνακα PIZZA. Θα επιλέξουμε τις πίτσες από αυτόν τον πίνακα των οποίων το γνώρισμα Ingredient έχει την τιμή “mushroom” και θα προβάλουμε το γνώρισμα Name, άρα:

$\pi_{Name} (\sigma_{Ingredient = \text{“mushroom”}}(PIZZA))$ .

Αντίστοιχα, σκεφτόμαστε για να διατυπώσουμε και τις υπόλοιπες ερωτήσεις.

(ii)  $\pi_{Ingredient} (\sigma_{Name = \text{“Special”}}(PIZZA))$

(iii)  $\pi_{Ingredient} (LIKES)$

(β) Τα αποτελέσματα των παραπάνω εκφράσεων δίνονται στην Εικόνα 3.19.

Name	Ingredient	Ingredient
Vegetarian	ham	mushroom
Special	bacon	ham
	mushroom	olive
		bacon
		pineapple

(α)

(β)

(γ)

**Εικόνα 3.19** Τα αποτελέσματα των εκφράσεων της Άσκησης Αυτοαξιολόγησης 3.3: (α) ερώτημα (i), (β) ερώτημα (ii), (γ) ερώτημα (iii)

### Απάντηση Άσκησης Αυτοαξιολόγησης 3.4

(i) Θα βρούμε πρώτα τις πίτσες που περιέχουν μανιτάρι και θα τις αφαιρέσουμε από όλες τις πίτσες, δηλαδή:  $\pi_{\text{Name}}(\text{PIZZA}) - \pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"mushroom"}}(\text{PIZZA}))$ .

Ας δούμε δύο συχνές λάθος απαντήσεις.

Η πρώτη είναι η έκφραση  $\pi_{\text{Name}}(\sigma_{\text{Ingredient} \neq \text{"mushroom"}}(\text{PIZZA}))$ . Η έκφραση αυτή μας δίνει ως αποτέλεσμα της πίτσες που έχουν και κάποιο συστατικό διαφορετικό του μανιταριού, αλλά δεν αποκλείει ότι αυτές έχουν και μανιτάρι. Για παράδειγμα, αν την εφαρμόσουμε στο στιγμιότυπο της Εικόνας 3.12, θα μας δώσει στο αποτέλεσμα και τη Vegetarian, αφού η συνθήκη της επιλογής  $\text{Ingredient} \neq \text{"mushroom"}$  ικανοποιείται για παράδειγμα από την πρώτη πλειάδα.

Ένα άλλο κοινό λάθος, είναι η έκφραση  $\pi_{\text{Name}}(\text{PIZZA}) - \sigma_{\text{Ingredient} = \text{"mushroom"}}(\text{PIZZA})$ . Για να δείτε γιατί είναι λάθος, χρησιμοποιήστε πάλι το στιγμιότυπο της Εικόνας 3.12.

(ii) Υπάρχουν δύο βασικοί τρόποι για να εκφράσει κανείς αυτήν την ερώτηση.

Ένας τρόπος είναι με χρήση της πράξης της ένωσης, δηλαδή:  $\pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"mushroom"}}(\text{PIZZA})) \cup \pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"ham"}}(\text{PIZZA}))$ .

Ένας άλλος τρόπος είναι με τη χρήση του OR στην συνθήκη της επιλογής, δηλαδή:  $\pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"mushroom"} \text{ OR } \text{Ingredient} = \text{"ham"}}(\text{PIZZA}))$ .

(iii) Όπως και στο ερώτημα (ii), μπορούμε να χρησιμοποιήσουμε έναν τελεστή συνόλου, συγκεκριμένα την τομή. Η αντίστοιχη έκφραση είναι  $\pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"mushroom"}}(\text{PIZZA})) \cap \pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"ham"}}(\text{PIZZA}))$ .

Προσέξτε όμως ότι δε μπορούμε να χρησιμοποιήσουμε AND στην συνθήκη της επιλογής. Η έκφραση  $\pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"mushroom"} \text{ AND } \text{Ingredient} = \text{"ham"}}(\text{PIZZA}))$  έχει ως αποτέλεσμα το κενό σύνολο σε οποιοδήποτε στιγμιότυπο. Ο λόγος είναι ότι όπως είπαμε, η συνθήκη της επιλογής εφαρμόζεται σε κάθε μια πλειάδα (γραμμή) της σχέσης. Επειδή, δε μπορεί να υπάρξει πλειάδα που να έχει ταυτόχρονα δύο διαφορετικές τιμές σε ένα γνώρισμα της, καμία πλειάδα δεν ικανοποιεί τη συνθήκη.

(iv) Θα χρησιμοποιήσουμε την πράξη της διαφοράς:  $\pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"mushroom"}}(\text{PIZZA})) - \pi_{\text{Name}}(\sigma_{\text{Ingredient} = \text{"ham"}}(\text{PIZZA}))$ .

### Απάντηση Άσκησης Αυτοαξιολόγησης 3.5

Ας πάρουμε πρώτα τα συστατικά που αρέσουν στο Δημήτρη χρησιμοποιώντας επιλογή:

$DIMITRIS-LIKES \leftarrow \sigma_{Student = \text{“Dimitris”}}(LIKES)$ .

Η πληροφορία που χρειαζόμαστε βρίσκετε σε δύο σχέσεις, τη σχέση PIZZA και τη σχέση DIMITRIS-LIKES, τις οποίες και θα συνδυάσουμε με καρτεσιανό γινόμενο. Ας τις μετονομάσουμε πρώτα για να αποφύγουμε ασάφειες με τα κοινά τους γνωρίσματα. Μπορείτε να δώσετε όποιο όνομα θέλετε στα γνωρίσματα, εδώ απλώς προσθέσαμε το όνομα της σχέσης,

$PIZZA1(Name, Pizza-ingredient) \leftarrow PIZZA(Name, Ingredient)$

$DIMITRIS-LIKES1(Student, Likes-ingredient) \leftarrow DIMITRIS-LIKES(Student, Ingredient)$

$Q1 \leftarrow PIZZA1 \times DIMITRIS-LIKES1$

Η Q1 περιέχει όπως είδαμε και «άσχετες» πλειάδες, πρέπει να κρατήσουμε αυτές που αφορούν πίτσες που περιέχουν το συστατικό που αρέσει στο φοιτητή Δημήτρη.

$Q2 \leftarrow \sigma_{Pizza-Ingredient = Likes-ingredient}(Q1)$ .

Τέλος, πρέπει να προβάλουμε τα ονόματα των επιλεγμένων πιτσών:

$Q3 \leftarrow \pi_{Name}(Q2)$ .

Θα μπορούσαμε να κάνουμε τις πράξεις με διαφορετική σειρά, για παράδειγμα να κάνουμε πρώτα το καρτεσιανό γινόμενο των PIZZA και LIKES και μετά να επιλέξουμε τις πλειάδες για τις οποίες ισχύει Student = Dimitris.

Επίσης, εδώ ορίσαμε ενδιάμεσα αποτελέσματα για να γίνει πιο κατανοητή η λύση. Θα μπορούσαμε να έχουμε μόνο μια έκφραση. Επίσης, θα μπορούσαμε να χρησιμοποιήσουμε τον τελεστή μετονομασίας.

Ενδεικτικά δίνουμε μια έκφραση ισοδύναμη της Q3 με χρήση του τελεστή μετονομασίας χωρίς ενδιάμεσα αποτελέσματα. Οι παρενθέσεις υποδεικνύουν την σειρά εκτέλεσης των πράξεων.

$\pi_{Name} (\sigma_{Pizza-Ingredient = Likes-ingredient} ((\rho_{(Name, Pizza-Ingredient)}(PIZZA) \times (\rho_{(Student, LIKES-Ingredient)}(\sigma_{Student = \text{“Dimitris”}}(LIKES))))))$

### Απάντηση Άσκησης Αυτοαξιολόγησης 3.6

(α)

(i) Η πληροφορία υπάρχει σε δύο σχέσεις, την SERVES και την PIZZA, τις οποίες θα πρέπει να συνενώσουμε. Πρώτα θα επιλέξουμε αυτές τις πίτσες που περιέχουν μανιτάρι, θα συνενώσουμε το αποτέλεσμα με τη σχέση SERVES και θα προβάλουμε το όνομα του μαγαζιού, δηλαδή:

$\pi_{Place} ((\sigma_{Ingredient = \text{“mushroom”}}(PIZZA)) \bowtie_{Name = Pizza-name} SERVES)$

Θα μπορούσαμε επίσης να μη συνενώσουμε όλη τη σχέση PIZZA αλλά μόνο το γνώρισμα Names, δηλαδή:

$\pi_{Place} (\pi_{Name} ((\sigma_{Ingredient = \text{“mushroom”}}(PIZZA))) \bowtie_{Name = Pizza-name} SERVES)$

(ii) Για αυτό το ερώτημα θα πρέπει να συνδυάσουμε πληροφορία και από τις τρεις σχέσεις, την SERVES, PIZZA και LIKES, άρα θα χρειαστούμε δύο συνενώσεις. Θα μπορούσαμε να κάνουμε τις συνενώσεις με οποιαδήποτε σειρά. Ας συνενώσουμε πρώτα το PIZZA και LIKES για να βρούμε τις πίτσες που έχουν συστατικά που αρέσουν στον Δημήτρη, όπως στην προηγούμενη άσκηση, μόνο που τώρα θα αντικαταστήσουμε τον τελεστή του καρτεσιανού γινομένου και τον τελεστή της επιλογής στο αποτέλεσμα του με τον τελεστή συνένωσης:

$Q \leftarrow \pi_{Name} (\rho_{(Name, Pizza-Ingredient)}(PIZZA) \bowtie_{Pizza-Ingredient = Likes-ingredient} (\rho_{(Student, LIKES-Ingredient)}(\sigma_{Student = \text{“Dimitris”}}(LIKES)))$

Στη συνέχεια, θα συνενώσουμε το αποτέλεσμα αυτής της συνένωσης με το SERVES και θα προβάλλουμε το όνομα του μαγαζιού

$\pi_{Place} (Q \bowtie_{Name = Pizza-name} SERVES)$

(β)

(i) Η απάντηση στο ερώτημα δεν είναι προφανής. Για να απαντήσετε σε αυτό το ερώτημα, ανατρέξτε στο στιγμιότυπο της Εικόνας 3.12. Οι πίτσες με τουλάχιστον δύο συστατικά εμφανίζονται στη σχέση PIZZA τουλάχιστον δυο φορές, δηλαδή υπάρχουν τουλάχιστον δυο πλειάδες (γραμμές) στη σχέση PIZZA με το ίδιο όνομα πίτσας και διαφορετικό συστατικό. Πως μπορούμε όμως να τις βρούμε εφαρμόζοντας τελεστές της σχεσιακής άλγεβρας;

Η ιδέα είναι να συνενώσουμε την PIZZA με τον εαυτό της και να αναζητήσουμε πλειάδες με την ίδια τιμή στα γνωρίσματα Name και διαφορετική τιμή στα γνωρίσματα Ingredient. Θα το δείξουμε σε στάδια με ενδιάμεσα αποτελέσματα για να γίνει πιο κατανοητό. Πρώτα ας μετονομάσουμε τα γνωρίσματα.

$PIZZA1(Name1, Ingredient1) \leftarrow PIZZA(Name, Ingredient)$

$PIZZA2(Name2, Ingredient2) \leftarrow PIZZA(Name, Ingredient)$

Τώρα μπορούμε να συνενώνουμε τις δύο σχέσεις όπως συζητήσαμε

$AT-LEAST-TWO \leftarrow \pi_{Name1} (PIZZA1 \bowtie_{Name1 = Name2 \text{ AND } Ingredient1 \neq Ingredient2} PIZZA2)$



(ii) Από το ερώτημα (i) έχουμε τις πίτσες με τουλάχιστον δύο συστατικά. Οι πίτσες που έχουν μόνο ένα συστατικό είναι όλες οι άλλες πίτσες. Άρα, αρκεί να αφαιρέσουμε από όλες τις πίτσες αυτές που έχουν τουλάχιστον δύο συστατικά.

$EXACTLY-ONE \leftarrow \pi_{Name} (PIZZA) - AT-LEAST\_TWO$

(iii) Οι πίτσες που έχουν ακριβώς δύο συστατικά είναι αυτές που έχουν τουλάχιστον δύο συστατικά αλλά δεν έχουν τρία συστατικά. Δηλαδή, πρέπει να υπολογίσουμε αυτές που έχουν τρία συστατικά και να τις αφαιρέσουμε από την  $AT-LEAST-TWO$ . Για να τις βρούμε, θα σκεφτούμε όπως στο ερώτημα (i). Οι πίτσες με τουλάχιστον τρία συστατικά εμφανίζονται στη σχέση  $PIZZA$  τουλάχιστον τρεις φορές, δηλαδή υπάρχουν τουλάχιστον τρεις πλειάδες στη σχέση  $PIZZA$  με το ίδιο όνομα πίτσας και διαφορετικό συστατικό. Για να τις βρούμε πρέπει να κάνουμε δύο συνενώσεις της  $PIZZA$  με τον εαυτό της. Ας το δούμε αναλυτικά:

$PIZZA3(Name3, Ingredient3) \leftarrow PIZZA(Name, Ingredient)$

$AT-LEAST-THREE \leftarrow \pi_{Name1} ((PIZZA1 \bowtie Name1 = Name2 \text{ AND } Ingredient1 \neq Ingredient2$   
 $PIZZA2) \bowtie Name1 = Name3 \text{ AND } Ingredient3 \neq Ingredient1 \text{ AND } Ingredient3 \neq Ingredient2 (PIZZA3))$

Και τελικά

$EXACTLY-TWO \leftarrow AT-LEAST-TWO - AT-LEAST-THREE$

### Απάντηση Άσκησης Αυτοαξιολόγησης 3.7

Υπάρχουν διάφορες τέτοιες ερωτήσεις. Για παράδειγμα, βρες (α) Τα μαγαζιά που σερβίρουν όλες τις πίτσες ή όλες τις πίτσες που έχουν πχ μανιτάρι, κλπ (β) Τις πίτσες που σερβίρονται από όλα τα μαγαζιά, κλπ (γ) Τις πίτσες που έχουν όλα τα συστατικά που αρέσουν στους φοιτητές ή όλα τα συστατικά που αρέσουν στο φοιτητή Δημήτρη, κοκ.

Ας διατυπώσουμε χρησιμοποιώντας διαίρεση την ερώτηση που μας δίνει τα μαγαζιά που σερβίρουν όλες τις πίτσες που έχουν μανιτάρι. Θα βρούμε πρώτα τις πίτσες με μανιτάρι και θα διαιρέσουμε τη  $SERVES$  με αυτές. Προσέξτε ότι τα γνωρίσματα πρέπει να έχουν το ίδιο όνομα, για αυτό θα μετονομάσουμε το γνώρισμα  $Name$  της  $PIZZA$  σε  $Pizza-name$  ώστε να έχει το ίδιο όνομα με το αντίστοιχο γνώρισμα της  $SERVES$ . Επίσης, πριν τη διαίρεση, θα πρέπει να κρατήσουμε από την  $SERVES$  μόνο τα γνωρίσματα  $Place$  και  $Pizza-Name$ .

$PIZZA-WITH-MUSHROOM(Pizza-name) \leftarrow \pi_{Name} (\sigma_{Ingredient='Mushroom'}(PIZZA))$

$\pi_{Place, Pizza-name}(SERVES) \div PIZZA-WITH-MUSHROOM$

Ας δούμε τώρα και τον εναλλακτικό τρόπο, τα μαγαζιά που δεν ανήκουν στο αποτέλεσμα είναι τα

$$Q' \leftarrow (\pi_{\text{Place}}(\text{SERVES}) \times \text{PIZZA-WITH-MUSHROOM}) - \pi_{\text{Place, Pizza-name}}(\text{SERVES})$$

Και τελικά

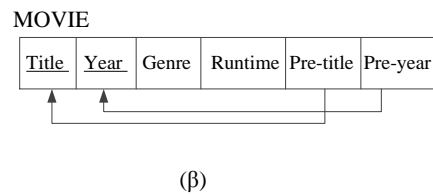
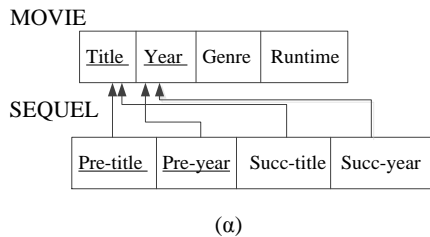
$$\pi_{\text{Place}}(\text{SERVES}) - \pi_{\text{Place}}(Q')$$

## **ΑΠΑΝΤΗΣΕΙΣ ΔΡΑΣΤΗΡΙΟΤΗΤΩΝ**

### **Απάντηση Δραστηριότητας 3.1**

(α) Το σχεσιακό σχήμα στην Εικόνα 3.20(α) ακολουθεί τη γενική μεθοδολογία μετατροπής μιας συσχέτισης στο σχεσιακό μοντέλο. Συγκεκριμένα, ορίσαμε για τη συσχέτιση SEQUEL μια σχέση με γνωρίσματα τα πρωτεύοντα κλειδιά των τύπων οντοτήτων που συσχετίζει. Επειδή η SEQUEL είναι αναδρομική συσχέτιση που συσχετίζει τη MOVIE (με ρόλο predecessor) με τη MOVIE (με ρόλο successor), το πρωτεύον κλειδί της MOVIE εμφανίζεται δυο φορές, μια για κάθε συμμετοχή του τύπου MOVIE στη συσχέτιση. Για να τα ξεχωρίσουμε, τα ονομάσαμε {Pre-title, Pre-year} και {Succ-title, Succ-year} αντίστοιχα. Επειδή η SEQUEL είναι 1-1 και τα δύο σύνολα είναι υπονήφια κλειδιά. Στην Εικόνα 3.20(α), θεωρήσαμε το {Pre-title, Pre-year} ως πρωτεύον κλειδί.

Επειδή η SEQUEL είναι 1-1, μπορούμε να τη συμπεριλάβουμε σε οποιαδήποτε από τις οντότητες που αυτή συσχετίζει, στην περίπτωση με τη MOVIE ως successor ή ως predecessor. Στην Εικόνα 3.20(β), τη συμπεριλάβαμε με τη MOVIE ως successor. Με απλά λόγια, με κάθε ταινία, διατηρούμε το πρωτεύον κλειδί {Pre-title, Pre-year} της ταινίας της οποίας είναι συνέχεια. Τα γνωρίσματα αυτά παίρνουν τιμή null για τις ταινίες που δεν αποτελούν συνέχεια κάποιας άλλης.



MOVIE

Title	Year	Genre	Runtime
Gravity	2013	drama	100
Ocean's Eleven	2001	drama	100
Amour	2012	drama	100
Ocean's Eleven	1960	drama	100
Ocean's Twelve	2004	drama	100
Ocean's Thirteen	2007	drama	100

SEQUEL

Pre-title	Pre-year	Succ-title	Succ-year
Ocean's Eleven	2001	Ocean's Twelve	2004
Ocean's Twelve	2004	Ocean's Thirteen	2007

(γ)

MOVIE

Title	Year	Genre	Runtime	Pre-title	Pre-Year
Gravity	2013	drama	100	null	null
Ocean's Eleven	2001	drama	100	null	null
Amour	2012	drama	100	null	null
Ocean's Eleven	1960	drama	100	null	null
Ocean's Twelve	2004	drama	100	Ocean's Eleven	2001
Ocean's Thirteen	2007	drama	100	Ocean's Twelve	2004

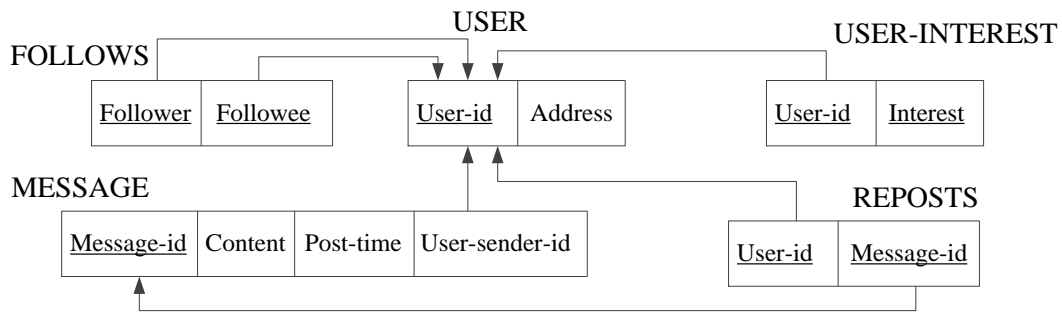
(δ)

**Εικόνα 3.20** Σχεσιακό σχήμα για τη βάση δεδομένων της Δραστηριότητας 3.1: (α) με σχήμα σχέσης για τη συσχέτιση SEQUEL, και (β) χωρίς σχήμα σχέσης για τη SEQUEL, και τα αντίστοιχα στιγμιότυπα: (γ) για το σχήμα (α) και (δ) για το σχήμα (β)

(β) Στην Εικόνα 3.20(γ) βλέπετε ένα στιγμιότυπο για το σχεσιακό σχήμα της Εικόνας 3.20(α) και στην Εικόνα 3.20(δ) ένα στιγμιότυπο για το σχεσιακό σχήμα της Εικόνας 3.20(β). Ο δεύτερος σχεδιασμός αποφεύγει τον ορισμό μιας ειδικής σχέσης για τη SEQUEL αλλά οδηγεί στην εισαγωγή τιμών null.

### Απάντηση Δραστηριότητας 3.2

Ένα πιθανό σχεσιακό σχήμα φαίνεται στην Εικόνα 3.21. Ας το δούμε αναλυτικά ξεκινώντας από τις οντότητες. Το διάγραμμα ΟΣ της Εικόνας 2.11 περιλαμβάνει δύο τύπους οντοτήτων: USER και MESSAGE. Στο σχεσιακό σχήμα έχουμε ορίσει αντίστοιχες σχέσεις. Επειδή το γνώρισμα Interest του USER είναι πλειότιμο, ορίσαμε τη σχέση USER-INTEREST.



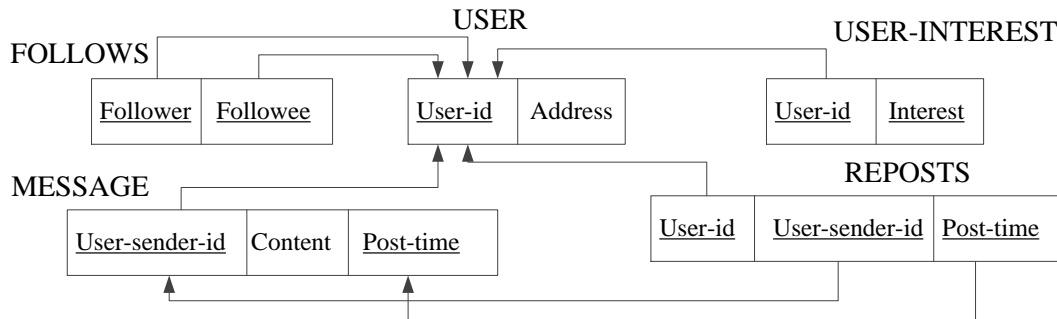
**Εικόνα 3.21** Σχεσιακό σχήμα για τη Δραστηριότητα 3.2

Σχετικά με τις συσχετίσεις, στο διάγραμμα ΟΣ, υπάρχουν οι συσχετίσεις FOLLOWS, POSTS και REPOSTS. Η FOLLOWS είναι μια N-M αναδρομική συσχέτιση και έχει μοντελοποιηθεί από τη σχέση FOLLOWS. Η σχέση REPOSTS είναι επίσης N-M και μοντελοποιείται από τη REPOSTS. Προσέξτε ότι και στις δύο περιπτώσεις, πρωτεύοντα κλειδιά των σχέσεων είναι το σύνολο που περιλαμβάνει τα πρωτεύοντα κλειδιά και των δύο τύπων οντοτήτων που συσχετίζονται. Η συσχέτιση POSTS είναι 1-N και υπάρχουν δύο τρόποι να μοντελοποιηθεί όπως περιγράφει η Παράγραφος 3.2.1. Επιλέξαμε να την ενσωματώσουμε στη σχέση MESSAGE επεκτείνοντας με το γνώρισμα User-sender-id που καταγράφει το User-id του χρήστη που έστειλε το μήνυμα. Ο λόγος είναι ότι η συμμετοχή του τύπου οντοτήτων USER στη συσχέτιση είναι ολική, δηλαδή όλα τα μηνύματα έχουν αποστολέα. Συνεπώς το γνώρισμα User-sender-id θα έχει πάντα μια τιμή διαφορετική της null. Εναλλακτικά, θα μπορούσαμε να μην έχουμε το γνώρισμα User-sender-id στο MESSAGE και να έχουμε έναν τύπο σχέσης POST(User-id, Message-id) με πρωτεύον κλειδί όπως φαίνεται το Message-id, αλλά αυτός ο σχεδιασμός απαιτεί περισσότερο χώρο.

### Απάντηση Δραστηριότητας 3.3

Η βασική διαφορά από την προηγούμενη δραστηριότητα είναι ότι το MESSAGE είναι ένας ασθενής τύπος οντότητας που προσδιορίζεται από το USER μέσω της συσχέτισης POST. Στην Εικόνα 3.22, το MESSAGE μοντελοποιείται από ένα σχήμα σχέσης που αποτελείται από τα γνωρίσματα του MESSAGE και το πρωτεύον κλειδί της προσδιορίζουσας συσχέτισης που το είπαμε User-sender-id (αντί απλώς User-id) για ευκολία στην κατανόηση. Πρωτεύον κλειδί της MESSAGE είναι το πρωτεύον κλειδί της προσδιορίζουσας οντότητας και το μερικό κλειδί της, δηλαδή το σύνολο

{User-sender-id, Post-time}. Τέλος, παρατηρείστε ότι η συσχέτιση REPOSTS στην οποία συμμετέχει το MESSAGE έχει τροποποιηθεί ώστε να περιέχει το νέο πρωτεύον κλειδί της MESSAGE.



**Εικόνα 3.22** Σχεσιακό σχήμα για τη Δραστηριότητα 3.3

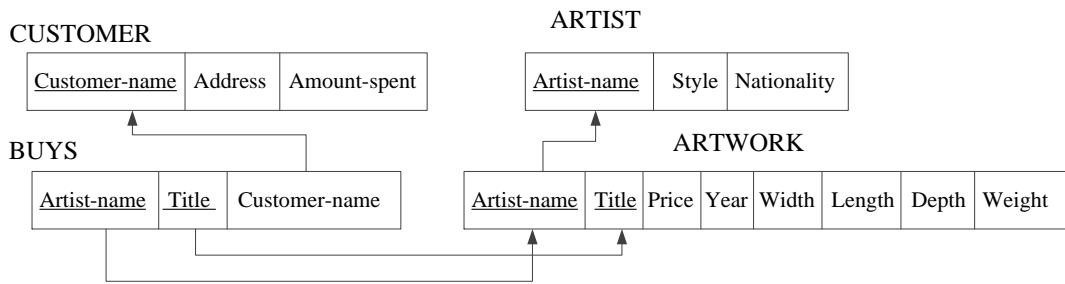
#### Απάντηση Δραστηριότητας 3.4

Ένα πιθανό σχεσιακό σχήμα φαίνεται στην Εικόνα 3.23. Στο διάγραμμα ΟΣ της Εικόνας 2.13 υπάρχουν δύο ισχυροί τύποι οντοτήτων: CUSTOMER και ARTIST, για τους οποίους έχουμε ορίσει αντίστοιχα σχήματα σχέσεων. Επιλέξαμε να αποθηκεύσουμε το παραγόμενο γνώρισμα Amount-spent.

Ο ασθενής τύπος οντότητας ARTWORK που προσδιορίζεται από τον ARTIST μέσω της συσχέτισης CREATES μοντελοποιείται από το σχήμα σχέσης ARTWORK που έχει ως γνωρίσματα τα γνωρίσματα του ARTWORK και το πρωτεύον κλειδί του ARTIST ως ξένο κλειδί.

Για τη συσχέτιση κλάσης/υποκλάσης επιλέξαμε να έχουμε μόνο ένα σχήμα σχέσης για την υπερκλάση, επειδή καμία από τις υποκλάσεις δεν συμμετέχει σε κάποια συσχέτιση. Τα γνωρίσματα Depth και Weight θα παίρνουν τιμές null για τις οντότητες τύπου PAINTING.

Τέλος, για την 1-N συσχέτιση BUYS δημιουργήσαμε ένα σχήμα σχέσης BUYS. Εναλλακτικά, αντί για τη σχέση BUYS, θα μπορούσαμε να προσθέσουμε το Customer-name ως γνώρισμα στο ARTWORK. Επειδή, η συμμετοχή του ARTWORK δεν είναι ολική, αυτό θα οδηγούσε σε null τιμές για τους πίνακες που δεν έχουν αγοραστεί.



**Εικόνα 3.23** Σχεσιακό σχήμα για τη Δραστηριότητα 3.4

### Κατάλογος Εικόνων

- Εικόνα 3. 1 Παράδειγμα σχήματος μιας σχεσιακής βάσης δεδομένων για κινηματογραφικές ταινίες
- Εικόνα 3. 2 Ένα στιγμιότυπο του σχήματος της σχεσιακής βάσης δεδομένων για τις κινηματογραφικές ταινίες
- Εικόνα 3. 3 Παράδειγμα μετατροπής διαγράμματος ΟΣ σε σχεσιακό μοντέλο: (α) σχήμα μιας βάσης δεδομένων σε διάγραμμα ΟΣ, (β) το αντίστοιχο σχεσιακό σχήμα της βάσης δεδομένων
- Εικόνα 3. 4 Εναλλακτική μοντελοποίηση (α) της ADVISES, (β) και (γ) της HEADS, και (δ) ενοποίηση οντοτήτων της Εικόνας 3.3(β)
- Εικόνα 3. 5 Μοντελοποίηση πλειότιμου γνώρισματος (α) το σχεσιακό σχήμα της Εικόνας 3.1 τροποποιημένο ώστε να υποστηρίζει το πλειότιμο γνώρισμα Genre και (β) στιγμιότυπα των σχέσεων MOVIE και MOVIE-GENRE
- Εικόνα 3. 6 Το σχεσιακό σχήμα για τη βάση δεδομένων για τις τηλεοπτικές σειρές που το διάγραμμα ΟΣ της δίνεται στην Εικόνα 3.7
- Εικόνα 3. 7 Σχεσιακό μοντέλο που ακολουθεί το γενικό τρόπο σχεδιασμού για τη βάση δεδομένων για τους εργαζόμενους που περιγράφεται από το διάγραμμα ΕΟΣ της Εικόνας 2.8
- Εικόνα 3. 8 Σχεσιακό μοντέλο χωρίς σχήμα σχέσης για την υπερκλάση για τη βάση δεδομένων για τους εργαζόμενους που περιγράφεται από το διάγραμμα ΕΟΣ της Εικόνας 2.8
- Εικόνα 3. 9 Σχεσιακό μοντέλο με χρήση μόνο μιας σχέσης για τη συσχέτιση κλάσης/υποκλάσης υπερκλάση για τη βάση δεδομένων για τους εργαζόμενους που περιγράφεται από το διάγραμμα ΕΟΣ της Εικόνας 2.8
- Εικόνα 3. 10 Παραδείγματα αποτελεσμάτων στο στιγμιότυπο της Εικόνας 3.2 , των εκφράσεων (α)  $\sigma_{Year > 2010}(MOVIE)$ , (β)  $\sigma_{Sex = "male" \text{ AND } Year > 1960}(ACTOR)$ , (γ)  $\sigma_{Genre = "drama" \text{ OR } Genre = "crime"}(MOVIE)$
- Εικόνα 3. 11 Παραδείγματα αποτελεσμάτων στο στιγμιότυπο της Εικόνας 3.2 , των εκφράσεων (α)  $\pi_{Genre}(MOVIE)$ , (β)  $\pi_{Name, Sex, Nationality}(ACTOR)$ , (γ)  $\pi_{Genre}(\sigma_{Year > 2000}(MOVIE))$

- Εικόνα 3. 12 Ένα στιγμιότυπο του σχήματος της σχεσιακής βάσης δεδομένων για τις πίτσες
- Εικόνα 3. 13 Παραδείγματα αποτελεσμάτων πράξεων συνόλου στο στιγμιότυπο της Εικόνας 3.2: (α)  $R$ , (β)  $S$ , (γ)  $R \cup S$ , (δ)  $R \cap S$  και (ε)  $R - S$ .
- Εικόνα 3. 14 Παράδειγμα καρτεσιανού γινομένου
- Εικόνα 3. 15 Παράδειγμα καρτεσιανού γινομένου στο στιγμιότυπο της Εικόνας 3.2
- Εικόνα 3. 16 Παράδειγμα συνενώσεων των σχέσεων  $R$  και  $S$  της Εικόνας 3.14: (α)  $\rho_{(A, R-B)}(R) \bowtie_{A > C} \rho_{(S-B, C)}(S)$ , (β)  $\rho_{(A, R-B)}(R) \bowtie_{R-B = C-B} \rho_{(S-B, C)}(S)$ , (γ)  $R * S$
- Εικόνα 3. 17 Παράδειγμα εξωτερικής συνένωσης των σχέσεων  $R$  και  $S$  της Εικόνας 3.14: (α)  $\rho_{(A, R-B)}(R) \Join_{R-B = C-B} \rho_{(S-B, C)}(S)$ , (β)  $\rho_{(A, R-B)}(R) \Join_{R-B = C-B} \rho_{(S-B, C)}(S)$ , (γ)  $\rho_{(A, R-B)}(R) \Join_{R-B = C-B} \rho_{(S-B, C)}(S)$ .
- Εικόνα 3. 18 Παραδείγματα διαίρεσης
- Εικόνα 3. 19 Τα αποτελέσματα των εκφράσεων της Άσκησης Αυτοαξιολόγησης 3.3: (α) ερώτημα (i), (β) ερώτημα (ii), (γ) ερώτημα (iii)
- Εικόνα 3. 20 Σχεσιακό σχήμα για τη βάση δεδομένων της Δραστηριότητας 3.1: (α) με σχήμα σχέσης για τη συσχέτιση SEQUEL, και (β) χωρίς σχήμα σχέσης για τη SEQUEL, και τα αντίστοιχα στιγμιότυπα: (γ) για το σχήμα (α) και (δ) για το σχήμα (β)
- Εικόνα 3. 21 Σχεσιακό σχήμα για τη Δραστηριότητα 3.2
- Εικόνα 3. 22 Σχεσιακό σχήμα για τη Δραστηριότητα 3.3
- Εικόνα 3. 23 Σχεσιακό σχήμα για τη Δραστηριότητα 3.4

### ***Κατάλογος Πινάκων***

Πίνακας 3. 1 Στιγμιότυπο σχέσης για την Άσκηση Αυτοαξιολόγησης 3.1.