Σχεσιακή άλγεβρα

Αθανάσιος Σταυρακούδης
http://stavrakoudis.econ.uoi.gr
astavrak@uoi.gr

Φεβρουάριος 2013

Περιεχόμενα

1 Σχεσιακή άλγεβρα και βασικές σχεσιακές πράξεις

Στόχος του κεφαλαίου είναι:

- Να εκτελείτε τις μοναδιαίες πράξεις προβολής και επιλογής.
- Να εκτελείτε τις συνολοθεωρητικές πράξεις ένωσης, διαφοράς και τομής.
- Να εκτελείτε πράξεις καρτεσιανού γινομένου συνόλων.
- Να εκτελείτε πράξεις σύζευξης σχέσεων.
- Να καταλάβετε τις διαφορές και ομοιότητες μεταξύ των διαφορετικών τύπων σύζευξης: φυσική, θήτα σύζευξη, εξωτερική, δεξιά εξωτερική, αριστερή εξωτερική, κ.α.
- Να εκτελείτε τη σχεσιακή πράξη της διαίρεσης.
- Να εκτελείτε πράξεις σύνοψης και συνάθροισης.
- Να εκτελείτε πράξεις ενημέρωσης σχέσεων: εισαγωγή, διαγραφή και τροποποίηση δεδομένων.

1.1 Κλειστότητα

Η σχεσιακή άλγεβρα και ο σχεσιακός λογισμός παρέχουν ένα σύνολο από τελεστές για πράξεις ανάμεσα σε σχέσεις. Όλες οι πράξεις στηρίζονται στην ιδιότητα της **κλειστότητας**, δηλαδή στο γεγονός πως το αποτέλεσμα οποιασδήποτε σχεσιακής πράξης είναι μια σχέση. Η ιδιότητα της κλειστότητας επιτρέπει τη λεγόμενη εμφώλευση (ένθεση) παραστάσεων: το αποτέλεσμα μιας σχεσιακής πράξης είναι σχέση, οπότε αυτό το αποτέλεσμα μπορεί να συμμετέχει ως τελεστέος σε μια νέα σχεσιακή παράσταση, σε μια νέα σχεσιακή πράξη δηλαδή.

Θυμηθείτε πως κάτι ανάλογο συμβαίνει στην άλγεβρα αριθμών: το σύνολο των αριθμών είναι κλειστό, δηλαδή πράξεις με αριθμούς, έχουν αποτέλεσμα αριθμούς. Λόγω αυτής της ιδιότητας μπορούν να γραφούν ένθετες παραστάσεις, όπως:

$$4 + (3 \cdot 2)$$

Το ίδιο μπορεί να γίνει και στην άλγεβρα σχέσεων: το αποτέλεσμα μιας σχεσιακής πράξης (σε αντιστοιχία με το $3 \cdot 2 = 6$) μπορεί να είναι τελεστέος σε μια άλλη σχεσιακή πράξη (σε αντιστοιχία με $4 + 6$).

Οι πράξεις με σχέσεις παράγουν σχέσεις. Το αποτέλεσμα της πράξης, αφού είναι σχέση, έχει μια επικεφαλίδα (σχήμα της σχέσης), επομένως πρέπει να μπορεί να οριστεί ο βαθμός της, δηλαδή το πλήθος των γνωρισμάτων της σχέσης αποτέλεσμα. Επίσης, έχει κορμό, δηλαδή ένα σύνολο από πλειάδες (συστοιχίες), επομένως πρέπει να έχει πληθικότητα. Θα πρέπει κανείς, για κάθε σχεσιακή πράξη, να μπορεί να ορίσει το σχήμα της σχέσης του αποτελέσματος, το βαθμό και την πληθικότητά του.

1.2 Συμβατότητα τύπου

Θα ορίσουμε σε αυτό το σημείο την έννοια της συμβατότητας τύπου, η οποία έχει εφαρμογή σε πολλές σχεσιακές πράξεις.

Δύο σχέσεις r και s, έχουν **συμβατότητα τύπου**, αν και μόνο αν:

- Έχουν τον ίδιο βαθμό, δηλαδή έχουν το ίδιο πλήθος γνωρισμάτων
- Τα αντίστοιχα γνωρίσματα έχουν το ίδιο πεδίο ορισμού

Προσοχή, δεν ενδιαφέρουν τα ονόματα των γνωρισμάτων, παρά μόνο το πλήθος τους και οι τύποι δεδομένων.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	20
2	b	10
3	c	20

s

A	B
5	b
2	b
3	c

t

A	B	C
5	b	a
2	b	b
3	c	b

u

Σχήμα 1.1: Παράδειγμα συμβατότητας τύπου οι σχέσεις r και s έχουν συμβατότητα τύπου. Όλα τα υπόλοιπα ζεύγη σχέσεων δεν έχουν.

Δείτε για παράδειγμα το σχήμα 1.1. Οι σχέσεις r και s έχουν συμβατότητα τύπου γιατί πληρούν και τις δύο προϋποθέσεις που έχουν αναφερθεί. Αντίθετα οι σχέσεις r και t δεν έχουν συμβατότητα τύπου επειδή δεν έχουν το ίδιο πλήθος γνωρισμάτων. Επίσης, ούτε οι σχέσεις r και u έχουν συμβατότητα τύπου. έχουν βέβαια το ίδιο πλήθος γνωρισμάτων, αλλά οι τύποι δεδομένων διαφέρουν. Η γνώρισμα C της σχέσης r περιέχει ακεραίους, ενώ το γνώρισμα C της σχέσης u περιέχει κείμενο. Επομένως δεν υπάρχει συμβατότητα τύπου.

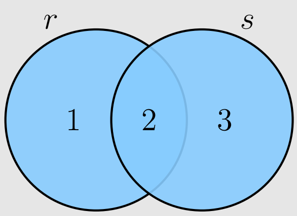
1.3 Συνολοθεωρητικές πράξεις

1.3.1 Ένωση

Ένωση δύο σχέσεων $r(R)$ και $s(S)$, που έχουν συμβατότητα τύπου, είναι μια νέα σχέση που έχει σχήμα (επικεφαλίδα) ίδιο με αυτό της r και s, και κορμό το σύνολο των κορμών των r και s, δηλαδή όλες τις πλειάδες που ανήκουν στην r, ή στην s, η και στις δύο πλειάδες. Η ένωση συμβολίζεται με $r \cup s$ ή $r \quad UNION \quad s$.

Ορισμός της ένωσης:

$$r \cup s = \{t \,|\, t \in r \text{ or } t \in s\} \tag{1.1}$$



$$r = \{1, 2\}$$
$$s = \{2, 3\}$$
$$r \cup s = \{1, 2, 3\}$$

Σχήμα 1.2: Η ένωση δύο συνόλων έχει ως μέλη τα στοιχεία είτε του ενός είτε του άλλου συνόλου.

Δείτε για παράδειγμα το σχήμα 1.2, το οποίο αναπαριστά σχηματικά την ένωση $r \cup s$. Αν $r = \{1, 2\}$ και $s = \{2, 3\}$ τότε $t = r \cup s = \{1, 2, 3\}$. Δηλαδή η ένωση συμπεριλαμβάνει όλα τα στοιχεία των δύο συνόλων. Προσέξτε πως το 2 είναι μέλος και των δύο συνόλων. Ωστόσο, στο αποτέλεσμα υπάρχει μόνο φορά.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	30
2	b	10
3	c	20

s

A	B	C
1	b	10
5	a	30
3	c	20
2	b	10

$r \cup s$

Σχήμα 1.3: Παράδειγμα σχεσιακής ένωσης: $t = r \cup s$

Ένα ακόμη παράδειγμα σχεσιακής ένωσης δίνεται στο σχήμα 1.3. Προσέξτε πως οι πλειάδες <5,a,30> και <3,c,20> βρίσκονται και στις δύο σχέσεις. Ωστόσο, στο αποτέλεσμα της σχεσιακής ένωσης θα υπάρχουν μόνο μία φορά η καθεμία από αυτές.

Για τη σχεσιακή πράξη της ένωσης ισχύει η αντιμεταθετική ιδιότητα:

$$r \cup s = s \cup r \tag{1.2}$$

Δηλαδή δεν έχει σημασία η σειρά των τελεστέων, με οποιαδήποτε σειρά και να γίνει η πράξη της τομής το αποτέλεσμα είναι το ίδιο.

Δείτε ένα παράδειγμα της αντιμεταθετικής ιδιότητας στο σχήμα 1.4. Υπενθυμίζετε πως η διάταξη των πλειάδων σε μία σχέση δεν έχει σημασία.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	30
2	b	10
3	c	20

s

A	B	C
1	b	10
5	a	30
3	c	20
2	b	10

$r \cup s$

A	B	C
5	a	30
2	b	10
3	c	20
1	b	10

$s \cup r$

Σχήμα 1.4: Παράδειγμα αντιμεταθετικής ιδιότητας της σχεσιακής ένωσης: $r \cup s = s \cup r$.

Για την σχεσιακή πράξη της ένωσης ισχύει επίσης και η προσεταιριστική ιδιότητα:

$$r \cup (s \cup t) = (r \cup s) \cup t \tag{1.3}$$

Λόγω αυτής της της ιδιότητας, είναι δυνατό να γραφεί η παράσταση χωρίς παρενθέσεις:

$$r \cup s \cup t$$

για να δηλώσει την ένωση τριών ή περισσότερων σχέσεων.

Η SQL υποστηρίζει την ένωση με τη χρήση της φράσης UNION. Έτσι, τα αποτελέσματα από δύο ερωτήματα SELECT, μπορούν να ενωθούν με τη φράση UNION. Για παράδειγμα, η εντολή SQL που αντιστοιχεί στο παράδειγμα του σχήματος 1.3 είναι:

```
SELECT r.A, r.B, r.C
  FROM r
 UNION DISTINCT
SELECT s.A, s.B, s.C
  FROM s
```

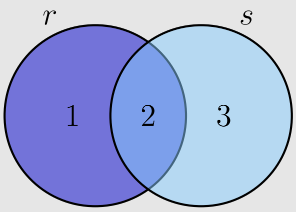
Η επιλογή DISTINCT (που στην SQL είναι προαιρετική) χρειάζεται εδώ για την απαλοιφή των διπλότυπων εγγραφών. Όπως είναι γνωστό τα αποτελέσματα ερωτημάτων SQL επιτρέπουν την εμφάνιση διπλότυπων εγγραφών, ενώ κάτι τέτοιο στην σχεσιακή άλγεβρα δεν ισχύει. Ωστόσο, όπως η σχεσιακή άλγεβρα απαιτεί τη συμβατότητα τύπου για την ένωση, έτσι και η SQL απαιτεί την ομοιογένεια των δύο αποτελεσμάτων για την ορθή εκτέλεση του ερωτήματος.

1.3.2 Διαφορά

Διαφορά δύο σχέσεων r και s, που έχουν συμβατότητα τύπου, είναι μια νέα σχέση που έχει σχήμα (επικεφαλίδα) ίδιο με αυτό της r και s, και κορμό τις πλειάδες που ανήκουν στην r αλλά όχι στην s. Η διαφορά συμβολίζεται με $R - S$ ή $r \quad MINUS \quad s$.

Ορισμός της διαφοράς:

$$r - s = \{t \,|\, t \in r \text{ and } t \notin s\} \tag{1.4}$$



$$
\begin{aligned}
r &= \{1,2\}\\
s &= \{2,3\}\\
r-s &= \{1\}\\
s-r &= \{3\}
\end{aligned}
$$

Σχήμα 1.5: Η διαφορά δύο συνόλων έχει ως μέλη τα στοιχεία του ενός συνόλου που δεν ανήκουν στο άλλο σύνολο.

Δείτε το σχήμα 1.5 που απεικονίζει ένα παράδειγμα διαφοράς δύο συνόλων. Προσέξτε πως $r - s = \{1\}$, δηλαδή το αποτέλεσμα της διαφοράς $r - s$ είναι εκείνα τα μέλη του συνόλου r που δεν ανήκουν στο σύνολο r. Από την άλλη πλευρά, προσέξτε πως $s-r = \{3\}$, δηλαδή το αποτέλεσμα της διαφοράς $s - r$ είναι εκείνα τα μέλη του συνόλου s που δεν ανήκουν στο σύνολο r.

Στη σχεσιακή πράξη της διαφοράς δεν ισχύει η αντιμεταθετική ιδιότητα:

$$r - s \neq s - r \tag{1.5}$$

Δηλαδή η σειρά των τελεστέων έχει σημασία. Ένα τέτοιο παράδειγμα φαίνεται στο σχήμα 1.6.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	30
2	b	10
3	c	20

s

A	B	C
1	b	10

$r - s$

A	B	C
2	b	10

$s - r$

Σχήμα 1.6: Παράδειγμα μη ισχύος της αντιμεταθετικής ιδιότητας της σχεσιακής πράξης της διαφοράς: $t_1 = r - s$ και $t_2 = s - r$. Προσέξτε πως $t1 \neq t2$.

Επίσης δεν ισχύει η προσεταιριστική ιδιότητα:

$$r - (s - t) \neq (s - r) - t \tag{1.6}$$

1.3.3 Τομή

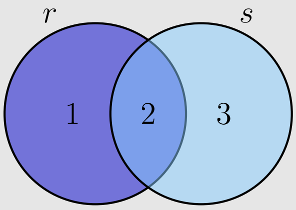
Τομή δύο σχέσεων $r(R)$ και $s(S)$, που έχουν συμβατότητα τύπου, είναι μια νέα σχέση που έχει σχήμα (επικεφαλίδα) ίδιο με αυτό της r και s, και κορμό τις πλειάδες που ανήκουν στην r και στην s, δηλαδή τις κοινές πλειάδες. Η τομή συμβολίζεται με $r \cap s$ ή $r \quad INTERSECT \quad s$.

Ορισμός της τομής:

$$r \cap s = \{t \,|\, t \in r \text{ and } t \in s\} \tag{1.7}$$

Δείτε για παράδειγμα το σχήμα 1.7, το οποίο αναπαριστά σχηματικά την τομή $r \cap s$. Αν $r = \{1,2\}$ και $s = \{2,3\}$ τότε $t = r \cap s = \{2\}$. Δηλαδή η τομή περιλαμβάνει μόνο τα κοινά στοιχεία των δύο συνόλων.

Ένα ακόμη παράδειγμα σχεσιακής ένωσης δίνεται στο σχήμα 1.8. Μόνο οι κοινές πλειάδες <5,a,30> και <3,c,20> των δύο σχέσεων υπεισέρχονται στο αποτέλεσμα. Η πλειάδα <1,b,10> της σχέσης



$$r = \{1, 2\}$$
$$s = \{2, 3\}$$
$$r \cap s = \{2\}$$

Σχήμα 1.7: Η τομή δύο συνόλων έχει ως μέλη τα κοινά στοιχεία των δύο συνόλων.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	30
2	b	10
3	c	20

s

A	B	C
5	a	30
3	c	20

$r \cap s$

Σχήμα 1.8: Παράδειγμα σχεσιακής τομής: $t = r \cap s$

r και η πλειάδα <2,b,10> της σχέσης s αποκλείονται από το αποτέλεσμα. Προσέξτε ότι στο αποτέλεσμα της σχεσιακής πράξης της τομής κάθε πλειάδα υπάρχει μία μόνο φορά.

Για τη σχεσιακή πράξη της τομής ισχύει η αντιμεταθετική ιδιότητα:

$$r \cap s = s \cap r \tag{1.8}$$

Δηλαδή δεν έχει σημασία η σειρά των τελεστέων, με οποιαδήποτε σειρά και να γίνει η πράξη της τομής το αποτέλεσμα είναι το ίδιο.

Ένα τέτοιο παράδειγμα φαίνεται στο σχήμα 1.9. Με όποια σειρά και αν τοποθετηθούν οι τελεστέοι της σχεσιακής πράξης, το αποτέλεσμα της τομής είναι το ίδιο.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	30
2	b	10
3	c	20

s

A	B	C
5	a	30
3	c	20

$r \cap s$

A	B	C
5	a	30
3	c	20

$s \cap r$

Σχήμα 1.9: Παράδειγμα της αντιμεταθετικής ιδιότητας της σχεσιακής πράξης της τομής: $r \cap s = s \cap r$.

Για τη σχεσιακή πράξη της τομής ισχύει επίσης η προσεταιριστική ιδιότητα:

$$r \cap (s \cap t) = (r \cap s) \cap t \tag{1.9}$$

Όπως και στην ένωση, έτσι και στην τομή, είναι δυνατό να γραφεί η παράσταση χωρίς παρενθέσεις:

$$r \cap s \cap t$$

για να δηλώσει την τομή τριών ή περισσότερων σχέσεων.

Όπως και στην ένωση, έτσι και στην τομή η SQL υποστηρίζει τη φράση *INTERSECT* που αντιστοιχεί στη σχεσιακή πράξη της τομής. Το παράδειγμα του σχήματος 1.8, μπορεί στην SQL να γραφεί ως:

```
    SELECT r.A, r.B, r.C
      FROM r
INTERSECT
    SELECT s.A, s.B, s.C
      FROM s
```

Ισχύει και εδώ η απαίτηση για ομοιογένεια στα αποτελέσματα των ερωτημάτων, σε αντιστοίχιση με την απαίτηση για συμβατότητα τύπου, στις παραστάσεις της σχεσιακής άλγεβρας.

Αξίζει να σημειωθεί πως η τομή είναι μια παράγωγη πράξη. Ισχύει:

Ορισμός της τομής:
$$r \cap s = r - (r - s) \tag{1.10}$$

Δηλαδή το αποτέλεσμα της τομής $r \cap s$ ισούται με το αποτέλεσμα της διαφοράς της r από τη διαφορά $r - s$. Ένα παράδειγμα ισχύος της σχέσης 1.10 φαίνεται στο σχήμα 1.10.

A	B	C
1	b	10
5	a	30
3	c	20

r

A	B	C
5	a	30
2	b	10
3	c	20

s

A	B	C
1	b	10

$r - s$

A	B	C
5	a	30
3	c	20

$r - (r - s)$

A	B	C
5	a	30
3	c	20

$r \cap s$

Σχήμα 1.10: Παράδειγμα τομής ως αποτέλεσμα δύο αφαιρέσεων: $r \cap s = r - (r - s)$.

1.3.4 Γινόμενο

Στη θεωρία συνόλων το καρτεσιανό γινόμενο ορίζεται ως:

Καρτεσιανό γινόμενο συνόλων:
$$A \times B = \{\ (a, b) : a \in A\ ,\ b \in B\} \tag{1.11}$$

Για παράδειγμα, αν $A = \{a, b\}$ και $B = \{1, 2, 3\}$ τότε

$$A \times B = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\}$$

Δηλαδή το γινόμενο δύο συνόλων A και B αποτελείται από όλους του συνδυασμούς των μελών των A και B. Προσέξτε πως το καρτεσιανό γινόμενο περιέχει διατεταγμένα ζεύγη.

Το καρτεσιανό γινόμενο της σχεσιακής άλγεβρας ορίζεται λίγο διαφορετικά. **Καρτεσιανό γινόμενο** δύο σχέσεων $r(R)$ και $s(S)$, είναι μια σχέση που έχει επικεφαλίδα το σύνολο των γνωρισμάτων των σχέσεων $r(R)$ και $s(S)$, και κορμό το σύνολο όλων των συνδυασμών των πλειάδων που ανήκουν στην $r(R)$ και στην $s(S)$. Το καρτεσιανό γινόμενο συμβολίζεται με $r \times s$ ή $r\ \ TIMES\ \ s$. Προσέξτε πως δεν υπάρχει καμία απαίτηση για συμβατότητα τύπου ανάμεσα στις r και s.

Καρτεσιανό γινόμενο σχέσεων:

$$r \times s = \{\, tu \,|\, t \in r \wedge u \in s \} \tag{1.12}$$

Το σχήμα ενός καρτεσιανού γινομένου προκύπτει μετά από μετονομασία των πιθανών κοινών γνωρισμάτων δύο σχέσεων. Για παράδειγμα, αν Y είναι ένα κοινό γνώρισμα των σχέσεων $r(R)$ και $s(S)$, τότε το σχήμα της σχέσης $r \times s$ είναι:

$$T = (R - S) \cup (S - R) \cup \{R.Y, S.Y \,|\, Y \in R \cap S\}$$

Αυτό σημαίνει πως στο γινόμενο τα κοινά γνωρίσματα υπάρχουν δύο φορές. Προσοχή, αυτό επιτρέπεται μόνο μετά από μετονομασία, τα μέλη ενός συνόλου πρέπει να είναι μοναδικά.

Αν η σχέση r είναι m_R βαθμού και έχει n_R πληθικότητα, και η σχέση s είναι m_S βαθμού και έχει n_S πληθικότητα, τότε το αποτέλεσμα είναι $m_R + m_S$ βαθμού και $n_R \cdot n_S$ πληθικότητας. Δηλαδή το αποτέλεσμα έχει βαθμό ίσο με το άθροισμα των βαθμών των σχέσεων r και s, και πληθικότητα το γινόμενο των πληθικοτήτων των σχέσεων r και s.

A	B
1	b
5	a
3	c

r

D	E	F
b	4	30
a	2	10

s

A	B	C	D	E
1	b	b	4	30
1	b	a	2	10
5	a	b	4	30
5	a	a	2	10
3	c	b	4	30
3	c	a	2	10

$r \times s$

Σχήμα 1.11: Παράδειγμα σχεσιακού καρτεσιανού γινομένου.

Το σχήμα 1.11 απεικονίζει ένα παράδειγμα σχεσιακού καρτεσιανού γινομένου, όπου οι δύο σχέσεις δεν έχουν γνώρισμα με κοινό όνομα. Αν υπάρχει γνώρισμα με κοινό όνομα στις δύο σχέσεις, τότε γίνεται μετονομασία του γνωρίσματος. Ένα τέτοιο παράδειγμα φαίνεται στο σχήμα 1.12.

A	B
1	b
5	a
3	c

r

A	B	F
b	4	30
a	2	10

s

R.A	R.B	S.A	S.B	E
1	b	b	4	30
1	b	a	2	10
5	a	b	4	30
5	a	a	2	10
3	c	b	4	30
3	c	a	2	10

$r \times s$

Σχήμα 1.12: Παράδειγμα σχεσιακού καρτεσιανού γινομένου με μετονομασία των κοινών γνωρισμάτων.

Το γινόμενο, από μόνο του, δεν έχει κάποια χρησιμότητα. Δεν προσθέτει καμία νέα πληροφορία στη βάση δεδομένων. Η κατανόησή του όμως, είναι απολύτως απαραίτητη διότι χρησιμοποιείται πολύ συχνά ως ενδιάμεσο αποτέλεσμα πράξεων, όπως είναι η σύζευξη.

Γενικά το γινόμενο ορίζεται για δύο σχέσεις που δεν έχουν κοινό όνομα στα γνωρίσματά τους. Ωστόσο, κάτι τέτοιο, δεν μπορεί να αποκλειστεί, είναι πιθανό να υπάρχουν γνωρίσματα με το ίδιο όνομα. Επειδή το αποτέλεσμα του γινομένου περιλαμβάνει στο σχήμα του (επικεφαλίδα) όλα τα γνωρίσματα των δύο σχέσεων, για να μην υπάρξει *σύγκρουση ονοματοδοσίας*, το γνωρίσματα αυτά δηλώνονται από τη σχετική τους θέση στις αρχικές σχέσεις.

Για το καρτεσιανό γινόμενο δεν ισχύει η αντιμεταθετική ιδιότητα:

$$r \times s \neq s \times r$$

Δηλαδή έχει σημασία η σειρά των τελεστέων.

Ισχύει ωστόσο η προσεταιριστική ιδιότητα:

$$r \times (s \times t) = (r \times s) \times t$$

Όπως και στην ένωση και στην τομή, έτσι και στο γινόμενο, είναι δυνατό να γραφεί η παράσταση χωρίς παρενθέσεις:

$$r \times s \times t$$

για να δηλώσει το γινόμενο τριών ή περισσότερων σχέσεων.

1.3.5 Παραδείγματα συνολοθεωρητικών πράξεων.

Ποδοσφαιριστές

Ας θεωρήσουμε ένα παιχνίδι ποδοσφαίρου, για παράδειγμα ανάμεσα στις ομάδες Milan[1] και Real.[2]

Στο παράδειγμά μας, έχουμε δύο σχέσεις με τους ποδοσφαιριστές των δύο ομάδων:
$Milan(\underline{number}, name)$ και $Real(\underline{number}, name)$.

Οι δύο αυτές σχέσεις, που έχουν συμβατότητα τύπου, περιγράφουν το αριθμό της φανέλας με τον οποίο αγωνίζεται κάθε ποδοσφαιριστής και το όνομά του. Η χρήση τους είναι πολύ απλή: βοηθούν τον αθλητικό δημοσιογράφο που καλύπτει τον αγώνα να διακρίνει κάθε ποδοσφαιριστή από το νούμερο της φανέλας του.

Στο συγκεκριμένο αγώνα, οι πλειάδες των δύο δύο σχέσεων δεν έχουν επικάλυψη: δεν υπάρχει δηλαδή κανείς ποδοσφαιριστής που να εμφανίζεται και στις δύο σχέσεις. Δεν μπορεί ένας ποδοσφαιριστής, στον ίδιο αγώνα, να αγωνιστεί και με τις δύο ομάδες.

Η ένωση των δύο σχέσεων:

$$Milan \cup Real$$

μας δίνει το σύνολο των ποδοσφαιριστών που συμμετέχουν στον ποδοσφαιρικό αγώνα, είτε από την μία ομάδα, είτε από την άλλη. Το αποτέλεσμα είναι το ίδιο με την ένωση:

$$Real \cup Milan$$

Η σειρά εμφάνισης δεν παίζει κάποιο ρόλο. Λόγω της αντιμεταθετικής ιδιότητας, ισχύει:

$$Milan \cup Real = Real \cup Milan$$

Η τομή των δύο σχέσεων:

$$Milan \cap Real$$

[1] http://en.wikipedia.org/wiki/A.C._Milan
[2] http://en.wikipedia.org/wiki/Real_Madrid_C.F.

θα επιστρέψει το κενό σύνολο. Αυτό γίνεται επειδή καμία πλειάδα της σχέσης $Milan$ δεν ταυτίζεται με κάποια πλειάδα της σχέση $Real$.

Το ίδιο θα συμβεί και με την τομή:

$$Real \cap Milan$$

θα αποδώσει επίσης το κενό σύνολο, καμία πλειάδα της $Real$ δεν ταυτίζεται με κάποια από της πλειάδες της $Milan$. Το γεγονός αυτό συνοψίζεται με την ισχύ της αντιμεταθετικής ιδιότητας:

$$Milan \cap Real = Real \cap Milan$$

Αν η τομή δύο σχέσεων είναι το κενό σύνολο, τότε η διαφορά των δύο σχέσεων δεν έχει κάποιο νόημα. Για παράδειγμα:

$$Milan - Real = Milan$$

όπως επίσης και:

$$Real - Milan = Real$$

Δηλαδή, αν από τους ποδοσφαιριστές της $Milan$ αφαιρέσουμε τους ποδοσφαιριστές της $Real$, θα μας μείνουν οι ποδοσφαιριστές της $Milan$. Το ίδιο θα συμβεί και με την ανάποδη σειρά, αν από τους ποδοσφαιριστές της $Real$ αφαιρέσουμε τους ποδοσφαιριστές της $Milan$, θα μας μείνουν οι ποδοσφαιριστές της $Real$.

Στη συγκεκριμένη, ακραία, αυτή περίπτωση, ισχύει:

$$Milan - Real = Milan - Real$$

Το γεγονός αυτό δεν πρέπει να σας μπερδεύει, η ισχύς της παραπάνω σχέσης οφείλεται στη συγκεκριμένη περίπτωση και δεν θα πρέπει να γενικευτεί σε άλλες περιπτώσεις. Στη σχεσιακή πράξη της διαφοράς, γενικά, δεν ισχύει η αντιμεταθετική ιδιότητα.

Διαχρονική καταγραφή των ποδοσφαιριστών των δύο ομάδων.

Ας γενικεύσουμε τώρα το παράδειγμα, ως προς τη χρονική του διάσταση. Ας μην περιοριστούμε στη διάρκεια ενός παιχνιδιού, αλλά ας καταγράψουμε στις δύο σχέσεις τους ποδοσφαιριστές των δύο ομάδων σε μια μεγάλη χρονική διάρκεια, για παράδειγμα κατά την τελευταία δεκαετία.

Σε αυτή την περίπτωση, η πληθικότητα των δύο σχέσεων θα είναι αρκετά μεγαλύτερη: περίπου 20–22 παίκτες ποδοσφαίρου συμμετέχουν από κάθε ομάδα σε ένα παιχνίδι, κάθε ομάδα ωστόσο αλλάζει περίπου 3–8 παίκτες κάθε περίοδο. Σε μια δεκαετία, μια ομάδα ενδέχεται να έχουν συμμετάσχει περίπου 50–100 ποδοσφαιριστές. Ένα μικρό δείγμα δεδομένων παρουσιάζεται στο σχήμα 1.13.

number	name
806	Marek Jankulovski
705	Kaká
580	Alessandro Nesta
322	Andriy Shevchenko

$Milan$

number	name
205	Claude Makélélé
108	Raúl Albiol
705	Kaká
203	Xabi Alonso

$Real$

Σχήμα 1.13: Ποδοσφαιριστές της Milan και της Real. Εδώ δίνεται ένα μικρό μόνο δείγμα δεδομένων. Ο αριθμός (number) είναι ένας αυθαίρετος κωδικός αρίθμησης των ποδοσφαιριστών και δεν έχει σχέση με τη φανέλα του ποδοσφαιριστή σε ένα συγκεκριμένο παιχνίδι. Το σχήμα παρουσιάζει ποιοι παίκτες έχουν παίξει με τις δύο ομάδες, ως ιστορικός πίνακας. Κάποιοι παίκτες δεν αγωνίζονται πλέον.

Επίσης υπάρχει και ένα άλλο ενδεχόμενο: κάποιοι ποδοσφαιριστές αλλάζουν ομάδα. Ενδέχεται λοιπόν κάποιος ποδοσφαιριστής της *Milan* να έχει εμφανιστεί, κάποια άλλη περίοδο, ως παίκτης της *Real*. Μπορεί βέβαια να συμβεί και το ανάποδο. Μία τέτοια περίπτωση είναι ο βραζιλιάνος ποδοσφαιριστής Kaká ο οποίος έχει αγωνιστεί και με τις δύο ομάδες.

Αν λοιπόν δεν ισχύει ο περιορισμός ενός ποδοσφαιρικού αγώνα, όπου κάθε ποδοσφαιριστής συμμετέχει σε μία μόνο ομάδα, αλλά εξεταστούν οι συμμετοχές των ποδοσφαιριστών σε ομάδες για μια μεγάλη χρονική περίοδο. τότε όλες οι απαντήσεις που δώσαμε πριν πρέπει να εξεταστούν εκ νέου.

Ας δούμε πρώτα την ένωση. Δηλαδή:

$$Milan \cup Real$$

Το αποτέλεσμα της ένωσης, για το δείγμα δεδομένων, απεικονίζεται στο σχήμα 1.14. Οι πλειάδες των δύο σχέσεων ενώνονται, στο αποτέλεσμα εμφανίζονται όλοι οι παίκτες. Ωστόσο, κάθε παίκτης εμφανίζεται μία φορά. Για παράδειγμα, ο Kaká εμφανίζεται μία φορά στο αποτέλεσμα.

number	name
806	Marek Jankulovski
705	Kaká
580	Alessandro Nesta
322	Andriy Shevchenko
205	Claude Makélélé
108	Raúl Albiol
203	Xabi Alonso

$Milan \cup Real$

Σχήμα 1.14: Ένωση των σχέσεων *Milan* και *Real*, όπως απεικονίζονται στο σχήμα 1.13. Στο αποτέλεσμα εμφανίζονται οι ποδοσφαιριστές που έχουν αγωνιστεί είτε με τη Milan είτε με την Real.

Η τομή των σχέσεων *Milan* και *Real*:

$$Milan \cap Real$$

θα δώσει ως αποτέλεσμα τις κοινές πλειάδες των δύο σχέσεων, δηλαδή τους ποδοσφαιριστές που έχουν αγωνιστεί και με τις δύο ομάδες.

Το αποτέλεσμα απεικονίζεται στο σχήμα 1.15. Για το συγκεκριμένο δείγμα δεδομένων, μόνο ο βραζιλιάνος Kaká έχει αγωνιστεί και με τις δύο ομάδες, οπότε η πλειάδα <505, Kaká> υπάρχει στο αποτέλεσμα της τομής.

number	name
705	Kaká

$Milan \cup Real$

Σχήμα 1.15: Τομή των σχέσεων *Milan* και *Real*, όπως απεικονίζονται στο σχήμα 1.15. Στο αποτέλεσμα εμφανίζονται οι ποδοσφαιριστές που έχουν αγωνιστεί και με τη Milan και με την Real.

Η διαφορά των σχέσεων *Milan* και *Real* θα μας δώσει στο αποτέλεσμα του ποδοσφαιριστές που έχουν αγωνιστεί μόνο με την μία από τις δύο ομάδες.

Για παράδειγμα, η διαφορά:

$$Milan - Real$$

θα επιστρέψει εκείνες τις πλειάδες της σχέσης *Milan* οι οποίες δεν βρίσκονται στη σχέση *Real*, δηλαδή τους ποδοσφαιριστές της Milan που δεν έχουν αγωνιστεί με την Real.

Ενώ η διαφορά:

$$Real - Milan$$

θα επιστρέψει εκείνες τις πλειάδες της σχέσης *Real* οι οποίες δεν βρίσκονται στη σχέση *Milan*, δηλαδή τους ποδοσφαιριστές της *Real* οι οποίοι δεν έχουν αγωνιστεί με την *Milan*.

Τα αποτελέσματα των δύο αυτών σχεσιακών πράξεων απεικονίζονται στο σχήμα 1.16. Παρατηρείστε πως τα αποτελέσματα είναι διαφορετικά, κάτι που επιβεβαιώνει την μη ισχύ της αντιμεταθετικής ιδιότητας στη σχεσιακή πράξη της διαφοράς.

number	name
806	Marek Jankulovski
580	Alessandro Nesta
322	Andriy Shevchenko

Milan − Real

number	name
205	Claude Makélélé
108	Raúl Albiol
203	Xabi Alonso

Real − Milan

Σχήμα 1.16: Διαφορά *Milan−Real* και *Real−Milan*. Ποδοσφαιριστές της Milan που δεν έχουν αγωνιστεί με τη Real και ποδοσφαιριστές της Real που δεν έχουν αγωνιστεί με την Milan, αντίστοιχα.

Ένα ενδιαφέρον αποτέλεσμα θα ήταν να είχαμε τους ποδοσφαιριστές των δύο ομάδων που δεν έχουν αγωνιστεί με άλλη ομάδα. Δηλαδή την ένωση των πλειάδων της σχέσης *Milan* και *Real* χωρίς τον συνυπολογισμό των κοινών τους πλειάδων. Στο συγκεκριμένο παράδειγμα του σχήματος 1.13, όλους τους ποδοσφαιριστές, εκτός του βραζιλιάνου Kaká που έχει αγωνιστεί και με τις δύο ομάδες.

Στη σχεσιακή άλγεβρα αυτό μπορεί πολύ εύκολα να υπολογιστεί ως εξής:

$$(Milan \cup Real) - (Milan \cap Real) \tag{1.13}$$

Δηλαδή, από την ένωση των σχέσεων *Milan* και *Real* αφαιρούμε την τομή των των σχέσεων *Milan* και *Real*. Το αποτέλεσμα θα είναι αυτό του σχήματος 1.17.

Ταινιοθήκη

Ας υποθέσουμε ένα φανταστικό σενάριο, μιας και το θέμα αυτής της ενότητας έχει να κάνει με τον κινηματογράφο.

Βρισκόμαστε στα 1955, στην Ιταλία. Εκείνη την εποχή δύο ταινίες του παραγωγού (και σεναριογράφου) Marcello Girosi έχουν κάνει μεγάλη επιτυχία: η ταινία Pane, amore e fantasia[3] και η ταινία Pane, amore e gelosia.[4]

[3] `http://www.imdb.com/title/tt0046159`
[4] `http://www.imdb.com/title/tt0047327`

number	name
806	Marek Jankulovski
580	Alessandro Nesta
322	Andriy Shevchenko
205	Claude Makélélé
108	Raúl Albiol
203	Xabi Alonso

$Milan \cup Real$

Σχήμα 1.17: Ένωση των των σχέσεων *Milan* και *Real*, με αποκλεισμό των κοινών πλειάδων, όπως υπολογίστηκαν με βάση τη σχέση 1.13.

Το σχήμα 1.18 δείχνει ένα αντιπροσωπευτικό δείγμα από τους ηθοποιούς που εμφανίστηκαν στις δύο ταινίες. Οι δύο σχέσεις έχουν από δύο γνωρίσματα, το κωδικό κάθε ηθοποιού[5] και το όνομά του.

Η σχέση f για την ταινία Pane, amore e fantasia:

$$f(actoID, name) \tag{1.14}$$

Η σχέση g για την ταινία Pane, amore e gelosia:

$$g(actoID, name) \tag{1.15}$$

actorID	name
0001120	Vittorio De Sica
0518178	Gina Lollobrigida
0581028	Marisa Merlini
0139214	Memmo Carotenuto
0728376	Roberto Risso
0681365	Tina Pica
0188022	Vittoria Crispo

f

actorID	name
0001120	Vittorio De Sica
0518178	Gina Lollobrigida
0581028	Marisa Merlini
0139214	Memmo Carotenuto
0681365	Tina Pica
0882237	Saro Urzì
0188022	Vittoria Crispo

g

Σχήμα 1.18: Ηθοποιοί που εμφανίστηκαν στις ταινίες Pane, amore e fantasia (δεξιά) και Pane, amore e fantasia (αριστερά). Το σχήμα δείχνει μέρος μόνο των ηθοποιών που εμφανίστηκαν. Το δείγμα ωστόσο είναι αντιπροσωπευτικό και επαρκεί πλήρως για τις ανάγκες του παραδείγματος της ενότητας.

Το 1955 θα κυκλοφορήσει και η τρίτη ταινία της τριλογίας, η ταινία Pane, amore e...[6]

Υποθετικά, ας βρεθούμε με τη φαντασία μας στην Ιταλία του 1955, πριν το γύρισμα της τρίτης ταινίας της τριλογίας «Pane, amore». Λίγο πριν το γύρισμα της τρίτης ταινίας της τριλογίας «Pane, amore» υπάρχουν προτάσεις για αλλαγές στη σύνθεση των ηθοποιών. Μερικές από τις προτάσεις φαίνονται στο σχήμα 1.19.

[5] Όπως δίνονται από την IMDB, `http://www.imdb.com`

[6] `http://www.imdb.com/title/tt0050817`

Η σχέση e για την ταινία Pane, amore e...:

$$e(actoID, name) \tag{1.16}$$

actorID	name
0001120	Vittorio De Sica
0000047	Sophia Loren
0655833	Lea Padovani
0681365	Tina Pica
0139213	Mario Carotenuto

Σχήμα 1.19: Προτεινόμενες συμμετοχές στην ταινία Pane, amore e...

Ας υποθέσουμε λοιπόν πως κάνουμε μια μελέτη για τις τρεις αυτές ταινίες.

Οι ηθοποιοί που έχουν παίξει στις δύο πρώτες ταινίες μπορούν να βρεθούν με την ένωση των σχέσεων f και g:

$$f \cup g$$

Το αποτέλεσμα της ένωσης φαίνεται στο σχήμα 1.20.

actorID	name
0001120	Vittorio De Sica
0518178	Gina Lollobrigida
0581028	Marisa Merlini
0139214	Memmo Carotenuto
0728376	Roberto Risso
0681365	Tina Pica
0188022	Vittoria Crispo
0882237	Saro Urzì

Σχήμα 1.20: Ηθοποιοί των δύο πρώτων ταινιών της τριλογίας «Pane, amore».

Πολλοί ηθοποιοί έχουν εμφανιστεί και στις δύο ταινίες, όπως φαίνεται στο σχήμα 1.18. Ωστόσο, στο αποτέλεσμα της ένωσης, σχήμα 1.20, κάθε ηθοποιός εμφανίζεται μόνο μία φορά.

Υπάρχουν αρκετοί ηθοποιοί που εμφανίστηκαν και στις δύο ταινίες. Αυτοί μπορούν να βρεθούν από την τομή:

$$f \cap g$$

Το αποτέλεσμα της τομής φαίνεται στο σχήμα 1.21 το οποίο απεικονίζει τους ηθοποιούς που εμφανίστηκαν στις δύο πρώτες ταινίες της τριλογίας «Pane, amore». Κάποιοι από τους ηθοποιούς της ταινίας Pane, amore e fantasia δεν εμφανίζονται στο αποτέλεσμα γιατί δεν έχουν συμμετοχή στην ταινία Pane, amore e gelosia. Το ίδιο συμβαίνει και για κάποιους από τους ηθοποιούς της ταινίας Pane, amore e gelosia, δεν εμφανίζονται στο αποτέλεσμα γιατί δεν έχουν παίξει στην ταινία Pane, amore e fantasia.

Ας δούμε τώρα τι συμβαίνει με τους ηθοποιούς της τρίτης ταινίας (Pane, amore e...). Οι ηθοποιοί της που έχουν παίξει σε κάποια από τις δύο πρώτες ταινίες είναι:

$$e \cap (f \cup g)$$

actorID	name
0001120	Vittorio De Sica
0518178	Gina Lollobrigida
0581028	Marisa Merlini
0139214	Memmo Carotenuto
0681365	Tina Pica
0188022	Vittoria Crispo

Σχήμα 1.21: Ηθοποιοί των δύο πρώτων ταινιών της τριλογίας «Pane, amore» που εμφανίστηκαν και στις δύο ταινίες.

actorID	name
0001120	Vittorio De Sica
0681365	Tina Pica

Σχήμα 1.22: Ηθοποιοί της τρίτης ταινίας της τριλογίας «Pane, amore» που εμφανίστηκαν σε κάποια από τις δύο πρώτες δύο ταινίες.

Το αποτέλεσμα φαίνεται στο σχήμα 1.22. Η διαπίστωση μπορεί να γίνει εύκολα συγκρίνοντας τα σχήματα 1.18 και 1.19.

Χρειάζεται προσοχή στη διατύπωση της παράστασης σχεσιακής άλγεβρας. Για παράδειγμα, η αλλαγή των παρενθέσεων:

$$(e \cap f) \cup g$$

θα δώσει τελείως διαφορετικό αποτέλεσμα: Στους κοινούς ηθοποιούς των δύο πρώτων ταινιών (τομή) προσθέτουμε (ένωση) τους ηθοποιούς της τρίτης ταινίας. Το αποτέλεσμα φαίνεται στο σχήμα 1.23.

actorID	name
0001120	Vittorio De Sica
0518178	Gina Lollobrigida
0581028	Marisa Merlini
0139214	Memmo Carotenuto
0681365	Tina Pica
0188022	Vittoria Crispo
0000047	Sophia Loren
0655833	Lea Padovani
0139213	Mario Carotenuto

Σχήμα 1.23: Αποτέλεσμα της σχεσιακής πράξης $(e \cap f) \cup g$.

Μπορούμε επίσης να βρούμε τους ηθοποιούς της τρίτης ταινίας που δεν έχουν εμφανιστεί σε καμία από τις δύο πρώτες:

$$e - (f \cup g)$$

actorID	name
0000047	Sophia Loren
0655833	Lea Padovani
0139213	Mario Carotenuto

Σχήμα 1.24: Αποτέλεσμα της σχεσιακής πράξης $(e \cap f) \cup g$.

Το αποτέλεσμα φαίνεται στο σχήμα 1.24, το οποίο απεικονίζει τους ηθοποιούς της ταινίας Pane, amore, e... που δεν εμφανίστηκαν ούτε στην ταινία Pane, amore, e fantasia ούτε στην ταινία Pane, amore, e gelosia.

Η αλλαγή της θέσης των παρενθέσεων θα δώσει εσφαλμένο αποτέλεσμα, παρόμοιο με αυτό που είδαμε λίγο παραπάνω.

Ένα άλλο συχνό λάθος είναι η χρήση της τομής αντί της ένωσης:

$$e - (f \cap g)$$

Το ερώτημα αυτό θα επιστρέψει στο αποτέλεσμα τους ηθοποιούς της τρίτης ταινίας που δεν έχουν εμφανιστεί και στις δύο προηγούμενες ταινίες. Προσοχή, το λάθος δεν είναι γραμματικό, είναι εννοιολογικό και αναφέρεται στη συγκεκριμένη ερώτηση.

Λεπτές διαφορές στη διατύπωση του ερωτήματος απαιτούν διαφορετική διατύπωση στην παράσταση σχεσιακής άλγεβρας. Με τον ίδιο τρόπο, μικρές διαφορές στην παράσταση σχεσιακής άλγεβρας (τελεστές, παρενθέσεις, κ.λπ.) προκαλούν αλλαγή στο αποτέλεσμα, στην ουσία απαντούν σε άλλο ερώτημα.

Τέλος, μπορούμε να βρούμε τους κοινούς ηθοποιούς και των τριών ταινιών:

$$f \cap g\ cape$$

actorID	name
0001120	Vittorio De Sica
0681365	Tina Pica

Σχήμα 1.25: Ηθοποιοί που εμφανίστηκαν σε όλες τις ταινίες της τριλογίας «Pane, amore».

Το αποτέλεσμα της τομής φαίνεται στο σχήμα 1.25. Υπενθυμίζεται πως για την τομή ισχύει η προσεταιριστική ιδιότητα, η σειρά των πράξεων δεν έχει σημασία:

$$(f \cap g) \cap e = f \cap (g \cap e)$$

Προτιμήσεις καταναλωτή

Μια συνοικιακή πιτσαρία δέχεται παραγγελίες πίτσας μέσω μηνυμάτων κινητού τηλεφώνου (SMS). Τα μηνύματα επεξεργάζονται μέσω Η/Υ και οι παραγγελίες προωθούνται προς εκτέλεση.

Τρία διαδοχικά απογεύματα, η πιτσαρία δέχθηκε παραγγελίες από τον ίδιο πελάτη.

Το περιεχόμενο της κάθε παραγγελίας φαίνεται στο σχήμα 1.26.

product
Margarita
Napolitana
Beer

*order*1

product
Calzone
Napolitana
Beer

*order*2

product
Calzone
Napolitana
Lasagne

*order*3

Σχήμα 1.26: Τρεις παραγγελίες πελατών μιας πιτσαρίας.

Ο υπεύθυνος πωλήσεων της πιτσαρίας ζητά τη βοήθειά σας για την ανάλυση αυτών των παραγγελιών.[7]

Τι είναι κοινό σε όλες τις παραγγελίες; Αυτό είναι ένα ερώτημα ένωσης:

$$order1 \cup order2 \cup order3$$

Το αποτέλεσμα της ένωσης φαίνεται στο σχήμα 1.27.

product
Margarita
Napolitana
Beer
Calzone
Lasagne

Σχήμα 1.27: Το αποτέλεσμα της ένωσης τριών παραγγελιών της πιτσαρίας: $order1 \cup order2 \cup order3$.

Τι κοινό έχουν και οι τρεις παραγγελίες; Αυτό είναι ένα ερώτημα τομής:

$$order1 \cap order2 \cap order3$$

Το αποτέλεσμα της τομής φαίνεται στο σχήμα 1.28.

product
Napolitana

Σχήμα 1.28: Το αποτέλεσμα της τομής των τριών παραγγελιών της πιτσαρίας: $order1 \cap order2 \cap order3$.

Τι υπάρχει στην τρίτη παραγγελία, που δεν υπάρχει στις δύο πρώτες; Αυτό είναι ένα ερώτημα διαφοράς:

$$order3 - (order1 \cup order2)$$

Το αποτέλεσμα της διαφοράς φαίνεται στο σχήμα 1.29.

[7] Μην παρασύρεστε από την απλότητα του συγκεκριμένου παραδείγματος. Η ανάλυση παραγγελιών και πωλήσεων είναι στοιχειώδης λειτουργία κάθε σοβαρής επιχείρησης στις μέρες μας. Κάτι τέτοιο απαιτεί βέβαια πολύ περισσότερες λειτουργίες από τα απλά παραδείγματα που δίνουμε εδώ. Δείτε το κεφάλαιο ;; για περισσότερες λεπτομέρειες.

product
Lasagne

Σχήμα 1.29: Τα προϊόντα της τρίτης παραγγελίας που δεν υπήρχαν στις πρώτες δύο παραγγελίες.

Τι υπάρχει στην πρώτες δύο παραγγελίες που δεν υπάρχει στην τρίτη; Αυτό είναι ένα ερώτημα διαφοράς:

$$(order1 \cup order2) - order3)$$

Το αποτέλεσμα της διαφοράς φαίνεται στο σχήμα 1.30.

product
Beer

Σχήμα 1.30: Τα προϊόντα των πρώτων δύο παραγγελιών που δεν υπήρχαν στην τρίτη παραγγελία.

Συμπέρασμα: Ο πελάτης παραγγέλνει σταθερά πίτσα «Ναπολιτάνα», αλλά φαίνεται να εγκαταλείπει την παραγγελία μπύρας.

1.4 Οι βασικές πράξεις προβολής και επιλογής

1.4.1 Προβολή

Προβολή μιας σχέσης r, με σχήμα R, πάνω στο υποσύνολο γνωρισμάτων της X ($Q \subseteq R$) είναι μια σχέση με σχήμα το σύνολο X και κορμό εκείνες τις πλειάδες που αντιστοιχούν σε μοναδικές τιμές για τα γνωρίσματα X.

Ορισμός της προβολής:

$$r[X] = \{t[X] \,|\, t \in r\} \tag{1.17}$$

Με απλά λόγια, η προβολή μιας σχέσης, προκύπτει από την αφαίρεση κάποιων γνωρισμάτων της και την απαλοιφή των πιθανών διπλότυπων εγγραφών από τις πλειάδες που προκύπτουν. Γιατί είναι πιθανό να συμβεί αυτό; Μα επειδή από τη λίστα γνωρισμάτων της προβολής μπορεί ενδεχομένως να λείπει το πρωτεύον ή άλλο υποψήφιο κλειδί.

Η προβολή συμβολίζεται με το ελληνικό γράμμα Π:

$$\Pi_{A_1, A_2, \ldots, A_m}(r)$$

Είναι δυνατόν να μην υπάρχει κανένα γνώρισμα στη λίστα γνωρισμάτων, και τότε υπονοείται πως η προβολή γίνεται σε όλα τα γνωρίσματα. Σε αυτή την περίπτωση έχουμε τη λεγόμενη ***ταυτοτική*** προβολή. Η προβολή δηλαδή μιας σχέσης στον εαυτό της, η οποία συμβολίζεται απλά με όνομα της σχέσης: $\Pi(r)$

Αυτό σημαίνει, πως το όνομα μιας σχέσης, είναι από μόνο του μια σχεσιακή παράσταση.

A	B	C
5	a	30
2	b	10
3	c	20
5	b	10

r

A	B
5	a
2	b
3	c
5	b

$\Pi_{A,B}(r)$

B	C
a	30
b	10
c	20

$\Pi_{B,C}(r)$

B
a
b
c

$\Pi_{B}(r)$

A	B	C
5	a	30
2	b	10
3	c	20
5	b	10

$\Pi(r)$

Σχήμα 1.31: Παραδείγματα προβολής.

Μερικά παραδείγματα προβολής φαίνονται στο σχήμα 1.31. Ας τα δούμε ένα-ένα. Στην πρώτη περίπτωση, $\Pi_{A,B}(r)$, έχουμε:

$$\Pi_{A,B}(r) = \begin{array}{cc} \hline A & B \\ \hline 5 & a \\ 2 & b \\ 3 & c \\ 5 & b \\ \hline \end{array}$$

Δηλαδή, από τα γνωρίσματα της σχέσης r επιλέγονται τα A, B και παραλείπεται το γνώρισμα C. Από τις πλειάδες της σχέσης r μένουν στο αποτέλεσμα όσες, για τα γνωρίσματα A, B έχουν μοναδικές τιμές.

Στην δεύτερη περίπτωση του σχήματος 1.31, $\Pi_{B,C}(r)$, έχουμε:

$$\Pi_{B,C}(r) = \begin{array}{cc} \hline B & C \\ \hline a & 30 \\ b & 10 \\ c & 20 \\ \hline \end{array}$$

Δηλαδή, από τα γνωρίσματα της σχέσης r επιλέγονται τα B, C και παραλείπεται το γνώρισμα A. Από τις πλειάδες της σχέσης r μένουν στο αποτέλεσμα όσες, για τα γνωρίσματα B, C, έχουν μοναδικές τιμές. Παρατηρείστε ότι η πλειάδα $\langle b, 10\rangle$ υπάρχει μία φορά στο αποτέλεσμα.

Στην τρίτη περίπτωση του σχήματος 1.31, $\Pi_B(r)$, έχουμε:

$$\Pi_{B}(r) = \begin{array}{c} \hline B \\ \hline a \\ b \\ c \\ \hline \end{array}$$

Δηλαδή, από τα γνωρίσματα της σχέσης r επιλέγεται το γνώρισμα B και παραλείπονται τα γνωρίσματα A, C. Από τις πλειάδες της σχέσης r μένουν στο αποτέλεσμα όσες, για το γνώρισμα B, έχουν μοναδικές τιμές. Παρατηρείστε ότι η πλειάδα $\langle b\rangle$ υπάρχει μία φορά στο αποτέλεσμα.

Στην τελευταία περίπτωση του σχήματος 1.31, $\Pi(r)$, έχουμε την ταυτοτική προβολή:

$$\Pi_{B}(r) = \begin{array}{ccc} \hline A & B & C \\ \hline 5 & a & 30 \\ 2 & b & 10 \\ 3 & c & 20 \\ 5 & b & 10 \\ \hline \end{array}$$

Δηλαδή, από τα γνωρίσματα της σχέσης r επιλέγονται όλα τα τα γνωρίσματα.

Η SQL χρησιμοποιεί τη φράση:

```
SELECT λίστα πεδίων
  FROM παράσταση πίνακα
```

για την επιλογή δεδομένων από πίνακες. Η ομοιότητα με την σχεσιακή πράξη της προβολής είναι προφανής. Για την επιλογή όλων των πεδίων ενός πίνακα, δηλαδή την ταυτοτική προβολή μιας σχέσης, με όρους σχεσιακής άλγεβρας, θα γράφουμε:

```
SELECT *
  FROM πίνακας
```

1.4.2 Επιλογή

Η επιλογή ή αλλιώς και περιορισμός μιας σχέσης $r(R)$, είναι μια σχέση που έχει το ίδιο σχήμα R με τη σχέση r και κορμό ένα υποσύνολο του κορμού της r που ικανοποιεί μια συνθήκη, πχ: $X\ \theta\ Y$.

Η επιλογή συμβολίζεται με:

$$A \quad WHERE \quad X \quad \theta \quad Y$$

ή αλλιώς: $\sigma_{condition}(r)$, όπου η συνθήκη περιορισμού είναι μια παράσταση που μπορεί να αποτιμηθεί σε TRUE, FALSE ή NULL. Οι παραστάσεις μπορούν επίσης να περιέχουν τους λογικούς τελεστές AND $\wedge$, OR $\vee$ και NOT $\neg$.

Ο τελεστής θ μπορεί να είναι ένας από $=, \neq, <, \leq, >, \geq$. Η τιμή ενός γνωρίσματος μπορεί να συγκριθεί με την τιμή ενός άλλου γνωρίσματος, ή με μια κυριολεκτική τιμή, ή ακόμα και με μια άλλη σχεσιακή παράσταση που αποδίδει κάποια τιμή (εμφώλευση ερωτημάτων).

Αν ϕ είναι μια παράσταση που χρησιμοποιεί τελεστές σύγκρισης ή/και λογικούς τελεστές, τότε η επιλογή μπορεί να οριστεί ως:

Ορισμός της επιλογής:

$$\sigma_{\phi}(r) = \{t \in r \,|\, t \text{ satisfies } \phi\} \tag{1.18}$$

Η SQL υποστηρίζει την πράξη της επιλογή με την προσθήκη της φράσης WHERE:

```
SELECT λίστα πεδίων
  FROM παράσταση πίνακα
 WHERE συνθήκη
```

empid	name	salary
101	Αθανασίου Μιχ.	1200
102	Βαφειάδης Νικ.	1150
104	Νικολοπούλου Ναν.	1570
108	Βασιλειάδη Μαρ.	1320

Σχήμα 1.32: Οι υπάλληλοι μιας εταιρείας, δείγμα δεδομένων.

Για παράδειγμα αν έχουμε τη σχέση $employees(empid, name, salary)$ των υπαλλήλων που καταγράφει τον κωδικό, το όνομα και το μισθό τους (σε €), όπως φαίνεται στο σχήμα 1.32, τότε μπορούμε να γράψουμε παραστάσεις όπως:

1. Να βρεθούν οι υπάλληλοι με μισθό μικρότερο από 1300 €:

$$\sigma_{salary<1300}(employees)$$

empid	name	salary
101	Αθανασίου Μιχ.	1200
102	Βαφειάδης Νικ.	1150

2. Να βρεθούν οι υπάλληλοι με μισθό μεταξύ 1200 και 1600 €:

$$\sigma_{salary\geq 1200 \wedge salary\leq 1600}(employees)$$

empid	name	salary
101	Αθανασίου Μιχ.	1200
104	Νικολοπούλου Ναν.	1570
108	Βασιλειάδη Μαρ.	1320

3. Να βρεθεί ο υπάλληλος με κωδικό 102:

$$\sigma_{empid=102}(employees)$$

empid	name	salary
102	Βαφειάδης Νικ.	1150

1.4.3 Παραδείγματα και ασκήσεις προβολής και επιλογής

Στην ενότητα αυτή θα δούμε παραδείγματα προβολής και επιλογής. Για τα παραδείγματα θα χρησιμοποιήσουμε τη σχέση *employees* με σχήμα:

$$employees(empid, firstname, lastname, depid, salary, hiredate)$$

που περιγράφει τους υπαλλήλους μιας εταιρείας. Μέρος από τα περιεχόμενα του πίνακα δίνεται στον πίνακα1.1. Τα πλήρη δεδομένα δίνονται στον πίνακα ;;, της σελίδας ;;.

Πίνακας 1.1: Μέρος από τα περιεχόμενα του πίνακα *employees*. Δείτε τον πίνακα ;; της σελίδας ;; για πλήρη απεικόνιση.

empid	**firstname**	**lastname**	**depid**	**salary**	**hiredate**
102	Νικηφόρος	Διαμαντίδης	6	1212.50	2003-06-02
109	Μαρία	Αθανασίου	1	2787.69	2000-01-26
153	Μαρία	Αλεβιζάτου	2	1321.92	2001-05-15
172	Χρήστος	Βλάσσης	3	1101.70	2000-07-04
189	Θεόδωρος	Αγγελίνας	6	1908.28	2000-06-19
...	...	...	...	...	...

Στα παραδείγματα θα δούμε πράξεις προβολής, επιλογής ή και συνδυασμούς των δύο πράξεων. Εκεί όπου το αποτέλεσμα έχει περισσότερες από 5 πλειάδες (εγγραφές πίνακα) οι επιπλέον πλειάδες δείχνονται με τρεις τελείες, κάτι που υπονοεί πως υπάρχουν υπάρχουν επιπλέον πλειάδες. Η αποκοπή γίνεται καθαρά για λόγους οικονομίας χώρου.

Οι απαντήσεις στα ερωτήματα δίνονται επίσης στα κεφάλαια ;;–;; με τη γλώσσα SQL. Έτσι μπορείτε να συνδυάσετε τις γνώσεις σας στη σχεσιακή άλγεβρα με τη γλώσσα ερωτημάτων SQL.

1. Να βρεθεί το επώνυμο των υπαλλήλων:

$$\Pi_{salary}(employees)$$

2. Να βρεθεί το όνομα και το επώνυμο όλων των υπαλλήλων:

$$\Pi_{firstname,lastname}(employees)$$

3. Να βρεθούν οι υπάλληλοι με μισθό μεγαλύτερο του 1500:

$$\sigma_{salary>1500}(employees)$$

4. Να βρεθεί το όνομα και το επώνυμο όλων των υπαλλήλων που παίρνουν μισθό μεγαλύτερο από 1500:

$$\Pi_{firstname,lastname}(\sigma_{salary>1500}(employees))$$

5. Να βρεθεί το ονοματεπώνυμο και ο μισθός όλων των υπαλλήλων που εργάζονται στο τμήμα με κωδικό 2:

$$\Pi_{firstname,lastname,salary}(\sigma_{depid=2}(employees))$$

6. Να βρεθεί το ονοματεπώνυμο, το τμήμα και ο μισθός όλων των υπαλλήλων που που εργάζονται σε τμήμα με κωδικό διαφορετικό του 1:

$$\Pi_{firstname,lastname,depid,salary}(\sigma_{depid<>1}(employees))$$

7. Να βρεθεί το όνομα και ο κωδικός των υπαλλήλων που προσλήφθηκαν πριν από την 1/8/2000:

$$\Pi_{firstname,lastname,empid}(\sigma_{hiredate<'2000-08-01'}(employees))$$

8. Να βρεθεί το επώνυμο και ο μισθός του υπαλλήλου με κωδικό 109 μετά την αύξηση 5% στο μισθό του:

$$\Pi_{lastname,salary*1.05}(\sigma_{empid=109}(employees))$$

9. Να βρεθούν όλες οι λεπτομέρειες των υπαλλήλων με μισθό μεγαλύτερο του 1200 ή με μισθό μικρότερο του 1600:

$$\sigma_{salary>1200 \vee salary<1600}(employees)$$

10. Να βρεθεί ο κωδικός, το όνομα, ο μισθό, και το τμήμα όλων των υπαλλήλων του τμήματος 2, αλλά και των υπαλλήλων από άλλα τμήματα με μισθό στο εύρος $[1330, 1630)$:

$$\sigma_{(salary>1330 \wedge salary<1630) \vee depid=2}(employees)$$

11. Να βρεθούν οι υπάλληλοι (κωδικός, επώνυμο, τμήμα) που εργάζονται στο τμήμα 2 και παίρνουν μισθό μικρότερο από 1200

$$\Pi_{empid,lastname,depid}\left(\sigma_{(depid=2 \vee salary<1200}(employees)\right)$$

12. Να βρεθούν οι υπάλληλοι που δεν εργάζονται στα τμήματα 2, 3, 4:

$$\sigma_{\neg(depid=2 \vee depid=3 \vee depid=4)}(employees)$$

1.5 Συζεύξη

1.5.1 Φυσική σύζευξη

Η σύζευξη μοιάζει με το γινόμενο, ωστόσο έχει μια σημαντική διαφορά: οι πλειάδες του αποτελέσματος δεν είναι όλοι οι πιθανοί συνδυασμοί, αλλά μόνο οι συνδυασμοί που ικανοποιούν κάποια συνθήκη. Υπάρχουν διάφορες παραλλαγές συζεύξεων, με σημαντικότερη και περισσότερο διαδεδομένη τη σύζευξη ισότητας, ή αλλιώς φυσική σύζευξη, η οποία στηρίζεται στην παρουσία κοινών γνωρισμάτων στις δύο σχέσεις.

Αν η r είναι σχέση με σχήμα $R = \{X, Y\}$ και s είναι σχέση με σχήμα $S = \{Y, Z\}$, τότε η **φυσική σύζευξη** των r και s είναι μια σχέση με σχήμα $R \cup S = \{X, Y, Z\}$ και κορμό το σύνολο των συνδυασμών των πλειάδων της r και s για τις οποίες οι τιμές στο κοινό γνώρισμα Y ταυτίζονται. Δηλαδή, μια πλειάδα της r, θα συνδυαστεί με μια πλειάδα της s, αν και μόνο αν, οι τιμές στο κοινό γνώρισμα Y ταυτίζονται μεταξύ τους. Η φυσική σύζευξη των σχέσεων r και s συμβολίζεται με $r \bowtie s$, ή $r \quad NATURAL \quad JOIN \quad s$, ή απλά $r \quad JOIN \quad s$.

Ορισμός της φυσικής σύζευξης:

$$r \bowtie s = \{t \,|\, \text{ υπάρχουν πλειάδες } u \in r \text{ και } v \in s \text{ έτσι ώστε } t[R] = u \text{ και } t[S] = v\} \tag{1.19}$$

Τα σύμβολα X, Y, Z δηλώνουν σύνολα γνωρισμάτων. Επομένως σύζευξη μπορεί να γίνει με περισσότερα από ένα γνωρίσματα, αν το Y για παράδειγμα, αποτελείται από δύο ή περισσότερα γνωρίσματα.

Το σχήμα 1.33 απεικονίζει ένα παράδειγμα σύζευξης. Μπορούμε να παρατηρήσουμε τα εξής:

1. Οι σχέσεις r με σχήμα $R = \{A, B, C\}$ και s με σχήμα $\{C, D\}$ έχουν ένα κοινό γνώρισμα, το C. Σύμφωνα με τα παραπάνω μπορούμε να πούμε πως $X = \{A, B\}$, $Y = \{C\}$ και $Z = \{D\}$.

A	B	C
1	b	10
5	a	30
3	c	20

r

C	D
20	1
10	3
20	3

s

A	B	C	D
1	b	10	3
3	c	20	1
3	c	20	3

$r \bowtie s$

Σχήμα 1.33: Παράδειγμα σχεσιακής σύζευξης: $r \bowtie s$.

2. Το σχήμα της σχέσης που προκύπτει από τη σύζευξη $r \bowtie s$ είναι $R \cup S = \{A, B, C, D\}$. Προσοχή, το S υπάρχει μία φορά ως γνώρισμα στο σχήμα της σχέσης που προκύπτει από τη σύζευξη.
3. Κάθε πλειάδα της r ενώνεται με την αντίστοιχη πλειάδα της s, όταν οι τιμές στο κοινό γνώριμα C ταυτίζονται.
4. Η πλειάδα $\langle 1, b, 10 \rangle$ της r ενώνεται με την πλειάδα $\langle 10, 3 \rangle$ της s γιατί οι τιμές στο κοινό γνώρισμα C ταυτίζονται.
5. Η πλειάδα $\langle 5, a, 30 \rangle$ της r δεν ενώνεται με καμία πλειάδα της s γιατί οι τιμές στο κοινό γνώρισμα C δεν ταυτίζονται.
6. Η πλειάδα $\langle 3, c, 20 \rangle$ της r ενώνεται με δύο πλειάδες της s γιατί υπάρχουν δύο πλειάδες της r, $\langle 20, 1 \rangle$ και $\langle 20, 3 \rangle$, για τις οποίες οι τιμές στο κοινό γνώρισμα C ταυτίζονται.

Για τη φυσική σύζευξη δεν ισχύει η αντιμεταθετική ιδιότητα:

$$r \bowtie s \neq s \bowtie r \tag{1.20}$$

εκτός και αν ακολουθείται από κάποια προβολή:

Σύζευξη και αντιμεταθετική ιδιότητα:

$$\Pi_L(r \bowtie s) = \Pi_L(s \bowtie r) \tag{1.21}$$

όπου L είναι λίστα γνωρισμάτων, $L \subseteq (X \cup Y \cup Z)$.

Ισχύει ωστόσο η προσεταιριστική ιδιότητα:

Σύζευξη και προσεταιριστική ιδιότητα:

$$r \bowtie (s \bowtie t) = (r \bowtie s) \bowtie t \tag{1.22}$$

οπότε μια απλοποιημένη παράσταση, χωρίς παρενθέσεις, είναι εξ΄ ίσου έγκυρη:

$$r \bowtie s \bowtie t = r \bowtie s \bowtie t \tag{1.23}$$

Η SQL υποστηρίζει τη (φυσική) σύζευξη με τον όρο NATURAL JOIN. Για παράδειγμα, για δύο πίνακες που έχουν τουλάχιστον ένα κοινό γνώρισμα, μπορούμε να γράψουμε:

```
SELECT *
  FROM πίνακας1 NATURAL JOIN πίνακας2
```

Η σύζευξη μπορεί να εκφραστεί ως παράγωγη πράξη του γινομένου και της επιλογής.

Σύζευξη ως καρτεσιανό γινόμενο :

$$r \bowtie' s = \sigma_{\phi}(r \times s) \tag{1.24}$$

όπου ϕ η συνθήκη σύζευξης:

$$\phi = \bigwedge_{Y \in R \cup S} R.Y = S.Y$$

r

A	B	C
1	b	10
5	a	30
3	c	20

s

C	D
20	1
10	3
20	3

$r \bowtie' s$

A	B	R.C	S.C	D
1	b	10	10	3
3	c	20	20	1
3	c	20	20	3

$r \bowtie s$

A	B	C	D
1	b	10	3
3	c	20	1
3	c	20	3

Σχήμα 1.34: Παράδειγμα σχεσιακής σύζευξης: $r \bowtie s$ ως συνδυασμός γινομένου και επιλογής $r \bowtie' s$. Το κοινό γνώρισμα, μετά τη μετονομασία σε $R.C$ και $S.C$, υπάρχει δύο φορές στο αποτέλεσμα $r \bowtie' s$.

Ένα παράδειγμα σύζευξης ως συνδυασμού γινομένου και επιλογής φαίνεται στο σχήμα 1.34. Παρατηρείστε ότι το κοινό γνώρισμα C στις δύο σχέσεις r και s του σχήματος 1.34 εμφανίζεται δύο φορές στο αποτέλεσμα του γινομένου, με διαφορετικό όνομα που προκύπτει από τη συνένωση του ονόματος της σχέσης και του ονόματος του γνωρίσματος: $R.C$ και $S.C$ αντίστοιχα. Αν θέλουμε να απαλειφθεί η διπλή εμφάνιση του γνωρίσματος $R.C$ από το αποτέλεσμα, πρέπει να χρησιμοποιήσουμε τη σχεσιακή πράξη της προβολής (Π).

Μπορούμε έτσι να δώσουμε έναν ακόμα εναλλακτικό ορισμό της φυσικής σύζευξης. Αν $Y = \{A_1, A_2, \ldots, A_n\}$ είναι τα κοινά γνωρίσματα στα δύο σχήματα R, S, των σχέσεων r και s αντίστοιχα, τότε η φυσική σύζευξης μπορεί να οριστεί και ως:

Φυσικής σύζευξη:

$$r \bowtie s = \Pi_{R \cup S}(\sigma_{R.A_1 = S.A_2 \wedge R.A_1 = S.A_2 \ldots \wedge R.A_n = S.A_n}(r \times s))\} \tag{1.25}$$

Υπάρχουν περιπτώσεις όπου ανάμεσα σε δύο σχέσεις r και s δεν υπάρχει κανένα κοινό γνώρισμα. Τότε το αποτέλεσμα της φυσικής σύζευξη ισούται με το καρτεσιανό γινόμενο:

Αν για δύο σχέσεις $r(R)$ και $s(S)$ ισχύει $R \cap S = \emptyset$, τότε

$$r \bowtie s = r \times s \tag{1.26}$$

1.5.2 Σύζευξη θήτα

Η σύζευξη θ ($\theta JOIN$) είναι γενίκευση της σύζευξης, και στη ουσία πρόκειται για συνδυασμό των πράξεων γινομένου και επιλογής, έτσι μπορεί να γραφεί:

$$(r \times s) \quad WHERE \quad X \quad \theta \quad Y$$

όπου X, Y ο τελεστής θ δεν είναι κατά ανάγκη ισότητα. Προσέξτε ότι τα γνωρίσματα που εμπλέκονται στη σύζευξη πρέπει να έχουν κοινό πεδίο ορισμού.

Αν η r είναι σχέση με σχήμα $\{A_1, A_2, \ldots, A_n\}$, s είναι σχέση με σχήμα $\{B_1, B_2, \ldots, B_m\}$, τα γνωρίσματα A_i και B_j έχουν το ίδιο πεδίο ορισμού, και θ είναι τελεστής σύγκρισης, $\theta \in \{=, \neq, <, \leq, >, \geq\}$, τότε η θ σύζευξη των R και S, $R \bowtie_{A_i \theta B_j} S$, είναι μια σχέση με σχήμα το σύνολο των γνωρισμάτων των R και S, $\{A_1, A_2, \ldots, A_n B_1, B_2, \ldots, B_m\}$ και κορμό το σύνολο των πλειάδων από κάθε συνδυασμό των πλειάδων των R και S, που ικανοποιούν τη συνθήκη $A_i \theta B_j$.

Το αποτέλεσμα είναι μια σχέση με βαθμό $m + n$, και πληθικότητα ανάμεσα στο 0 και στο $n_R \times n_S$.

Αν κάποια πλειάδα έχει τιμή NULL στο γνώρισμα που συμμετέχει στη σύζευξη (A_i ή B_j), τότε δεν συμμετέχει στο αποτέλεσμα. Με αυτή την έννοια το αποτέλεσμα της σύζευξης, ενδεχομένως, περιέχει λιγότερη πληροφορία, από ότι οι επιμέρους σχέσεις.

Αν ο τελεστής θ είναι το $=$ τότε η σύζευξη καλείται **ισοσύζευξη**, και πρακτικά είναι φυσική σύζευξη.

Ορισμός της σύζευξης θ:

$$r \bowtie_\theta s = \{\sigma_\theta r \times s\} \tag{1.27}$$

A	B	C
1	b	10
5	a	30
3	c	20

r

C	D
20	1
10	3
20	3

s

A	B	R.C	S.C	D
1	b	10	10	3
3	c	20	20	1
3	c	20	20	3

$r \bowtie_{R.C=S.C} s$

A	B	R.C	S.C	D
5	a	30	10	3
3	c	20	10	3

$r \bowtie_{R.C>S.C} s$

Σχήμα 1.35: Παράδειγμα σχεσιακής σύζευξης θ: $r \bowtie_{R.C=S.C} s$ και $r \bowtie_{R.C>S.C} s$.

1.5.3 Πλήρης εξωτερική σύζευξη

Η εξωτερική σύζευξη είναι επέκταση της σύζευξης ανάμεσα σε δύο σχέσεις, στην περίπτωση που υπάρχουν πλειάδες στη μία σχέση, χωρίς ταιριαστές τιμές στην άλλη σχέση στο γνώρισμα (ή τα γνωρίσματα) της σύζευξης. Η εξωτερική σύζευξη δίνει αποτέλεσμα ό,τι και η εσωτερική, αλλά και κάτι επιπλέον: πλειάδες από τη μία ή την άλλη σχέση που δεν έχουν ταιριαστές τιμές με την άλλη σχέση. Αν υπάρχουν τέτοιες πλειάδες, τότε θα εμφανιστούν στο αποτέλεσμα της εξωτερικής σύζευξης, αλλά όχι στο αποτέλεσμα της φυσικής ή εσωτερικής σύζευξης.

Η εξωτερική σύζευξη μπορεί να είναι πλήρης εξωτερική σύζευξη, αριστερή εξωτερική σύζευξη ή δεξιά αριστερή σύζευξη. Θα εξετάσουμε πρώτα την πλήρη εξωτερική σύζευξη και στις δύο επόμενες υποενότητες την αριστερή και δεξιά εξωτερική σύζευξη.

Για παράδειγμα θεωρείστε τις δύο σχέσεις του σχήματος 1.36, που παριστάνουν ένα δείγμα από τα υποκαταστήματα (b, branches) και τους πελάτες (c, customers) μιας εταιρείας. Χρησιμοποιούμε μια απλουστευμένη έκδοση της αντίστοιχης βάσης δεδομένων, όπου μόνο για τους πελάτες δίνονται μόνο το όνομα και η πόλη είναι ορατά για τους πελάτες. Επίσης, μόνο η πόλη και ο αντίστοιχος κωδικός δίνονται για τα υποκαταστήματα.

Έστω πως θέλουμε να βρούμε το αποτέλεσμα της εξωτερικής σύζευξης των δύο σχέσεων με βάση την πόλη:

$$b ⟗ c$$

Το αποτέλεσμα της εξωτερικής σύζευξης $b ⟗ c$ φαίνεται στο σχήμα 1.36. Για λόγους σύγκρισης δίνεται επίσης το αποτέλεσμα της φυσικής σύζευξης. Μπορούμε να παρατηρήσουμε τα εξής:

id	city
1	Αθήνα
2	Πάτρα
3	Θεσσαλονίκη

Υποκαταστήματα (b)

name	city
Νίκος	Πάτρα
Βάσω	Κοζάνη
Αγγελική	Πάτρα
Βασίλης	Αθήνα

Πελάτες (c)

id	city	name
1	Αθήνα	Βασίλης
2	Πάτρα	Νίκος
2	Πάτρα	Αγγελική
3	Θεσσαλονίκη	NULL
NULL	NULL	Βάσω

Εξωτερική σύζευξη $b ⟗ c$

id	city	name
1	Αθήνα	Βασίλης
2	Πάτρα	Νίκος
2	Πάτρα	Αγγελική

Φυσική σύζευξη $b \bowtie c$

Σχήμα 1.36: Πελάτες και υποκαταστήματα μιας εταιρείας

1. Για τις δύο σχέσεις, b και c, γίνεται σύζευξη με βάση την τιμή στο γνώρισμα *city*, που είναι κοινό στις δύο σχέσεις. Οπότε ουσιαστικά γίνεται λόγος για φυσική σύζευξη. Το σχήμα της σχέσης $b ⟗ c$ είναι το ίδιο με αυτό της φυσικής σύζευξης. Στο αποτέλεσμα της εξωτερικής σύζευξης $b ⟗ c$ υπάρχουν όλες οι πλειάδες που υπάρχουν και στο αποτέλεσμα της φυσικής σύζευξης $b \bowtie c$.

2. Υπάρχει το υποκατάστημα 3 στη Θεσσαλονίκη, ενώ δεν υπάρχει κανείς πελάτης από τη Θεσσαλονίκη. Το υποκατάστημα 3 δεν θα εμφανιστεί στο αποτέλεσμα μιας φυσικής ή εσωτερικής σύζευξης. Εμφανίζεται ωστόσο στο αποτέλεσμα της εξωτερικής σύζευξης.

3. Υπάρχει η πελάτης με όνομα Βάσω από την Κοζάνη, χωρίς στην πόλη αυτή να υπάρχει υποκατάστημα. Η Βάσω δεν θα εμφανιστεί στο αποτέλεσμα μιας φυσικής ή εσωτερικής σύζευξης. Εμφανίζεται ωστόσο στο αποτέλεσμα της εξωτερικής σύζευξης.

4. Τα υποκαταστήματα σε Αθήνα και Πάτρα θα εμφανιστούν στο αποτέλεσμα μιας φυσικής ή εσωτερικής σύζευξης, επειδή υπάρχουν πελάτες τόσο στην Αθήνα (Βασίλης) όσο και στην Πάτρα (Νίκος, Αγγελική).

5. Συνέπεια του προηγούμενου είναι ότι, στο αποτέλεσμα μιας φυσικής ή εσωτερικής σύζευξης, θα εμφανιστούν ο Βασίλης (Αθήνα) και οι Νίκος, Αγγελική (Πάτρα), επειδή στις πόλεις τους υπάρχουν υποκαταστήματα.

6. Το αποτέλεσμα της εξωτερικής σύζευξης θα περιέχει όλες τις πλειάδες της εσωτερικής σύζευξης, καθώς και επιπλέον τις πλειάδες που πιθανώς δεν έχουν ταιριαστές τιμές. Τέτοιες είναι το υποκατάστημα στη Θεσσαλονίκη και η πελάτης με όνομα Βάσω.

7. Οι πλειάδες της μίας σχέσης, που δεν έχουν ταιριαστές τιμές στην άλλη σχέση, θα συμπληρώσουν με τιμές NULL τα γνωρίσματα των σχέσεων (πεδία στους πίνακες) στο αποτέλεσμα της εξωτερικής σύζευξης. Έτσι η πλειάδα $\langle 3, \text{Θεσσαλονίκη}\rangle$ της σχέσης b (υποκατάστημα) θα υπάρχει στο αποτέλεσμα της εξωτερικής σύζευξης, όμως θα έχει τιμές NULL στα γνωρίσματα της σχέσης c (Πελάτες). Επίσης, η πλειάδα $\langle \text{Βάσω}, \text{Κοζάνη}\rangle$ Βάσω, θα υπάρχει στο αποτέλεσμα της εξωτερικής σύζευξης, αλλά με τιμές NULL στα γνωρίσματα της σχέσης b (Υποκατάστημα).

Αν r είναι σχέση με σχήμα $R = \{X, Y\}$ και s είναι μία σχέση με σχήμα $S = \{Y, Z\}$, τότε η εξωτερική σύζευξη $t = r \mathbin{⟗} s$ έχει σχήμα $T = \{X, Y, Z\}$ και κορμό που αποτελείται από τις πλειάδες:

$$r \mathbin{⟗} s = (r \bowtie s) \cup ((r - \Pi_R(r \bowtie s)) \times w_1) \cup w_2 \times ((s - \Pi_S(r \bowtie s)) \qquad (1.28)$$

όπου w_1 είναι μία σχέση με σχήμα $R - S$ και μία πλειάδα με τιμές $\{null, null, \ldots, null\}$, και w_2 είναι μία σχέση με σχήμα $S - R$ και μία πλειάδα με τιμές $\{null, null, \ldots, null\}$.

1.5.4 Αριστερή εξωτερική σύζευξη

Η αριστερή εξωτερική σύζευξη μοιάζει με την εξωτερική σύζευξη, αλλά έχει μια διαφορά: μόνο οι αταίριαστες πλειάδες της αριστερής σχέσης υπάρχουν στο αποτέλεσμα. Δηλαδή, υπάρχουν στο αποτέλεσμα οι ταιριαστές πλειάδες ανάμεσα στις δύο σχέσεις, αυτές δηλαδή που προέρχονται από τη φυσική σύζευξη, καθώς και όλες οι πλειάδες της αριστερής σχέσης που δεν έχουν τιμή σύζευξης με τη δεξιά σχέση. Τα γνωρίσματα της δεξιάς σχέσης συμπληρώνονται στο αποτέλεσμα με τιμές NULL.

Αν r είναι σχέση με σχήμα $R = \{X, Y\}$ και s είναι μία σχέση με σχήμα $S = \{Y, Z\}$, τότε η αριστερή εξωτερική σύζευξη $t = r \mathbin{⟕} s$ έχει σχήμα $T = \{X, Y, Z\}$ και κορμό που αποτελείται από τις πλειάδες:

$$r \mathbin{⟗} s = (r \bowtie s) \cup ((r - \Pi_R(r \bowtie s)) \times w) \qquad (1.29)$$

όπου w είναι μία σχέση με σχήμα $R - S$ και μία πλειάδα με τιμές $\{null, null, \ldots, null\}$.

Η **αριστερή εξωτερική σύζευξη** (ή απλώς αριστερή σύζευξη):

$$r \mathbin{⟕} s$$

έχει σαν αποτέλεσμα μια σχέση με:

- Σχήμα όμοιο αυτό της φυσικής σύζευξης $r \times s$.
- Κορμό τις πλειάδες που προκύπτουν από την ένωση των πλειάδων:
  - της φυσικής σύζευξης $r \bowtie s$
  - όλων των πλειάδων της r (αριστερής σχέσης) που δεν είναι στο αποτέλεσμα της φυσικής σύζευξης, με NULL τιμές στα γνωρίσματα της s (δεξιάς σχέσης)

1.5.5 Δεξιά εξωτερική σύζευξη

Όπως και η από αριστερά εξωτερική σύζευξη, έτσι και η από δεξιά εξωτερική σύζευξη, μοιάζει με την εξωτερική σύζευξη. Στη περίπτωση αυτή μόνο οι αταίριαστες πλειάδες της από δεξιά σχέσης υπάρχουν στο αποτέλεσμα. Δηλαδή, υπάρχουν στο αποτέλεσμα οι ταιριαστές πλειάδες ανάμεσα στις δύο σχέσεις, αυτές δηλαδή που προέρχονται από τη φυσική σύζευξη, και όλες οι πλειάδες της από δεξιά σχέσης, που δεν έχουν τιμή σύζευξης με τη δεξιά σχέση. Τα γνωρίσματα της δεξιάς σχέσης, συμπληρώνονται στο αποτέλεσμα με τιμές NULL.

Αν r είναι σχέση με σχήμα $R = \{X, Y\}$ και s είναι μία σχέση με σχήμα $S = \{Y, Z\}$, τότε η δεξιά εξωτερική σύζευξη $t = r \mathbin{⟖} s$ έχει σχήμα $T = \{X, Y, Z\}$ και κορμό που αποτελείται από τις πλειάδες:

$$r \mathbin{⟗} s = (r \bowtie s) \cup w \times ((s - \Pi_S(r \bowtie s)) \qquad (1.30)$$

όπου w είναι μία σχέση με σχήμα $S - R$ και μία πλειάδα με τιμές $\{null, null, \ldots, null\}$.

id	city
1	Αθήνα
2	Πάτρα
3	Θεσσαλονίκη

Υποκαταστήματα (b)

name	city
Νίκος	Πάτρα
Βάσω	Κοζάνη
Αγγελική	Πάτρα
Βασίλης	Αθήνα

Πελάτες (c)

id	city	name
1	Αθήνα	Βασίλης
2	Πάτρα	Νίκος
2	Πάτρα	Αγγελική
3	Θεσσαλονίκη	NULL

Αριστερή σύζευξη $b ⟕ c$

id	city	name
1	Αθήνα	Βασίλης
2	Πάτρα	Νίκος
2	Πάτρα	Αγγελική

Φυσική σύζευξη $b \Join c$

Σχήμα 1.37: Αριστερή εξωτερική σύζευξη ανάμεσα στα υποκαταστήματα (b) και τους πελάτες (c) μιας εταιρείας.

Περισσότερο επεξηγηματικά, μπορούμε να πούμε πως η **δεξιά εξωτερική σύζευξη** (ή απλώς δεξιά σύζευξη):

$$r ⟖ s$$

έχει σαν αποτέλεσμα μια σχέση με :

- Σχήμα όμοιο με αυτό της φυσικής σύζευξης $r ⟖ s$.
- Κορμό τις πλειάδες που προκύπτουν από την ένωση των πλειάδων:
  - της φυσικής σύζευξης $r \Join s$
  - όλων των πλειάδων της s (δεξιάς σχέσης) που δεν είναι στο αποτέλεσμα της φυσικής σύζευξης, με NULL τιμές στα γνωρίσματα της r (αριστερής σχέσης)

Η από δεξιά εξωτερική σύζευξη είναι ισοδύναμη με την από αριστερά εξωτερική σύζευξη, αν αλλαχθεί η σειρά εμφάνισης των σχέσεων. Δηλαδή, η ισχύει:

$$r ⟖ s = s ⟕ r$$

Το σχήμα 1.38 δείχνει ένα παράδειγμα δεξιάς εξωτερικής σύζευξης. Για λόγους σύγκρισης παρατίθεται επίσης το αποτέλεσμα της φυσικής σύζευξης των δύο σχέσεων. Παρατηρείστε πως στο αποτέλεσμα της δεξιάς σύζευξης $r ⟖ s$ υπάρχουν όλες οι πλειάδες της φυσικής σύζευξης $r \Join s$. Υπάρχουν επίσης η πλειάδα της s $\langle$Βάσω, Κοζάνη$\rangle$ για την οποία δεν υπάρχει ταιριαστή πλειάδα στην σχέση r. Το γνώρισμα id συμπληρώνεται με τιμή NULL στο αποτέλεσμα της δεξιάς σύζευξης.

1.6 Διαίρεση

Η διαίρεση είναι μια πολύ ενδιαφέρουσα πράξη. Μοιάζει με τη διαίρεση στην άλγεβρα αριθμών, με την έννοια πως έχει διαιρετέο και διαιρέτη. Το αποτέλεσμά της είναι μια σχέση με (γενικά) λιγότερα γνωρίσματα από το διαιρετέο (όσα τα γνωρίσματα του διαιρέτη), και (γενικά) λιγότερες πλειάδες (το

id	city
1	Αθήνα
2	Πάτρα
3	Θεσσαλονίκη

Υποκαταστήματα (b)

name	city
Νίκος	Πάτρα
Βάσω	Κοζάνη
Αγγελική	Πάτρα
Βασίλης	Αθήνα

Πελάτες (c)

id	city	name
1	Βασίλης	Αθήνα
2	Νίκος	Πάτρα
2	Αγγελική	Πάτρα
NULL	Βάσω	Κοζάνη

Δεξιά σύζευξη $b \rtimes c$

id	city	name
1	Αθήνα	Βασίλης
2	Πάτρα	Νίκος
2	Πάτρα	Αγγελική

Φυσική σύζευξη $b \bowtie c$

Σχήμα 1.38: Δεξιά εξωτερική σύζευξη ανάμεσα στα υποκαταστήματα (b) και τους πελάτες (c) μιας εταιρείας.

πολύ όσες έχει ο διαιρέτης). Από τη σχέση διαιρετέου αφαιρούνται εκείνα τα γνωρίσματα που δεν υπάρχουν στο διαιρέτη, και εκείνες οι πλειάδες του διαιρετέου που δεν αντιστοιχούν στο διαιρέτη. Ο διαιρέτης της σχεσιακής πράξης της διαίρεσης με απλά λόγια «οδηγεί» το διαιρετέο σε ένα συνδυασμό προβολής (λιγότερα γνωρίσματα) και περιορισμού (λιγότερες πλειάδες).

Το πλέον συνηθισμένο πεδίο εφαρμογής της διαίρεσης είναι σε πράξεις αναζήτησης δεδομένων με τη φράση «όλα», με αναφορά σε δύο σχέσεις (πίνακες). Οι δύο σχέσεις (διαιρετέος και διαιρέτης) έχουν τουλάχιστον ένα κοινό γνώρισμα, το οποίο δεν εμφανίζεται στο αποτέλεσμα.

Ορισμός της διαίρεσης:
Αν η r είναι σχέση με σχήμα $R = \{X, Y\}$ και s είναι μια σχέση με σχήμα $S = \{Y\}$, δηλαδή $S \subseteq R$, τότε, το αποτέλεσμα της διαίρεσης $t = r \div s$, είναι μια σχέση t με σχήμα $\{X\}$. Δηλαδή, η t είναι μια σχέση με σχήμα τη διαφορά $R - S = \{X\}$, δηλαδή εκείνα τα γνωρίσματα της r που δεν ανήκουν στην s.
Ο κορμός της t αποτελείται από εκείνες τις πλειάδες της σχέσης r για οποίες τα (κοινά) γνωρίσματα $\{X\}$ των σχέσεων r και s έχουν τιμές που ταυτίζονται. Μια πλειάδα t, με γνωρίσματα $R - S = \{X\}$, ανήκει στο αποτέλεσμα όταν:

1. $t \in \Pi_{R-S}(r)$
2. $\forall\, t_s \in s, \quad \exists\, t_r \in r$ έτσι ώστε:
   - $t_r[S] = t_s[S]$
   - $t_r[R - S] = t$

Δείτε το παράδειγμα του σχήματος 1.39, που αναπαριστά τη διαίρεση: $\Sigma = A \div \Pi$. Σε μια σχολική τάξη, διάφοροι μαθητές παίρνουν μέρος σε αθλήματα. Στο επόμενο σχολικό πρωτάθλημα, περιλαμβάνονται τα αθλήματα ποδοσφαίρου και στίβου. Η διαίρεση: *Αθλούνται προς Πρωτάθλημα*, βρίσκει εκείνους τους μαθητές που θα πάρουν μέρος σε όλα τα αγωνίσματα του πρωταθλήματος (και στο στίβο και στο ποδόσφαιρο). Ας εξετάσουμε το αποτέλεσμα, βήμα προς βήμα:

- Διαιρετέος είναι η σχέση *Αθλούνται* και διαιρέτης η σχέση *Πρωτάθλημα*). Το αποτέλεσμα είναι η σχέση *Συμμετοχή*. Δηλαδή:

Όνομα	*Άθλημα*
Νίκος	Ποδόσφαιρο
Ανδρέας	Κολύμπι
Βασίλης	Στίβος
Ανδρέας	Ποδόσφαιρο
Χρήστος	Κολύμπι
Βασίλης	Ποδόσφαιρο
Νίκος	Στίβος
Γιώργος	Ποδόσφαιρο
Χρήστος	Στίβος

(α΄) Αθλούνται

Άθλημα
Ποδόσφαιρο
Στίβος

(β΄) Πρωτάθλημα

Όνομα
Νίκος
Βασίλης

(γ΄) Συμμετοχή

Σχήμα 1.39: Παράδειγμα διαίρεσης. Η σχέση Συμμετοχή προκύπτει από τη διαίρεση της σχέσης Αθλούνται με τη σχέση Πρωτάθλημα.

$$\text{Συμμετοχή} = \text{Αθλούνται} \div \text{Πρωτάθλημα}$$

- Το αποτέλεσμα έχει σχήμα τη διαφορά των σχημάτων των δύο σχέσεων. Η σχέση *Συμμετοχή* έχει στο σχήμα της εκείνα τα γνωρίσματα της σχέσης *Αθλούνται*, που δεν υπάρχουν στη σχέση *Πρωτάθλημα*, δηλαδή αφαιρείται το γνώρισμα *Άθλημα* και στο αποτέλεσμα μένει μόνο το γνώρισμα *Όνομα*.

- Για το κορμό της σχέσης *Συμμετοχή*, πρέπει να βρεθούν εκείνες οι πλειάδες της σχέσης *Αθλούνται*, για τις οποίες το *Άθλημα* υπάρχει στις πλειάδες που αντιστοιχούν σε όλες τις πλειάδες της σχέση *Πρωτάθλημα*. Δηλαδή, θα ληφθούν τα ονόματα των μαθητών που αθλούνται σε όλα τα αθλήματα τα οποία υπάρχουν στη λίστα αθλημάτων για το πρωτάθλημα.

- Ακόμα πιο απλά, θέλουμε εκείνες τις πλειάδες της σχέσης $\Pi_{\text{Όνομα}}(\text{Αθλούνται})$ για τις οποίες υπάρχει πλήρης αντιστοίχηση, μέσω της σχέσης $\Pi_{\text{Άθλημα}}(\text{Αθλούνται})$, με τις πλειάδες της σχέσης $\Pi_{\text{Άθλημα}}(\text{Πρωτάθλημα})$: τα ονόματα των μαθητών που συμμετέχουν σε όλα τα αθλήματα.

- Ο Νίκος συμμετέχει στο πρωτάθλημα, γιατί αγωνίζεται και στο ποδόσφαιρο στο στίβο, στα δύο αθλήματα του πρωταθλήματος.

- Ο Ανδρέας συμμετέχει στο πρωτάθλημα, γιατί δεν αγωνίζεται στο ποδόσφαιρο, το οποίο είναι στο πρωτάθλημα.

- Ο Βασίλης συμμετέχει στο πρωτάθλημα, γιατί αγωνίζεται και στο ποδόσφαιρο στο στίβο, στα δύο αθλήματα του πρωταθλήματος.

- Ο Χρήστος δεν συμμετέχει στο πρωτάθλημα, γιατί δεν αγωνίζεται στο στίβο, το οποίο είναι στο πρωτάθλημα.

- Ο Γιώργος δεν συμμετέχει στο πρωτάθλημα, γιατί δεν αγωνίζεται στο στίβο, το οποίο είναι στο πρωτάθλημα.

Μπορούμε να κάνουμε κάποιες επιπλέον παρατηρήσεις. Ξεκινήσαμε με το ερώτημα «Να βρεθούν όλα τα ονόματα που θα λάβουν μέρος στο πρωτάθλημα». Δεδομένα είναι δύο σχέσεις που περιέχουν τα αθλήματα του πρωταθλήματος (διαιρέτης) και το άθλημα κάθε μαθητή (διαιρετέος). Δηλαδή έχουμε δύο σχέσεις με ένα κοινό γνώρισμα. Το αποτέλεσμα είναι μια σχέση (*Συμμετοχή*) που προέρχεται από το διαιρετέο (*Αθλούνται*) χωρίς το κοινό γνώρισμα *Άθλημα*, δηλαδή το κοινό γνώρισμα διαιρετέου και διαιρέτη. Από τις πλειάδες της σχέσης *Αθλούνται* μένουν στο αποτέλεσμα μόνο εκείνες για τις οποίες η αντίστοιχη τιμή στο γνώρισμα *Άθλημα* ταυτίζεται με όλες τις τιμές που υπάρχουν στη σχέση *Πρωτάθλημα*. Ο Χρήστος, για παράδειγμα, που αθλείται μόνο στο κολύμπι, δεν συμμετέχει στο πρωτάθλημα. Ο Νίκος συμμετέχει στο πρωτάθλημα, γιατί αγωνίζεται σε όλα τα αθλήματα του πρωταθλήματος. Ο Ανδρέας που αγωνίζεται μόνο στο ποδόσφαιρο, αλλά όχι στο στίβο, δεν συμμετέχει στο πρωτάθλημα. Παρατηρήστε επίσης ότι, ενώ κάποιοι αθλούνται σε δύο αθλήματα, κανείς δεν έχει διπλή πλειάδα στο πρωτάθλημα.

Η διαίρεση είναι το αντίστροφο του γινομένου στην άλγεβρα των αριθμών, και το ίδιο μπορεί να ειπωθεί για τη σχεσιακή άλγεβρα. Έτσι, αν $t = r \div s$, τότε η σχέση $t \times s$, έχει συμβατότητα τύπου με την r.

Η διαίρεση είναι παράγωγη πράξη, και μπορεί να οριστεί με βάση άλλες σχεσιακές πράξεις. Ένας εναλλακτικός τρόπος να ορίσουμε τη διαίρεση δύο σχέσεων r και s με σχήματα R και S αντίστοιχα, όταν $S \subset R$, είναι:

Ορισμός της διαίρεσης:

$$r \div s = \Pi_{R-S}(r) - \Pi_{R-S}((\Pi_{R-S}(r) \times s) - \Pi_{R-S,S}(r)) \tag{1.31}$$

Αν πάρουμε ως $R = \{\text{Όνομα, Άθλημα}\}$ και $S = \{\text{Άθλημα}\}$ των σχέσεων του σχήματος 1.39, δηλαδή $r =$ Αθλούνται και $s =$ Πρωτάθλημα τότε ο παραπάνω ορισμός μπορεί να αναλυθεί ως εξής:

$$\begin{aligned} \text{Αθλούνται} \div \text{Πρωτάθλημα} = \Pi_{\text{Όνομα}}(\text{Αθλούνται}) \\ -\Pi_{\text{Όνομα}}((\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Πρωτάθλημα}) - \Pi_{\text{Όνομα, Άθλημα}}(\text{Αθλούνται})) \end{aligned}$$

Ας δούμε αναλυτικά πως αναλύεται η παραπάνω παράσταση:

1. Παίρνουμε την προβολή $\Pi_{R-S}(r)$, δηλαδή:

   $$\Pi_{\{\text{Όνομα, Άθλημα}\}-\{\text{Άθλημα}\}}(\text{Αθλούνται}),$$

   που αντιστοιχεί στην προβολή των ονομάτων του πίνακα *Αθλούνται* του σχήματος 1.39. Ισχύει:

   $$R - S = \{\text{Όνομα, Πρωτάθλημα}\} - \{\text{Πρωτάθλημα}\} = \{\text{Όνομα}\}$$

   Η προβολή $\Pi_{R-S}(r) = \Pi_{\{\text{Όνομα}\}}(\text{Πρωτάθλημα})$ θα επιστρέψει:

Όνομα
Νίκος
Ανδρέας
Βασίλης
Χρήστος
Γιώργος

   Δηλαδή τα ονόματα (με απαλοιφή διπλοεγγραφών) των αθλητών.

2. Από το προηγούμενο αποτέλεσμα πρέπει να αφαιρέσουμε τις πλειάδες:

$$\Pi_{R-S}((\Pi_{R-S}(r) \times s) - \Pi_{R-S,S}(r))$$

ή:

$$\Pi_{\text{Όνομα}}((\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Πρωτάθλημα}) - \Pi_{\text{Όνομα, Άθλημα}}(\text{Αθλούνται})$$

3. Το αποτέλεσμα $\Pi_{R-S}(r) \times s$ ή:

$$\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Πρωτάθλημα}$$

είναι ίσο με:

Όνομα
Νίκος
Ανδρέας
Βασίλης
Χρήστος
Γιώργος

$\Pi_{\text{Όνομα}}(\text{Αθλούνται})$

Άθλημα
Ποδόσφαιρο
Στίβος

Πρωτάθλημα

Όνομα	Άθλημα
Νίκος	Ποδόσφαιρο
Ανδρέας	Ποδόσφαιρο
Βασίλης	Ποδόσφαιρο
Χρήστος	Ποδόσφαιρο
Γιώργος	Ποδόσφαιρο
Νίκος	Στίβος
Ανδρέας	Στίβος
Βασίλης	Στίβος
Χρήστος	Στίβος
Γιώργος	Στίβος

$\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Πρωτάθλημα}$

4. Η σχέση $\Pi_{R-S,S}(r)$ είναι ίδια με την r. Αυτό που γίνεται είναι απλά μια πιθανή αναδιάταξη των γνωρισμάτων της r έχει ώστε να έχει συμβατότητα τύπου με το γινόμενο $\Pi_{R-S}(r) \times s$. Δηλαδή, τοποθετούνται πρώτα τα γνωρίσματα της r που δεν ανήκουν στην s $(R-S)$ και μετά τα κοινά γνωρίσματα των r και s. Στην περίπτωσή μας αυτό είναι:

$$\Pi_{\text{Όνομα, Άθλημα}}(\text{Αθλούνται})$$

Στη συγκεκριμένη περίπτωση η αναδιάταξη είναι περιττή. Το αποτέλεσμα είναι το ίδιο με αυτό του σχήματος 1.39:

Όνομα	Άθλημα
Νίκος	Ποδόσφαιρο
Ανδρέας	Κολύμπι
Βασίλης	Στίβος
Ανδρέας	Ποδόσφαιρο
Χρήστος	Κολύμπι
Βασίλης	Ποδόσφαιρο
Νίκος	Στίβος
Γιώργος	Ποδόσφαιρο
Χρήστος	Στίβος

Ωστόσο, η προβολή $\Pi_{R-S,S}(r)$ πρέπει να γίνεται πάντοτε, ώστε να εξασφαλίζεται σίγουρα η συμβατότητα τύπου για την επόμενη πράξη της αφαίρεσης.

5. Μπορούμε τώρα να κάνουμε την αφαίρεση $(\Pi_{R-S}(r) \times s) - \Pi_{R-S,S}(r)$ που στην περίπτωσή μας αντιστοιχεί στην αφαίρεση:

$$(\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Άθλημα}) - \Pi_{\text{Όνομα, Άθλημα}}(\text{Αθλούνται})$$

Το αποτέλεσμα είναι:

Όνομα	Άθλημα
Νίκος	Ποδόσφαιρο
Ανδρέας	Ποδόσφαιρο
Βασίλης	Ποδόσφαιρο
Χρήστος	Ποδόσφαιρο
Γιώργος	Ποδόσφαιρο
Νίκος	Στίβος
Ανδρέας	Στίβος
Βασίλης	Στίβος
Χρήστος	Στίβος
Γιώργος	Στίβος

$(\Pi_{R-S}(r) \times s)$

Όνομα	Άθλημα
Νίκος	Ποδόσφαιρο
Ανδρέας	Κολύμπι
Βασίλης	Στίβος
Ανδρέας	Ποδόσφαιρο
Χρήστος	Κολύμπι
Βασίλης	Ποδόσφαιρο
Νίκος	Στίβος
Γιώργος	Ποδόσφαιρο
Χρήστος	Στίβος

$\Pi_{R-S,S}(r)$

Όνομα	Άθλημα
Χρήστος	Ποδόσφαιρο
Ανδρέας	Στίβος
Γιώργος	Στίβος

$(\Pi_{R-S}(r) \times s) - \Pi_{R-S,S}(r)$

6. Η προβολή $\Pi_{R-S}((\Pi_{R-S}(r) \times s) - \Pi_{R-S,S}(r))$ που στην περίπτωσή μας είναι:

$$\Pi_{\text{Όνομα}}((\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Πρωτάθλημα}) - \Pi_{\text{Όνομα, Άθλημα}}(\text{Αθλούνται}))$$

θα δώσει το αποτέλεσμα:

Όνομα
Χρήστος
Ανδρέας
Γιώργος

και στην ουσία πρόκειται για τους μαθητές που δεν συμμετέχουν στο πρωτάθλημα.

7. Το τελικό αποτέλεσμα βρίσκεται από τη διαφορά:

$$\begin{aligned} \text{Αθλούνται} \div \text{Πρωτάθλημα} = \Pi_{\text{Όνομα}}(\text{Αθλούνται}) \\ -\Pi_{\text{Όνομα}}((\Pi_{\text{Όνομα}}(\text{Αθλούνται}) \times \text{Πρωτάθλημα}) - \Pi_{\text{Όνομα, Άθλημα}}(\text{Αθλούνται})) \end{aligned}$$

Όνομα
Νίκος
Ανδρέας
Βασίλης
Χρήστος
Γιώργος

Όνομα
Χρήστος
Ανδρέας
Γιώργος

Όνομα
Νίκος
Βασίλης

Αθλούνται ÷ Πρωτάθλημα

1.7 Σύνοψη και συναρτήσεις συνάθροισης

Η σύνοψη (ή αλλιώς ομαδοποίηση) ομαδοποιεί πλειάδες μιας σχέσης με βάση κοινές τιμές σε ένα ή περισσότερα γνωρίσματα. Σε κάθε ένα από τα υποσύνολα των πλειάδων που προκύπτουν μπορεί να εφαρμοστεί μια συναθροιστική συνάρτηση (αγγρεγατινγ φυνςτιον), όπως η καταμέτρηση του πλήθους ή ο υπολογισμός του μέσου όρου. Το σημαντικό, σε σχέση με άλλες σχεσιακές πράξεις, είναι ότι από πολλές πλειάδες μιας σχέσεις λαμβάνονται λιγότερες στο αποτέλεσμα, όχι εξαιτίας της εφαρμογής κάποιας συνθήκης, αλλά λόγω ομαδοποίησης.

area	employee	amount
Ημαθίας	Ευθυμίου	10890
Μαγνησίας	Αλεξανρίδης	2400
Καβάλας	Αλεξανρίδης	780
Μαγνησίας	Ευθυμίου	20100
Τρικάλων	Πετρίδης	4400
Πιερίας	Πετρίδης	1820
Καβάλας	Ευθυμίου	2400

Σχήμα 1.40: Πινακοποιημένη μορφή της σχέσης *sales* με δείγμα δεδομένων

Θεωρείστε για παράδειγμα τη σχέση $sales = \{area, employee, amount\}$ που κρατά στοιχεία για το ύψος πωλήσεων που πετυχαίνουν εμπορικοί αντιπρόσωποι μιας εταιρείας. Δείγμα των δεδομένων της σχέσης φαίνεται στο σχήμα 1.40. Εύκολα μπορεί να εξαχθεί το συμπέρασμα πως (τουλάχιστον στο δείγμα) υπάρχουν 4 υπάλληλοι που δραστηριοποιούνται σε 5 νομούς.

1. Ποιο είναι το σύνολο (άθροισμα) των πωλήσεων σε όλους τους νομούς;
2. Ποιο είναι το άθροισμα των πωλήσεων ανά εμπορικό αντιπρόσωπο (υπάλληλο) της εταιρείας;
3. Επίσης, ποιο το πλήθος υπαλλήλων ανά νομό;

Τέτοιες, και άλλες πολλές (αλλά παρόμοιες) ερωτήσεις μπορούν να απαντήσουν τα ερωτήματα με τις συναρτήσεις συνάθροισης, με ή χωρίς ομαδοποίηση.

Υπάρχουν 5 κύριες συναρτήσεις συνάθροισης:

- COUNT(), επιστρέφει πλήθος εγγραφών.
- SUM(), υπολογίζει και επιστρέφει το άθροισμα αριθμητικού γνωρίσματος,

- AVG(), υπολογίζει και επιστρέφει το μέσο όρο αριθμητικού γνωρίσματος,
- MIN(), επιστρέφει τη μικρότερη τιμή.
- MAX(), επιστρέφει τη μεγαλύτερη τιμή.

Για παράδειγμα, το άθροισμα όλων των πωλήσεων είναι:

$$\mathscr{G}_{\mathbf{sum}(amount)}(sales)$$

Στην SQL γράφεται:

```
  SELECT SUM(amount)
    FROM sales
```

Η τιμή που λαμβάνεται από το προηγούμενο ερώτημα είναι μία, και αντιστοιχεί στην απλή άθροιση των τιμών του γνωρίσματος *amount*. Η χρήση της συναθροιστικής συνάρτησης SUM(), σε συνδυασμό με την ομαδοποίηση ως προς κάποιο γνώρισμα, πχ το *employees*, μπορεί να δώσει απάντηση σε περισσότερο πολύπλοκα ερωτήματα, όπως το να βρεθεί το σύνολο των πωλήσεων ανά υπάλληλο:

$$_{employee}\mathscr{G}_{\mathbf{sum}(amount)}(sales)$$

Ή στην SQL:

```
  SELECT employee, SUM(amount)
    FROM sales
GROUP BY employee;
```

Η παραπάνω σχεσιακή παράσταση μπορεί να αναλυθεί ως εξής:

1. Λαμβάνεται ο κορμός της σχέσης *sales*, δηλαδή όλες οι πλειάδες.

2. Ο αριστερός δείκτης *employee* από το καλλιγραφικό G, σημαίνει πως οι πλειάδες πρέπει να ομαδοποιηθούν ως προς την κοινή τιμή σε αυτό το γνώρισμα

3. Υπάρχουν 4 υπάλληλοι-πωλητές. Επομένως, ο κορμός της *sales* διαιρείται σε 4 υποσύνολα. Καθένα από αυτά έχει την ίδια τιμή στο γνώρισμα *employee*. Δηλαδή, ένα υποσύνολο πλειάδων για κάθε ένα υπάλληλο. Αυτό ακριβώς σημαίνει η φράση «ανά υπάλληλο»

4. Ο δεξιός δείκτης *sum(amount)* από το $\mathscr{G}$ (καλλιγραφικό G), σημαίνει πως στο σύνολο των πλειάδων της *sales* θα υπολογιστεί το άθροισμα. Αφού έχει προηγηθεί διάσπαση σε 4 υποσύνολα, το άθροισμα θα υπολογιστεί μία φορά για κάθε ένα υποσύνολο ξεχωριστά. Δηλαδή θα υπολογιστούν 4 υπο-αθροίσματα: ένα για κάθε ξεχωριστό υπάλληλο.

5. Το αποτέλεσμα της σχεσιακής παράστασης έχει σχήμα το ({*employee, sum(amount)* }) και κορμό τις 4 πλειάδες που αντιστοιχούν στα ονόματα των 4 υπαλλήλων και των 4 αθροισμάτων στο ποσό των πωλήσεων, όπως φαίνεται στο σχήμα ;;.

employee	sum(amount)
Ευθυμίου	33390
Αλεξανρίδης	3180
Πετρίδης	6220
Σταθόπουλος	19500

Σχήμα 1.41: Άθροισμα πωλήσεων ανά υπάλληλο

1.8 Ενημέρωση

Εκτός από τις πράξεις επιλογής, η σχεσιακή άλγεβρα έχει ανάγκη από τις πράξεις ενημέρωσης της βάσης δεδομένων. Με αυτές υπάρχει η δυνατότητα:

- **Εισαγωγής** δεδομένων στις σχέσεις, δηλαδή εισαγωγής μιας ή περισσοτέρων πλειάδων. Η πληθικότητα της σχέσης, δηλαδή το πλήθος των πλειάδων της, αυξάνεται μετά την εισαγωγή.

- **Διαγραφής** δεδομένων από τις σχέσεις, δηλαδή απαλοιφή μιας ή περισσότερων πλειάδων της σχέσης. Η πληθικότητα της σχέσης μειώνεται μετά την διαγραφή.
- **Τροποποίησης** δεδομένων στις σχέσεις, δηλαδή αλλαγή στις τιμές των γνωρισμάτων μιας σχέσης. Η πληθικότητα της σχέσης δεν μεταβάλλεται μετά την τροποποίηση.

Να έχετε πάντα υπ΄ όψη σας πως οι πράξεις ενημέρωσης είναι πράξεις συνόλων. Αυτό το οποίο εισάγεται, διαγράφεται ή τροποποιείται είναι μια πλειάδα, δηλαδή ένα σύνολο από τιμές που αντιστοιχούν στα γνωρίσματα της σχέσης.

Κωδικός	Τίτλος	Έτος
0083658	Blade Runner	1982
0034583	Casablanca	1942
0053779	La Dolce Vita	1960
0087884	Paris Texas	1984

Σχήμα 1.42: Πινακοποιημένη μορφή της σχέσης *movies* με δείγμα δεδομένων.

Δείτε για παράδειγμα το σχήμα 1.42, το οποίο παρουσιάζει ένα δείγμα δεδομένων από τη σχέση *movies* = {Κωδικός, Τίτλος, Έτος}. Με άλλα λόγια, έχουμε μια σχέση όπου αποθηκεύονται στοιχεία για τον τίτλο και το έτος παραγωγής ταινιών. Μια συγκεκριμένη ταινία, σύμφωνα με το σχεσιακό μοντέλο δεδομένων της σχέσης *movies*, είναι ένα σύνολο από τρεις τιμές, που αντιστοιχούν στα τρία γνωρίσματα της σχέσης. Έτσι κάθε ταινία, δηλαδή κάθε πλειάδα που αντιστοιχεί σε μία ταινία, είναι ένα σύνολο τιμών. Η σχέση *movies* έχει ως σώμα ένα σύνολο από πλειάδες: όλες τις ταινίες που περιέχει. Δηλαδή κάθε ταινία είναι ένα σύνολο τιμών, αλλά είναι και μέλος ενός συνόλου: είναι μέλος όλων των πλειάδων της σχέσης *movies*.

Κάθε πράξη ενημέρωσης της βάσης (εισαγωγή, διαγραφή, τροποποίηση) αφορά πλειάδες μιας σχέσης και όχι μεμονωμένες τιμές κάποιου γνωρίσματος. Επίσης οι πράξεις ενημέρωσης δεν μεταβάλλουν το σχεσιακό σχήμα. Δεν προστίθενται, αφαιρούνται ή τροποποιούνται γνωρίσματα κάποιας σχέσης, αλλά πλειάδες της σχέσης. Δηλαδή οι πράξεις ενημέρωσης αφορούν το σώμα μιας σχέσης, εδώ είναι που γίνεται η ενημέρωση.

Ας δούμε τις τρεις πράξεις ενημέρωσης αναλυτικά.

1.8.1 Εισαγωγή

Η εισαγωγή δεδομένων στη βάση σημαίνει την εισαγωγή μιας πλειάδας σε κάποια σχέση. Θυμηθείτε ότι η πλειάδα είναι ένα σύνολο τιμών που αντιστοιχούν στα γνωρίσματα μιας σχέσης. Επομένως υπάρχει πάντα ο περιορισμός της σωστής αντιστοίχισης, καθώς επίσης και της σχεσιακής ακεραιότητας δεδομένων: κάθε τιμή της πλειάδας πρέπει να ανήκει στο πεδίο ορισμού του γνωρίσματος. Η πλειάδα προς εισαγωγή μπορεί να προκύψει είτε από νέα δεδομένα που εισάγει ο χρήστης της βάσης, είτε από το αποτέλεσμα ενός ερωτήματος.

Γενικά, η εισαγωγή των E (σχεσιακής έκφρασης) στη σχέση r, γράφετε ως:

$$r \leftarrow r \cup E$$

Για παράδειγμα η εισαγωγή της ταινίας *Blade Runner* του 1982 με κωδικό 0083658 γίνεται με σχεσιακή πράξη:

$$movies \leftarrow movies \cup \{0083658, \text{'Blade Runner'}, 1982\}$$

Που σημαίνει πως η σχέση *movies* αποκτά μια νέα πλειάδα και η πληθικότητά της αυξάνεται κατά μία μονάδα. Σε περίπτωση που υπάρξει κάποιο λάθος, όπως για παράδειγμα παραβίαση του κανόνα πρωτεύοντος κλειδιού, τότε η εισαγωγή αποτυγχάνει **συνολικά**. Αυτό σημαίνει πως αποτυγχάνει η εισαγωγή μιας πλειάδας στη σχέση, ή μιας εγγραφής στον πίνακα, και όχι η εισαγωγή κάποιας τιμής ενός γνωρίσματος.

Στην SQL η εισαγωγή δεδομένων γίνεται με την εντολή INSERT. Για παράδειγμα:

```
INSERT
  INTO movies (code, title, year)
VALUES (0083658, 'Blade Runner', 1982);
```

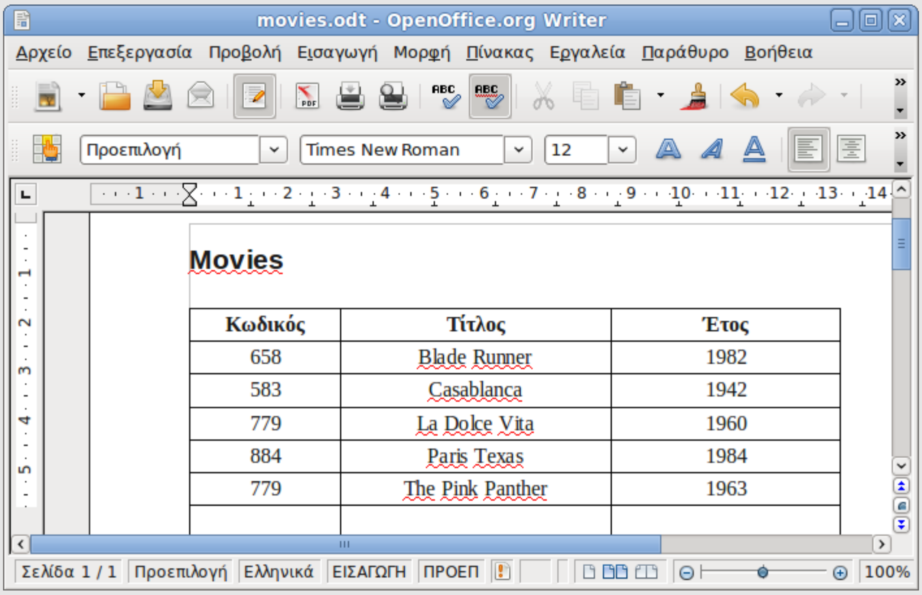
Κωδικός	Τίτλος	Έτος
0083658	Blade Runner	1982
0034583	Casablanca	1942
0053779	La Dolce Vita	1960
0087884	Paris Texas	1984
	The Pink Panther	1960

Σχήμα 1.43: Πινακοποιημένη μορφή της σχέσης *movies* με δείγμα δεδομένων, όπως θα ήταν μετά την προσπάθεια εισαγωγής δεδομένων που παραβιάζουν κάποιο κανόνα ακεραιότητας δεδομένων, εδώ τον κανόνα πρωτεύοντος κλειδιού.

Για παράδειγμα θεωρείστε την πλειάδα $\langle 0053779, \text{'The Pink Panther'}, 1963 \rangle$ που αντιστοιχεί στην εισαγωγή της διάσημης (και πολύ αγαπημένης) κωμωδίας στη σχέση *movies*. Η απόπειρα εκτέλεσης της πράξης:

$$movies \leftarrow movies \cup \{053779, \text{'The Pink Panther'}, 1963\}$$

θα αποτύχει γιατί η τιμή 053779 που αντιστοιχεί στο πρωτεύον κλειδί, υπάρχει ήδη σε άλλη πλειάδα της σχέσης *movies*. Αυτό που θα προκύψει είναι ένα μήνυμα λάθους από το Σύστημα Διαχείρισης Βάσεων Δεδομένων, και όχι μια σχέση στο σώμα της οποίας υπάρχει μια πλειάδα χωρίς πρωτεύον κλειδί. Το σχήμα 1.43 δεν θα προκύψει ως πίνακας μιας σχεσιακής βάσης ποτέ, και αναπαριστά με εσφαλμένο τρόπο το αποτέλεσμα της εισαγωγής. Έτσι λοιπόν, μετά από μια αποτυχημένη εισαγωγή πλειάδων σε σχέση, δεν μεταβάλλεται η πληθικότητά της και η σχέση παραμένει όπως και πριν.



Movies

Κωδικός	Τίτλος	Έτος
658	Blade Runner	1982
583	Casablanca	1942
779	La Dolce Vita	1960
884	Paris Texas	1984
779	The Pink Panther	1963

Σχήμα 1.44: Οι ταινίες σε έναν επεξεργαστή κειμένου. Καταχρηστικά, είναι δυνατή η παρουσία διπλότυπων τιμών στο κλειδί, και η γενικότερη εισαγωγή λανθασμένων δεδομένων. Οι επεξεργαστές κειμένου σε καμία περίπτωση δεν είναι σχεσιακές βάσεις, με όλα τα συνεπακόλουθα!

Αυτό είναι ένα πολύ συχνό λάθος των αρχάριων χρηστών των βάσεων δεδομένων: η εισαγωγή δεδομένων δεν γίνεται ποτέ τιμή προς τιμή, όπως για παράδειγμα σε ένα λογιστικό φύλλο ή σε έναν επεξεργαστή κειμένου, όπως φαίνεται στο σχήμα 1.44. Οι βάσεις δεδομένων λειτουργούν με διαφορετικό τρόπο. Όπως φαίνεται στο σχήμα 1.44, σε αυτά τα προγράμματα υπάρχει η δυνατότητα καταχώρισης δεδομένων χωρίς τους ελέγχους του σχεσιακού μοντέλου. Πιθανότατα, όσοι από εσάς έχετε δουλέψει σε αντίστοιχα περιβάλλοντα επεξεργασίας κειμένου, θα θυμάστε επίσης τη διαδικασία αποθήκευσης των αλλαγών που κάνετε στο κείμενο, π.χ. με την εισαγωγή μίας επιπλέον ταινίας. Στις βάσεις δεδομένων τίποτα τέτοιο δεν ισχύει. Αν μια πλειάδα εισαχθεί με επιτυχία, τότε δεν χρειάζεται καμία επιπλέον ενέργεια, η έννοια της αποθήκευσης δεν υπάρχει! Αν αποτύχει η εισαγωγή της πλειάδας, τότε καμία από τις τιμές δεν εισέρχεται στη βάση. Είτε θα εισαχθεί το σύνολο των τιμών της πλειάδας, είτε τίποτα.

1.8.2 Διαγραφή

Η διαγραφή δεδομένων από τη βάση σημαίνει διαγραφή μιας (ή περισσότερων) πλειάδας σε κάποια σχέση. Όπως έχει ειπωθεί και στην προηγούμενη ενότητα για την Εισαγωγή δεδομένων, η πλειάδα είναι σύνολο τιμών που αντιστοιχούν στα γνωρίσματα μιας σχέσης. Αυτό που διαγράφεται είναι το σύνολο των τιμών που απαρτίζουν την πλειάδα, και όχι κάποια τιμή από την πλειάδα.

Γενικά, η διαγραφή των δεδομένων που προκύπτουν από μια σχεσιακή παράσταση E στη σχέση r, γράφετε ως:

$$r \leftarrow r - E$$

Εδώ εννοείται πως η σχεσιακή παράσταση E έχει συμβατότητα τύπου με τη σχέση r. Η διαγραφή δεδομένων, όπως είναι φανερό, εμπλέκει τη σχεσιακή πράξη της αφαίρεσης.

Για παράδειγμα η διαγραφή της ταινίας *Blade Runner* του 1982 με κωδικό 0083658 γίνεται με σχεσιακή πράξη:

$$movies \leftarrow movies - \sigma_{\text{Κωδικός}=0083658}(movies)$$

Η SQL χρησιμοποιεί την εντολή DELETE για να διαγράψει εγγραφές από πίνακες:

```
DELETE
  FROM movies;
```

Αυτό χρειάζεται προσοχή, επειδή η παραπάνω εντολή διαγράφει όλες τις εγγραφές του πίνακα! Η παρακάτω εντολή διαγράφει την ταινία *Blade Runner* του 1982 με κωδικό 0083658 αφήνοντας ανέπαφες τις υπόλοιπες:

```
DELETE
  FROM movies
 WHERE code = 0083658;
```

Όπως και σε άλλες σχεσιακές πράξεις, η παράσταση E δεν είναι υποχρεωτικό να βασίζεται στο πρωτεύον κλειδί. Για παράδειγμα, να διαγραφούν οι ταινίες με έτος παραγωγής μέχρι και το 1970:

$$movies \leftarrow movies - \sigma_{\text{Ετος} \leq 1970}(movies)$$

Ή με την SQL:

```
DELETE
  FROM movies
 WHERE year <= 1970;
```

Η επιτυχής διαγραφή πλειάδων από τη σχέση έχει ως αποτέλεσμα τη μείωση της πληθικότητάς της. Σε περίπτωση που σχεσιακή παράσταση E επιστρέψει κενό σύνολο, τότε η εφαρμογή της διαγραφής δεν επηρεάζει καμία από τις πλειάδες της σχέσης. Αυτό σημαίνει πως η πληθικότητα της σχέσης θα μείνει αμετάβλητη.

1.8.3 Τροποποίηση

Η τροποποίηση (UPDATE) αφορά την αλλαγή κάποιων τιμών μιας πλειάδας (ή και περισσότερων πλειάδων). Μετά από κάθε πράξη τροποποίησης η πληθικότητα της σχέσης μένει αμετάβλητη: η τροποποίηση ούτε αυξάνει, ούτε μειώνει τις πλειάδες μιας σχέσης. Απλά αλλάζει (τροποποιεί) κάποιες τιμές σε κάποια από τα γνωρίσματα μιας σχέσης.

Για την τροποποίηση χρησιμοποιείται ο τελεστής γενικευμένης προβολής:

$$r \leftarrow \Pi_{F_1, F_2, \ldots, F_n}(r)$$

όπου F_i, είναι είτε η τιμή ενός γνωρίσματος, αν δεν είναι στη λίστα τροποποίησης ή μια σχεσιακή παράσταση που τροποποιεί τα δεδομένα της σχέσης.

Αυτό βέβαια θα αλλάζει τα δεδομένα σε όλες τις πλειάδες μιας σχέσης, κάτι που γενικά δεν είναι επιθυμητό. Είναι εξαιρετικά σπάνιο να επιβληθεί η ίδια τιμή σε κάποιο γνώρισμα, σε όλες τις πλειάδες μιας σχέσης. Αν γνωρίζετε λεπτομέρειες από την ιστορία του σινεμά, θα έχετε παρατηρήσει πως το σχήμα 1.42 εσφαλμένα δείχνει ως έτος παραγωγής της ταινίας *Casablanca* το 1952 (το σωστό είναι 1942). Για να αλλάξει μια (εδώ λανθασμένη) τιμή σε κάποια άλλη χρειάζεται η πράξη της τροποποίησης. Για παράδειγμα, η τροποποίηση:

$$movies \leftarrow \Pi_{\text{Κωδικός, Τίτλος, Έτος}=1942}(movies)$$

θα αλλάξει την χρονολογία παραγωγής στην τιμή 1942 σε όλες τις πλειάδες, δηλαδή για όλες τις ταινίες. Στην SQL η τροποποίηση δεδομένων σε πίνακες της βάσης, γίνεται με την εντολή UPDATE:

```
UPDATE movies
   SET year = 1942;
```

Η εφαρμογή αυτής της εντολής (ή της αντίστοιχης σχεσιακής πράξης) έχει ως αποτέλεσμα την ορθή τροποποίηση της χρονολογίας παραγωγής της ταινίας *Casablanca*. Ωστόσο, έχει και ένα σοβαρό μειονέκτημα: τη λανθασμένη τροποποίηση όλων των υπόλοιπων χρονολογιών, όπως φαίνεται στη σχήμα 1.45.

Κωδικός	Τίτλος	Έτος
0083658	Blade Runner	1942
0034583	Casablanca	1942
0053779	La Dolce Vita	1942
0087884	Paris Texas	1942

Σχήμα 1.45: Πινακοποιημένη μορφή της σχέσης *movies*, μετά την εφαρμογή της καθολικής τροποποίησης της χρονολογίας.

Συνήθως η τροποποίηση αφορά ένα υποσύνολο των πλειάδων μιας σχέσης. Αυτό σημαίνει πως πρέπει να προηγηθεί κάποιος περιορισμός στις πλειάδες, πριν εφαρμοστεί η τροποποίηση. Έτσι, για να τροποποιηθεί το έτος παραγωγής της ταινίας Casablanca στο 1942 (το σχήμα 1.42 εσφαλμένα δείχνει το έτος 1952) θα πρέπει να γράψουμε:

$$movies \leftarrow \pi_{code,title,year=1942}(\sigma_{code=583}(movies))$$

Ή με την SQL:

```
UPDATE movies
   SET year = 1942
 WHERE code = 0034583;
```

Προσοχή λοιπόν στην τροποποίηση δεδομένων! Σχεδόν πάντα θα πρέπει να συνοδεύεται με περιορισμό των πλειάδων.

Τροποποίηση υπάρχει και εκεί όπου πολλές φορές μοιάζει με εισαγωγή δεδομένων. Θεωρείστε πως η σχέση *movies* επιτρέπει τιμές NULL στην χρονολογία. Αυτό σημαίνει πως, για παράδειγμα είναι έγκυρη η σχεσιακή παράσταση, που αφορά την εισαγωγή της ταινίας La Dolce Vita:

$$movies \leftarrow movies \cup \{0087884, \text{'Paris Texas'}, \text{NULL}\}$$

Όπως έχει αναφερθεί, NULL τιμή σημαίνει είτε πως κάποια άλλη τιμή δεν έχει κάποια εφαρμογή, είτε (όπως θεωρούμε πως συμβαίνει εδώ) δεν είναι γνωστή η τιμή. Έτσι λοιπόν, τα δεδομένα της σχέσης *movies* θα είναι όπως φαίνεται στο σχήμα 1.46.

Κωδικός	Τίτλος	Έτος
0083658	Blade Runner	1942
0034583	Casablanca	1942
0053779	La Dolce Vita	1942
0087884	Paris Texas	NULL

Σχήμα 1.46: Πινακοποιημένη μορφή της σχέσης *movies* χωρίς την χρονολογία της ταινίας *Paris Texas*.

Στην περίπτωση που εξετάζουμε, το έτος παραγωγής προσωρινά ήταν άγνωστο και η καταχώρηση της ταινίας έγινε χωρίς κάποια τιμή, δηλαδή NULL. Όταν κάποια στιγμή γίνει γνωστή αυτή η τιμή (1984), τότε μπορεί να εισαχθεί στη βάση. Θα πρέπει ωστόσο να γίνει κατανοητό πως, σε όρους σχεσιακού μοντέλου, αυτό αποτελεί τροποποίηση των δεδομένων της βάσης και όχι εισαγωγή δεδομένων. Όπως επεξηγήθηκε στην παράγραφο 1.8.1, εισαγωγή δεδομένων στη βάση σημαίνει εισαγωγή πλειάδων σε κάποια σχέση και αύξηση της πληθικότητάς της. Αντίθετα, στην περίπτωση που εξετάζουμε, η εισαγωγή της χρονολογίας παραγωγής της ταινίας *Paris Texas*, δηλαδή της τιμής 1984 στο γνώρισμα *year*, δεν σημαίνει εισαγωγή πλειάδας στη σχέση (επομένως ούτε και αύξηση της πληθικότητάς της), αλλά τροποποίηση πλειάδας. Η τιμή NULL θα αλλάξει σε 1984. Αυτό δεν μεταβάλει την πληθικότητα της σχέσης, απλά τροποποιεί κάποιες τιμές σε ένα γνώρισμα της σχέσης. Η πράξη που γίνεται είναι η τροποποίηση και όχι η εισαγωγή. Στη σχεσιακή άλγεβρα θα γράφουμε:

$$movies \leftarrow \Pi_{\text{Κωδικός, Τίτλος, Έτος}=1984}(\sigma_{\text{Κωδικός}=0087884}(movies))$$

Ή με την SQL:

```
UPDATE movies
   SET year = 1984
 WHERE code = 0087884;
```

Σε καμία περίπτωση δεν μπορεί να γίνει εισαγωγή, όπως:

$$movies \leftarrow movies \cup \{884, \text{'Paris Texas'}, 1984\}$$

Ή με την SQL:

```sql
INSERT
  INTO movies (code, title, year)
VALUES (884, 'Paris Texas', 1984);
```

Αυτό δεν μπορεί να συμβεί γιατί η ταινία Paris Texas με κωδικό (πρωτεύον κλειδί) 0087884, υπάρχει ως πλειάδα της σχέσης movies, και δεν μπορεί να γίνει διπλότυπη εισαγωγή (παραβίαση του κανόνα ακεραιότητας των οντοτήτων και του πρωτεύοντος κλειδιού).