


Σχεσιακή Άλγεβρα

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά

1




Εισαγωγή

Στα προηγούμενα μαθήματα:

- Εννοιολογικός Σχεδιασμός Βάσεων Δεδομένων (με χρήση του Μοντέλου Οντοτήτων/Συσχετίσεων)
- Λογικός Σχεδιασμός Βάσεων Δεδομένων (με χρήση του Σχεσιακού Μοντέλου)
- Αντιστοιχία (μετατροπή) ανάμεσα στα μοντέλα

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά

2



Εισαγωγή

Μετά τη φάση του σχεδιασμού, καταλήγουμε σε ένα σχεσιακό σχήμα.


Δυο ερωτήματα

1. Είναι ο σχεδιασμός μας καλός;
Θεωρία Κανονικών Μορφών
2. Πως θα υλοποιήσουμε (προγραμματίσουμε) την εφαρμογή μας χρησιμοποιώντας ένα ΣΔΒΔ;
Σχεσιακή Άλγεβρα (θεωρητικό υπόβαθρο) - SQL

Θα αρχίσουμε από το ερώτημα 2 - για να δούμε γρήγορα πως η θεωρία βρήκε εφαρμογή σε πραγματικά συστήματα.

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά

3



Εισαγωγή

Τι χρειαζόμαστε: (Η Γενική Εικόνα)

Μια γλώσσα ορισμού δεδομένων ΓΟΔ (για τον ορισμό των σχημάτων)


ένας μεταφραστής της ΓΟΔ επεξεργάζεται τις εντολές της ΓΟΔ, αναγνωρίζει τις περιγραφές των δομικών στοιχείων του σχήματος και αποθηκεύει την περιγραφή του σχήματος στον κατάλογο του ΣΔΒΔ

Μια γλώσσα χειρισμού δεδομένων ΓΧΔ

- γλώσσα ενημέρωσης
- γλώσσα ερωτήσεων (επερωτήσεων) (*Query Language*)

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά

4



Γλώσσες Ερωτήσεων: Εισαγωγή


Γλώσσες Ερωτήσεων (Query Languages): Επιτρέπουν τον χειρισμό και την εύρεση πληροφορίας από μια βάση δεδομένων

Με τη διατύπωση ερωτήσεων στον τρέχων στιγμιότυπο της βάσης δεδομένων (querying)

Το σχεσιακό μοντέλο υποστηρίζει απλές και ισχυρές γλώσσες ερωτήσεων (σε αντίθεση με το μοντέλο Ο/Σ)

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά

5



Εισαγωγή

«Υψηλού επιπέδου» γλώσσες (επε)-ερωτήσεων (SQL - QBE)
Η SQL είναι και ΓΟΔ και ΓΧΔ

«Χαμηλού επιπέδου» γλώσσες ερωτήσεων (σχεσιακή άλγεβρα - σχεσιακός λογισμός)

Θεωρία πως από στιγμιότυπα -> νέα στιγμιότυπα με εφαρμογή πράξεων - τελεστών (σχεσιακή άλγεβρα) με λογικούς κανόνες (σχεσιακός λογισμός)

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά

6



Υπενθύμιση

(Σχεσιακή) Βάση Δεδομένων ένα σύνολο από σχέσεις (πίνακες)

Γραμμές: πλειάδες

Στήλες: Γνωρίσματα

Κλειδί (υπερ-κλειδί, υποψήφιο, πρωτεύον)



Σχεσιακή άλγεβρα: έναν απλό τρόπο δημιουργίας νέων σχέσεων από υπάρχουσες.

Ένα σύνολο από **πράξεις** που όταν εφαρμοστούν σε **σχέσεις** (πίνακες) μας δίνουν **νέες σχέσεις**

Μια ερώτηση εφαρμόζεται σε ένα **στιγμιότυπο σχέσης** και το **αποτέλεσμα** της ερώτησης είναι πάλι ένα στιγμιότυπο σχέσης



Ποιοι είναι κατάλληλοι τελεστές;

Ελάχιστος αριθμός;



Οι πράξεις της σχεσιακής άλγεβρας:

1. Πράξεις που αφαιρούν κομμάτια από μια σχέση είτε **επιλέγοντας γραμμές** είτε **προβάλλοντας στήλες**
2. Οι συνηθισμένες **πράξεις συνόλου** - ένωση, τομή, διαφορά
3. Πράξεις που **συνδυάζουν πλειάδες από δύο σχέσεις**
4. **Μετονομασία γνωρισμάτων**



Η πράξη της επιλογής (select)

Επιλογή ενός υποσυνόλου των πλειάδων μιας σχέσης που ικανοποιεί μια συνθήκη επιλογής

$\sigma_{\langle \text{συνθήκη επιλογής} \rangle}$ (<όνομα σχέσης>)



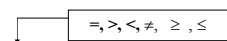
Επιλογή ενός υποσυνόλου των πλειάδων μιας σχέσης που ικανοποιεί μια συνθήκη επιλογής

$\sigma_{\langle \text{συνθήκη επιλογής} \rangle}$ (<όνομα σχέσης>)

συνθήκη

προτάσεις της μορφής

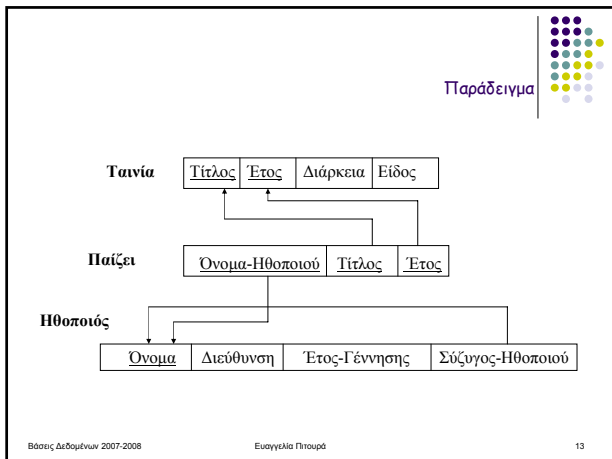
<όνομα γνωρίσματος>



<τελεστής σύγκρισης>

<όνομα γνωρίσματος> ή <σταθερή τιμή από το πεδίο ορισμού του γνωρίσματος>

συνδυασμένες με AND, OR, NOT



Η Πράξη της Επιλογής

Παραδείγματα

| τίτλος | χρόνος | διάρκεια | είδος |
|---------------|--------|----------|---------|
| Star Wars | 1997 | 124 | έγχρωμη |
| Mighty Ducks | 1991 | 104 | έγχρωμη |
| Wayne's World | 1992 | 95 | έγχρωμη |

1. *Ταινίες με διάρκεια μεγαλύτερη των 100 λεπτών*

σ διάρκεια > 100 (Ταινία)

| τίτλος | χρόνος | διάρκεια | είδος |
|--------------|--------|----------|---------|
| Star Wars | 1997 | 124 | έγχρωμη |
| Mighty Ducks | 1991 | 104 | έγχρωμη |

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πλουρά 14

Η Πράξη της Επιλογής

| τίτλος | χρόνος | διάρκεια | είδος |
|---------------|--------|----------|---------|
| Star Wars | 1997 | 124 | έγχρωμη |
| Mighty Ducks | 1991 | 104 | έγχρωμη |
| Wayne's World | 1992 | 95 | έγχρωμη |

2. *Ταινίες με διάρκεια μεγαλύτερη των 100 λεπτών που γυρίστηκαν μετά το 1995*

σ διάρκεια > 100 AND χρόνος > 1995 (Ταινία)

| τίτλος | χρόνος | διάρκεια | είδος |
|-----------|--------|----------|---------|
| Star Wars | 1997 | 124 | έγχρωμη |

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πλουρά 15

- Η Πράξη της Επιλογής
- Η συνθήκη επιλογής εφαρμόζεται ανεξάρτητα σε κάθε πλειάδα
 - Ο τελεστής είναι μοναδιαίος
 - Ο βαθμός της σχέσης που προκύπτει ίδιος με τον βαθμό της αρχικής σχέσης
 - Πλήθος πλειάδων μικρότερο ή ίσο με την αρχική σχέση: ποσοστό που επιλέγονται - επιλεκτικότητα (*selectivity*)
- Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πλουρά 16

Η Πράξη της Επιλογής

Ιδιότητες

- **αντιμεταθετική**

$$\sigma_{\langle \text{συνθ1} \rangle} (\sigma_{\langle \text{συνθ2} \rangle} (R)) = \sigma_{\langle \text{συνθ2} \rangle} (\sigma_{\langle \text{συνθ1} \rangle} (R))$$

- $\sigma_{\langle \text{συνθ1} \rangle} (\sigma_{\langle \text{συνθ2} \rangle} (\dots \sigma_{\langle \text{συνθn} \rangle} (R) \dots)) =$

$$\sigma_{\langle \text{συνθ1} \rangle \text{ AND } \langle \text{συνθ2} \rangle \dots \text{ AND } \langle \text{συνθn} \rangle} (R)$$

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πλουρά 17

Η Πράξη της Προβολής

Η πράξη της προβολής (project)

Επιλογή συγκεκριμένων στηλών (γνωρισμάτων)

Π«λίστα γνωρισμάτων» («όνομα σχέσης»)

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πλουρά 18



Παραδείγματα

| τίτλος | χρόνος | διάρκεια | είδος |
|---------------|--------|----------|---------|
| Star Wars | 1997 | 124 | έγχρωμη |
| Mighty Ducks | 1991 | 104 | έγχρωμη |
| Wayne's World | 1992 | 95 | έγχρωμη |



1. Τίτλος, χρόνος, διάρκεια των ταινιών

π τίτλος, χρόνος, διάρκεια (Ταινία)

| τίτλος | χρόνος | διάρκεια |
|---------------|--------|----------|
| Star Wars | 1997 | 124 |
| Mighty Ducks | 1991 | 104 |
| Wayne's World | 1992 | 95 |



2. Είδος ταινιών

π είδος (Ταινία)

$\frac{\text{είδος}}{\text{έγχρωμη}}$

Προσοχή: απαλοιφή διπλότιμων

Γιατί:

Με βάση τον αριθμό το αποτέλεσμα είναι σχέση (δηλαδή, σύνολο πλειάδων)



- Τα γνωρίσματα έχουν την ίδια διάταξη
- Ο τελεστής είναι μοναδιαίος
- Ο βαθμός της σχέσης είναι ίσος με τον αριθμό γνωρισμάτων στη <λίστα γνωρισμάτων>
- Πλήθος πλειάδων μικρότερο ή ίσο (πότε;) με την αρχική σχέση



Ιδιότητες

- αντιμεταθετική;
- π <λίστα1> (π <λίστα2> (R)) = ?



Παράδειγμα

Διάρκειες μεγαλύτερες των 100 λεπτών

π διάρκεια (σ διάρκεια > 100 (Ταινία))

| διάρκεια |
|----------|
| 124 |
| 104 |

Πράξεις συνόλου

- Ένωση (\cup)
- Τομή (\cap)
- Διαφορά ($-$)

Συμβατότητα ως προς την ένωση

Δύο σχέσεις $R(A_1, A_2, \dots, A_n)$ και $S(B_1, B_2, \dots, B_n)$ είναι συμβατές ως προς την ένωση όταν

1. Έχουν τον ίδιο βαθμό n
2. $\forall i, \text{dom}(A_i) = \text{dom}(B_i)$

• Σύμβαση: η προκύπτουσα σχέση έχει τα ίδια ονόματα γνωρισμάτων με την πρώτη σχέση

• Απαλοιφή διπλότιμων

Οι πράξεις της σχεσιακής άλγεβρας:

- ✓ 1. Πράξεις που αφαιρούν κομμάτια από μια σχέση είτε *επιλέγοντας* γραμμές είτε *προβάλλοντας* στήλες
- ✓ 2. Οι συνηθισμένες πράξεις συνόλου - ένωση, τομή, διαφορά
3. Πράξεις που συνδυάζουν πλειάδες από δύο σχέσεις
4. Μετονομασία γνωρισμάτων

| | | |
|----------|-------------------|---------------------|
| | <u>A</u> <u>B</u> | Παράδειγμα |
| R | 1 2 | $\sigma_{A > B}(R)$ |
| | 1 4 | $\pi_A(R)$ |
| | 2 1 | $R \cup S$ |
| | 6 5 | $R \cap S$ |
| | <u>B</u> <u>C</u> | $R - S$ |
| S | 2 3 | $S - R$ |
| | 2 5 | |
| | 1 4 | |

Μετονομασία

• όνομα στην ενδιάμεση σχέση

$R \leftarrow$

Παράδειγμα

$MEΓΑΛΗΣ_ΔΙΑΡΚΕΙΑΣ \leftarrow \sigma_{\text{διάρκεια} > 100}(\text{Ταινία})$

• μετονομασία γνωρισμάτων

$R(\text{λίστα-με-νέα-ονόματα}) \leftarrow$

Παράδειγμα

$MEΓΑΛΗΣ_ΔΙΑΡΚΕΙΑΣ(\text{όνομα ταινίας}, \text{έτος παραγωγής}, \text{διάρκεια}, \text{είδος}) \leftarrow \sigma_{\text{διάρκεια} > 100}(\text{Ταινία})$

| όνομα ταινίας | έτος παραγωγής | διάρκεια | είδος |
|---------------|----------------|----------|---------|
| Star Wars | 1997 | 124 | έγχρωμη |
| Mighty Ducks | 1991 | 104 | έγχρωμη |

Καρτεσιανό Γινόμενο

(ή χιαστί γινόμενο (cross product) ή χιαστί συνένωση (cross join))

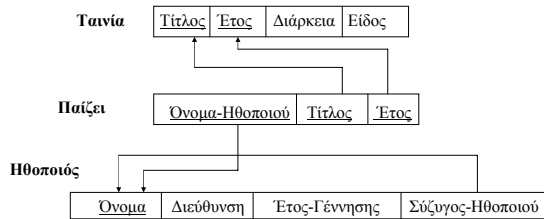
$$R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$$

αποτέλεσμα η σχέση Q: $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$

- $n + m$ γνωρίσματα
- $n_R * n_S$ πλειάδες

R x S

| | | | | | | | | | | |
|---|---|---|---|----|----|---|-----|-----|----|----|
| | R | | S | | | | | | | |
| | A | B | B | C | D | A | R.B | S.B | C | D |
| 1 | 2 | | 2 | 5 | 6 | 1 | 2 | 2 | 5 | 6 |
| 1 | 2 | | 4 | 7 | 8 | 1 | 2 | 4 | 7 | 8 |
| 3 | 4 | | 4 | 7 | 8 | 1 | 2 | 9 | 10 | 11 |
| | | | 9 | 10 | 11 | 3 | 4 | 2 | 5 | 6 |
| | | | | | | 3 | 4 | 4 | 7 | 8 |
| | | | | | | 3 | 4 | 9 | 10 | 11 |



| | | | | |
|----------------|-----------------------|-----------------|--------------|---------------|
| Τίτλος | Έτος | Διάρκεια | Είδος | |
| Παραμύθι | 1990 | 90 | Έγχρωμη | Ταινία |
| Παραμύθι | 1930 | 120 | Ασπρόμαυρη | |
| Φυγή | 2000 | 98 | Ασπρόμαυρη | |
| Άνοιξη | 1998 | 101 | Έγχρωμη | |
| | Όνομα-Ηθοποιού | Τίτλος | Έτος | |
| Αλίκη Πατπά | Παραμύθι | 1930 | | Παιζαί |
| Μαρία Γεωργίου | Παραμύθι | 1990 | | |
| Κώστας Χρήστου | Φυγή | 2000 | | |
| Μαρία Στεργίου | Άνοιξη | 1998 | | |

| | | | | | | | | |
|---------------|----------------------|--------------------|-----------------|--------------|---------------|-----------------------|----------------------|--------------------|
| Ταινία | Ταινία_Τίτλος | Ταινία_Έτος | Διάρκεια | Είδος | Παιζαί | Όνομα-Ηθοποιού | Παιζαί_Τίτλος | Παιζαί_Έτος |
| Παραμύθι | Παραμύθι | 1990 | 90 | Έγχρωμη | Παραμύθι | Αλίκη Πατπά | Παραμύθι | 1930 |
| Παραμύθι | Παραμύθι | 1990 | 90 | Έγχρωμη | Παραμύθι | Μαρία Γεωργίου | Παραμύθι | 1990 |
| Φυγή | Φυγή | 2000 | 98 | Ασπρόμαυρη | Φυγή | Κώστας Χρήστου | Φυγή | 2000 |
| Άνοιξη | Άνοιξη | 1998 | 101 | Έγχρωμη | Άνοιξη | Μαρία Στεργίου | Άνοιξη | 1998 |
| ... | ... | ... | ... | ... | ... | Κατερίνα Αποστόλου | Φυγή | 2000 |
| ... | ... | ... | ... | ... | ... | Αλίκη Πατπά | Παραμύθι | 1930 |
| ... | ... | ... | ... | ... | ... | Μαρία Γεωργίου | Παραμύθι | 1990 |
| ... | ... | ... | ... | ... | ... | Κώστας Χρήστου | Φυγή | 2000 |
| ... | ... | ... | ... | ... | ... | Μαρία Στεργίου | Άνοιξη | 1998 |
| ... | ... | ... | ... | ... | ... | Κατερίνα Αποστόλου | Φυγή | 2000 |
| ... | ... | ... | ... | ... | ... | Αλίκη Πατπά | Παραμύθι | 1930 |

Παράδειγμα

Για κάθε ηθοποιό το όνομα και τον τίτλο-έτος για όλες τις έγχρωμες ταινίες στις οποίες παίζει

Παράδειγμα

Για κάθε ηθοποιό το όνομα και τον τίτλο-έτος για όλες τις έγχρωμες ταινίες στις οποίες παίζει

π όνομα, τίτλος, έτος (σ είδος = "έγχρωμη" AND π παίζει.τίτλος = Ταινία.τίτλος AND π παίζει.έτος = Ταινία.έτος (π παίζει \times Ταινία))

ή

π όνομα, τίτλος, έτος (σ παίζει.τίτλος = Ταινία.τίτλος AND π παίζει.έτος = Ταινία.έτος (π παίζει \times (σ είδος = "έγχρωμη" (Ταινία))))

Συνένωση (ή Θήτα συνένωση) (join)

συνδυασμός σχετιζόμενων πλειάδων

$$R \bowtie \langle \text{συνθήκη συνένωσης} \rangle S$$

($\equiv \sigma$ <συνθήκη συνένωσης> (R \times S))

Συνθήκη συνένωσης

Προτάσεις της μορφής

$$A_i \langle \text{τελεστής σύγκρισης} \rangle B_j$$

όπου A_i γνώρισμα της R, B_j γνώρισμα της S, και $\text{dom}(A_i) = \text{dom}(B_j)$ συνδυασμένες με AND

- το αποτέλεσμα είναι οι συνδυασμοί πλειάδων που ικανοποιούν τη συνθήκη
- η συνθήκη αποτιμάται για κάθε συνδυασμό
- αποτέλεσμα σχέση Q με n + m γνώρισμα
- πλειάδες με τιμή null σε γνώρισμα συνένωσης δεν εμφανίζονται στο αποτέλεσμα

$$U \bowtie_{A < D} V$$

| U | | | V | | |
|---|---|---|---|---|----|
| A | B | C | B | C | D |
| 1 | 2 | 3 | 2 | 3 | 4 |
| 6 | 7 | 8 | 2 | 3 | 5 |
| 9 | 7 | 8 | 7 | 8 | 10 |

| A | U.B | U.C | V.B | V.C | D |
|---|-----|-----|-----|-----|----|
| 1 | 2 | 3 | 2 | 3 | 4 |
| 1 | 2 | 3 | 2 | 3 | 5 |
| 1 | 2 | 3 | 7 | 8 | 10 |
| 6 | 7 | 8 | 7 | 8 | 10 |
| 9 | 7 | 8 | 7 | 8 | 10 |

$$U \bowtie_{A < D \text{ AND } U.B \neq V.B} V$$

Παράδειγμα

Για κάθε ηθοποιό το όνομα και τον τίτλο-έτος για όλες τις έγχρωμες ταινίες στις οποίες παίζει

π όνομα, τίτλος, έτος (σ παίζει.τίτλος = Ταινία.τίτλος AND π παίζει.έτος = Ταινία.έτος (π παίζει \times (σ είδος = "έγχρωμη" (Ταινία))))

π όνομα, τίτλος, έτος (π παίζει \bowtie παίζει.τίτλος = Ταινία.τίτλος AND π παίζει.έτος = Ταινία.έτος (σ είδος = "έγχρωμη" (Ταινία))))

Συνένωση Ισότητας (equijoin)

όταν χρησιμοποιείται μόνο τελεστής ισότητας

Συνθήκη συνένωσης

Προτάσεις της μορφής

$$A_i = B_j$$

όπου A_i γνώρισμα της R, B_j γνώρισμα της S, και $\text{dom}(A_i) = \text{dom}(B_j)$ συνδυασμένες με AND

Συνένωση Ισότητας

| R | | S | | |
|---|---|---|----|----|
| A | B | B | C | D |
| 1 | 2 | 2 | 5 | 6 |
| 3 | 4 | 4 | 7 | 8 |
| | | 9 | 10 | 11 |

$$R \bowtie S$$

$$R.B = S.B$$

Φυσική Συνένωση

Φυσική Συνένωση

συνένωση ισότητας όπου *παράλείπουμε* το γνώρισμα της δεύτερης σχέσης από το αποτέλεσμα

όταν διαφορετικό όνομα - μετονομασία

$$R^*_{(λίστα1, λίστα2)} S$$

επιλεκτικότητα συνένωσης : μέγεθος αποτελέσματος / ($n_r * n_s$)

Φυσική Συνένωση

| R | | S | | | R * S | | | |
|---|---|---|----|----|-------|---|---|---|
| A | B | B | C | D | A | B | C | D |
| 1 | 2 | 2 | 5 | 6 | 1 | 2 | 5 | 6 |
| 3 | 4 | 4 | 7 | 8 | 3 | 4 | 7 | 8 |
| | | 9 | 10 | 11 | | | | |

Φυσική Συνένωση

| U | | | V | | | U * V | | | |
|---|---|---|---|---|----|-------|---|---|----|
| A | B | C | B | C | D | A | B | C | D |
| 1 | 2 | 3 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 6 | 7 | 8 | 2 | 3 | 5 | 1 | 2 | 3 | 5 |
| 9 | 7 | 8 | 7 | 8 | 10 | 6 | 7 | 8 | 10 |
| | | | | | | 9 | 7 | 8 | 10 |

Φυσική Συνένωση

Παράδειγμα

Για κάθε ηθοποιό το όνομα και τον τίτλο-έτος για όλες τις έγχρωμες ταινίες στις οποίες παίζει

π όνομα, τίτλος, έτος (σ Παίζει.τίτλος = Ταινία.τίτλος AND Παίζει.έτος = Ταινία.έτος) (Παίζει x (σ είδος = "έγχρωμη" (Ταινία)))

π όνομα, τίτλος, έτος (Παίζει \bowtie Παίζει.τίτλος = Ταινία.τίτλος AND Παίζει.έτος = Ταινία.έτος) (σ είδος = "έγχρωμη" (Ταινία))

π όνομα, τίτλος, έτος (Παίζει * (σ είδος = "έγχρωμη" (Ταινία)))

είναι η τρίτη έκφραση πριν την προβολή ισοδύναμη των άλλων δύο:

Σχεσιακή Άλγεβρα

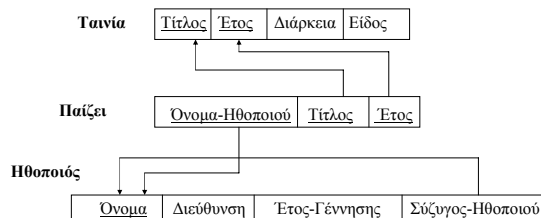
| | | A | B | Παράδειγμα | |
|---|--|---|---|---------------------------|------------------------------|
| R | | 1 | 2 | | |
| | | 1 | 4 | | |
| | | 2 | 1 | | |
| | | 6 | 5 | | |
| S | | B | C | $R \times S$ | $R \bowtie_{R.a \geq S.b} S$ |
| | | 2 | 3 | | |
| | | 2 | 5 | $R \bowtie_{R.a = S.b} S$ | $R * S$ |
| | | 1 | 4 | | |



Πλήρες σύνολο πράξεων

- επιλογή (σ)
- προβολή (π)
- ένωση (\cup)
- διαφορά ($-$)
- καρτεσιανό γινόμενο (\times)

- Επίσης
τομή (\cap)
συνένωση \bowtie
συνένωση ισότητας
φυσική συνένωση ($*$)



Παράδειγμα

Όλες τις ταινίες (τίτλο, έτος) με ηθοποιό τη Βουγιουκλάκη

Όλες τις ταινίες (τίτλο, έτος) μεταξύ 1956 και 1975 με ηθοποιό τη Βουγιουκλάκη



Παράδειγμα

Για κάθε ηθοποιό το όνομα του και τον τίτλο-έτος για όλες τις (έγχρωμες) ταινίες στις οποίες παίζει μαζί με τον σύζυγο του/της

Ονόματα ηθοποιών που δεν έπαιξαν σε καμία ταινία μεταξύ 1995 και 2000



Μερικά ακόμα απλά παραδείγματα

- Τις ταινίες (όλα τα γνωρίσματα) που γυρίστηκαν το 2005
- Μόνο τον τίτλο των ταινιών που γυρίστηκαν το 2005
- Τους ηθοποιούς (ονόματα) που έπαιξαν σε ταινίες που γυρίστηκαν το 2005
- Τους ηθοποιούς (ονόματα) που έπαιξαν σε ταινίες που γυρίστηκαν το 2005, αλλά δεν έπαιξαν σε καμία ταινία που γυρίστηκε το 2004



Γλώσσες Ερωτήσεων (Query Languages): Επιτρέπουν τον χειρισμό και την εύρεση πληροφορίας από μια βάση δεδομένων

Το *σχεσιακό μοντέλο* υποστηρίζει απλές και ισχυρές γλώσσες ερωτήσεων

Σχεσιακή Άλγεβρα (ανακεφαλαίωση)

Μια ερώτηση εφαρμόζεται σε ένα **στιγμιότυπο σχέσης** και το αποτέλεσμα της ερώτησης είναι πάλι ένα στιγμιότυπο σχέσης

Το σχήμα της σχέσης εισόδου είναι ορισμένο

Το σχήμα του αποτελέσματος είναι επίσης ορισμένο

Γλώσσες Ερωτήσεων: Εισαγωγή

Δύο μαθηματικές γλώσσες **ερωτήσεων** αποτελούν τη βάση για τις πραγματικές γλώσσες ερωτήσεων (π.χ., SQL) και για την υλοποίησή τους

• **Σχεσιακή Άλγεβρα**: Λειτουργική "operational" (**database byte-code!**): αποτελείται από ένα σύνολο τελεστών και περιγράφει τα βήματα για τον υπολογισμό του αποτελέσματος

• **Σχεσιακός Λογισμός** (calculus): Επιτρέπει στους χρήστες να περιγράψουν τι θέλουν αλλά *όχι πώς* να το υπολογίσουν

Αυτές οι τυπικές γλώσσες επηρέασαν τις εμπορικές γλώσσες (SQL, QBE) που θα δούμε στα επόμενα μαθήματα

Σχεσιακή Άλγεβρα

Γλώσσες Ερωτήσεων != Γλώσσες Προγραμματισμού!

- Δεν αναμένεται να είναι "Turing complete".
- Δεν αναμένεται να χρησιμοποιηθούν για "δύσκολους υπολογισμούς".
- Υποστηρίζουν εύκολη και αποδοτική προσπέλαση σε μεγάλα σύνολα δεδομένων.

Σχεσιακή Άλγεβρα (ανακεφαλαίωση)

Οι πράξεις της σχεσιακής άλγεβρας:

1. Πράξεις που αφαιρούν κομμάτια από μια σχέση είτε *επιλέγοντας* γραμμές (σ) είτε *προβάλλοντας* στήλες (π)
2. Οι συνηθισμένες πράξεις συνόλου: ένωση, τομή, διαφορά
3. Πράξεις που συνδυάζουν πλειάδες από δύο σχέσεις
4. Μετονομασία γνωρισμάτων

Σχεσιακή Άλγεβρα (ανακεφαλαίωση)

Πλήρες σύνολο πράξεων

επιλογή (σ)
προβολή (π)
ένωση (\cup)
διαφορά ($-$)
καρτεσιανό γινόμενο (\times)

Επίσης
τομή (\cap)
συνένωση \bowtie
συνένωση ισότητας
φυσική συνένωση ($*$)

ΠΡΟΤΙΜΑ(ΠΟΤΗΣ, ΜΠΥΡΑ)
ΣΥΧΝΑΖΕΙ(ΠΟΤΗΣ, ΜΑΓΑΖΙ)
ΣΕΡΒΙΡΕΙ(ΜΑΓΑΖΙ, ΜΠΥΡΑ)

Παράδειγμα

1. Τους πότες που συχνάζουν σε μαγαζιά που σερβίρουν μπίρα «Guinness»
2. Τα μαγαζιά που σερβίρουν μπίρα «Guinness» ή μπίρα «Leffe Brunex» ή και τα δύο
3. Τα μαγαζιά που σερβίρουν μπίρα «Guinness» και μπίρα «Leffe Brunex»
4. Τα μαγαζιά που σερβίρουν **μόνο** μπίρα «Guinness»
5. Μαγαζιά που σερβίρουν **τουλάχιστον δύο** διαφορετικές μπίρες. (**μόνο μία:**)
6. Μαγαζιά που σερβίρουν **ακριβώς δύο** διαφορετικές μπίρες.
7. Τα μαγαζιά που σερβίρουν μπίρες που προτιμά ο πότης «Δημήτρης».
8. Τα μαγαζιά που σερβίρουν **όλες** τις μπίρες που προτιμά ο Δημήτρης



$$R \div S$$

παράδειγμα: βρες τα μαγαζιά που σερβίρουν ΟΛΕΣ τις μπίρες που αρέσουν στο Δημήτρη

R (ΣΕΡΒΙΡΕΙ): Όλα τα μαγαζιά και οι μπίρες που σερβίρουν

S: Όλες οι μπίρες που προτιμά ο Δημήτρης

Q: Τα μαγαζιά που εμφανίζονται στη σχέση ΣΕΡΒΙΡΕΙ (R) με τα υπόλοιπα γνωρίσματα να παίρνουν **όλες τις τιμές** του S



$$R \div S$$

Χρήσιμη όταν **για κάθε**,

παράδειγμα: βρες τον ηθοποιό που παίζει σε όλες (σε κάθε) ταινία που παίζει και ο George Clooney.

R (Παίζει): Όλοι οι ηθοποιοί και οι ταινίες που παίζουν

S: Όλες τις ταινίες που παίζει ο George Clooney

Q: Οι ηθοποιοί που (το όνομα τους) εμφανίζονται στη σχέση Παίζει (R) με υπόλοιπα γνωρίσματα να παίρνουν **όλες τις τιμές** του S



| R | A | B |
|---|----------------|----------------|
| | a ₁ | b ₁ |
| | a ₁ | b ₃ |
| | a ₁ | b ₄ |
| | a ₂ | b ₂ |
| | a ₂ | b ₄ |
| | a ₃ | b ₂ |

S

$$R \div S$$

$$Z = \{A, B\} \quad X = \{B\}$$

$$R(Z) \div S(X), X \subseteq Z$$

Q(Y)?

$$Y = Z - X \quad Y = \{A\}$$

$$t \in Q, \exists t_{R1} \in R, t_{R1}[Y] = t$$

$$\forall t_S \in S, \exists t_R \in R, t_R[X] = t_S \text{ και } t_R[Y] = t$$



$$R \div S;$$

| R | A | B | C |
|---|----------------|----------------|----------------|
| | a ₁ | b ₁ | c ₁ |
| | a ₁ | b ₁ | c ₂ |
| | a ₂ | b ₂ | c ₂ |
| | a ₂ | b ₁ | c ₁ |
| | a ₂ | b ₂ | c ₁ |
| | a ₃ | b ₁ | c ₁ |
| | a ₃ | b ₁ | c ₂ |

Παράδειγμα

S

A

a₁

a₂

a₃



$$R \div S;$$

| R | A | B | C |
|---|----------------|----------------|----------------|
| | a ₁ | b ₁ | c ₁ |
| | a ₁ | b ₁ | c ₂ |
| | a ₂ | b ₂ | c ₂ |
| | a ₂ | b ₂ | c ₃ |
| | a ₂ | b ₂ | c ₁ |
| | a ₃ | b ₁ | c ₁ |
| | a ₃ | b ₁ | c ₂ |

Παράδειγμα

S

A

a₁

a₂

a₃

B

b₁

b₂



$$R(Z) \div S(X), X \subseteq Z$$

Το αποτέλεσμα είναι μια καινούργια σχέση Q(Y) όπου Y = Z - X και

t ∈ Q(Y) αν

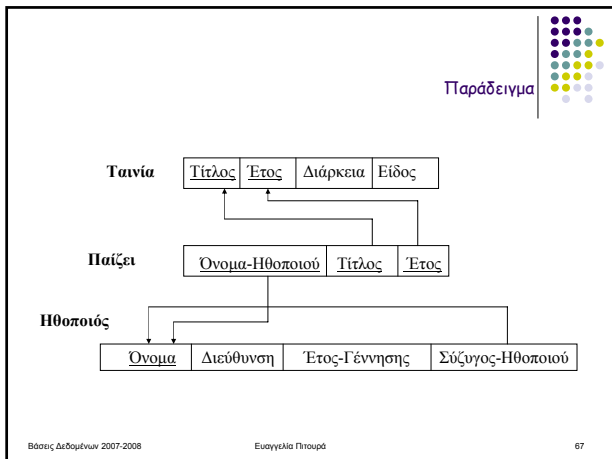
$$\exists t_{R1} \in R, t_{R1}[Y] = t \text{ και}$$

$$\forall t_S \in S, \exists t_R \in R, t_R[X] = t_S \text{ και } t_R[Y] = t$$

• αναλογία με τη διαίρεση ακεραίων

διαίρεση ακεραίων: R / S το αποτέλεσμα Q τέτοιο ώστε: Q * S ≤ R

διαίρεση σχέσεων: R ÷ S το αποτέλεσμα Q τέτοιο ώστε ...



Διαίρεση

παράδειγμα: βρες τον ηθοποιό που παίζει σε όλες (σε κάθε) ταινία που παίζει και ο George Clooney.

S: Όλες τις ταινίες που παίζει ο George Clooney
 Q: Οι ηθοποιοί που (το όνομα τους) εμφανίζονται στη σχέση Παίζει (R) με υπόλοιπα γνωρίσματα να παίρνουν όλες τις τιμές του S

$S \leftarrow \pi_{\text{τίτλος, έτος}} (\sigma_{\text{Όνομα Ηθοποιού} = \text{George Clooney}} (\text{Παίζει}))$
 $Q \leftarrow \text{Παίζει} \div S$

Χωρίς να χρησιμοποιήσω την πράξη της διαίρεσης:

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά 68

Διαίρεση

Ισοδύναμη έκφραση για το $Q(Y) \leftarrow R(Z) \div S(X)$

- Υπολογισμός των πλειάδων που δεν πρέπει να είναι στο αποτέλεσμα.

Μια πλειάδα γ αποκλείεται από το αποτέλεσμα αν και μόνον αν: όταν της συνάψουμε μια τιμή x από το S, η πλειάδα $\langle \gamma, x \rangle$ δεν ανήκει στο R

$T_1 \leftarrow (\pi_{\gamma} (R) \times S) - R$
 $Q \leftarrow \pi_{\gamma} (R) - \pi_{\gamma} (T_1)$

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά 69

Διαίρεση

Παράδειγμα (εφαρμογή ισοδύναμης έκφρασης): βρες τον ηθοποιό που παίζει σε όλες (σε κάθε) ταινία που παίζει και ο George Clooney.

Μια πλειάδα γ αποκλείεται από το αποτέλεσμα αν όταν τις συνάψουμε μια τιμή x από το S, η πλειάδα $\langle \gamma, x \rangle$ δεν ανήκει στο R

$T_1 \leftarrow (S \times \pi_{\gamma} (R)) - R$
 $Q \leftarrow \pi_{\gamma} (R) - \pi_{\gamma} (T_1)$

$S \leftarrow \pi_{\text{τίτλος, έτος}} (\sigma_{\text{Όνομα Ηθοποιού} = \text{George Clooney}} (\text{Παίζει}))$
 $T_1 \leftarrow (\pi_{\text{ηθοποιός}} (\text{Παίζει}) \times S) - \text{Παίζει}$ (μόνον μόνο οι ηθοποιοί που δεν παίζουν σε κάποια ταινία που παίζει ο Clooney!)
 $Q \leftarrow \pi_{\text{ηθοποιός}} (\text{Παίζει}) - \pi_{\text{ηθοποιός}} (T_1)$

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά 70

Αναδρομική Κλειστότητα

R

| | | | |
|---------|-----------|--------|--------------|
| Αρ_Ταυτ | Διεύθυνση | Μισθός | Προϊστάμενος |
|---------|-----------|--------|--------------|

Δεν είναι δυνατόν να βρούμε όλους τους υφισταμένους που επιτηρεί σε οποιοδήποτε επίπεδο ένας συγκεκριμένος προϊστάμενος (π.χ., Αρ_Ταυτ = M20200)

$\Pi_1 (\text{Προϊστ1}) \leftarrow \pi_{\text{Αρ_Ταυτ}} (\sigma_{\text{Προϊστάμενος} = \text{M20200}} (R))$
 $\Pi_2 (\text{Προϊστ2}) \leftarrow \pi_{\text{Αρ_Ταυτ}} (\Pi_1 \bowtie \text{Προϊστ1} = \text{Προϊστάμενος} (R))$

Παρόμοια, μπορώ να βρω π.χ τους συμπρωταγωνιστές του George Clooney (ηθοποιούς που έπαιξαν σε τουλάχιστον μια ταινία μαζί του), τους συμπρωταγωνιστές των συμπρωταγωνιστών του κλπ άλλα μέχρι ενός βάθους

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά 71

Εξωτερική Συνένωση

Όταν θέλουμε να κρατήσουμε στο αποτέλεσμα όλες τις πλειάδες - και αυτές που δεν ταιριάζουν) είτε της σχέσης στα αριστερά (αριστερή εξωτερική συνένωση) είτε της σχέσης στα δεξιά (δεξιά εξωτερική συνένωση)

| R | | S | | R * S | | | A | | | C | | | B | | |
|---|---|---|---|-------|---|---|---|---|------|---|------|---|---|---|---|
| A | C | A | B | A | C | B | A | C | B | A | C | B | A | C | B |
| 1 | 6 | 1 | 3 | 1 | 6 | 3 | 1 | 6 | 3 | 1 | 6 | 3 | 1 | 6 | 3 |
| 2 | 4 | 1 | 5 | 1 | 6 | 5 | 1 | 6 | 5 | 1 | 6 | 5 | 1 | 6 | 5 |
| | | 3 | 9 | | | | 2 | 4 | null | 3 | null | 9 | | | |

Βάσεις Δεδομένων 2007-2008 Ευαγγελία Πιτουρά 72