



Βάσεις δεδομένων

(6^ο μάθημα)

Βασίλης Ευθυμίου
vefthym@hua.gr



(Συνέχεια)



ΣΧΕΣΙΑΚΗ ΑΛΓΕΒΡΑ





Περιεχόμενα

- Συνέχεια στη Σχεσιακή άλγεβρα
 - Συνένωση
 - Θήτα
 - Ισότητας
 - Φυσική
 - Διαίρεση
 - Σύνθετες λειτουργίες
 - Συναθροιστικές συναρτήσεις
 - Ομαδοποίηση
 - Εξωτερική συνένωση



ΣΤ) Συνένωση - join

ή θήτα συνένωση

- Συνδυάζει σχετιζόμενες πλειάδες

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

- Ισοδύναμη της επιλογής στο καρτεσιανό γινόμενο

$$R \bowtie_{\langle \text{συνθήκη συνένωσης} \rangle} S \equiv \sigma_{\langle \text{συνθήκη συνένωσης} \rangle} (R \times S)$$

- Η συνθήκη συνένωσης περιλαμβάνει προτάσεις **A_i <τελεστής σύγκρισης> B_j** (με $\text{dom}(A_i) = \text{dom}(B_j)$) συνδυασμένες με AND

Παράδειγμα

R

A	B	C
1	2	3
6	7	8
9	7	8

S

B	C	D
2	3	4
2	4	5
7	8	10

R ⋈ **A < D** **S**

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

R ⋈ **A < D AND R.B ≠ S.B** **S**

A	R.B	R.C	S.B	S.C	D
1	2	3	7	8	10



Ζ) Συνένωση ισότητας - equijoin

- Χρησιμοποιείται μόνο ο τελεστής ισότητας
- Συνθήκη συνένωσης
 - Προτάσεις της μορφής: $A_i = B_j$
 - συνδυασμένες με AND
 - όπου
 - A_i γνώρισμα της R ,
 - B_j γνώρισμα της S ,
 - και $\text{dom}(A_i) = \text{dom}(B_j)$

Παράδειγμα

R

A	B	C
1	2	3
6	7	8
9	7	8

S

B	C	D
2	3	4
2	4	5
7	8	10

$R \bowtie_{R.B=S.B} S$

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

$R \bowtie_{R.B=S.B \text{ AND } R.C=S.C} S$

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



Η) Φυσική συνένωση

- Συνένωση ισότητας όπου παραλείπουμε το γνώρισμα της δεύτερης σχέσης από το αποτέλεσμα (π.χ. αν $R.B=S.B$ κρατούμε μόνο ένα B)
- Για διαφορετικό όνομα κάνουμε μετονομασία

$$R * S$$

- Επιλεκτικότητα συνένωσης = μέγεθος αποτελέσματος / ($n_r * n_s$)

Παράδειγμα

R

A	B	C
1	2	3
6	7	8
9	7	8

S

B	C	D
2	3	4
2	4	5
7	8	10

R * S

A	B	C	D
1	2	3	4
6	7	8	10
9	7	8	10

Παράδειγμα

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

π όνομα, τίτλος, έτος $(\sigma_{\text{Παίζει.τίτλος} = \text{Ταινία.τίτλος} \text{ AND } \text{Παίζει.έτος} = \text{Ταινία.έτος}} (\sigma_{\text{είδος} = \text{"έγχρωμη"}} (\text{Ταινία})))$

π όνομα, τίτλος, έτος $(\text{Παίζει} \bowtie \sigma_{\text{Παίζει.τίτλος} = \text{Ταινία.τίτλος} \text{ AND } \text{Παίζει.έτος} = \text{Ταινία.έτος}} (\sigma_{\text{είδος} = \text{"έγχρωμη"}} (\text{Ταινία})))$

π όνομα, τίτλος, έτος $(\text{Παίζει} * (\sigma_{\text{είδος} = \text{"έγχρωμη"}} (\text{Ταινία})))$

Άσκηση

R

A	B
1	2
1	4
2	1

S

B	C
2	3
2	5
1	4

R $\bowtie$ **R.a** $\geq$ **S.b** **S**

R $\bowtie$ **R.a** $=$ **S.b** **S**

R * **S**

Συνοψίζοντας

- επιλογή (σ)
- προβολή (π)
- μετονομασία ($\leftarrow$)
- ένωση ($\cup$)
- διαφορά ($-$)
- τομή ($\cap$)
- καρτεσιανό γινόμενο ($\times$)
- Συνένωση ($\bowtie$)
- συνένωση ισότητας
- φυσική συνένωση ($*$)



Ασκήσεις

- Όλα τα έργα (όνομα, τοποθεσία) για το τμήμα *Research*
- Όλα τα ονόματα και διευθύνσεις των εργαζομένων στο τμήμα *Research*
- Όλα τα ονόματα έργων στο *Houston* για το τμήμα *Headquarters*
- Για κάθε τμήμα το όνομα του και την τοποθεσία στην οποία βρίσκεται
- Τους υπαλλήλους (όλα τα γνωρίσματα) που εργάζονται στο *Bellaire*
- Ονόματα εργαζομένων που ασφαλίζουν τις συζύγους τους
- Ονόματα εργαζομένων του τμήματος *Research* που ασφαλίζουν τις συζύγους τους
- Ονόματα εργαζομένων που δεν έχουν κανένα προστατευόμενο μέλος
- Τους εργαζόμενους (ονόματα) και τους προϊσταμένους τους (ονόματα)
- Τους εργαζόμενους που δουλεύουν σε τμήμα στο *Bellaire* και απασχολήθηκαν σε έργο στο *Bellaire* (ονόματα υπαλλήλων και έργων και ώρες εργασίας)

Από το βιβλίο
Εικ. 5.6

EMPLOYEE	FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
	John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
	Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
	Alicia	J	Zelaya	999887777	1968-07-19	3321 Castle, Spring, TX	F	25000	987654321	4
	Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
	Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
	Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
	Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
	James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	null	1

DEPARTMENT	DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
	Research	5	333445555	1988-05-22
	Administration	4	987654321	1995-01-01
	Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS	DNUMBER	DLOCATION
	1	Houston
	4	Stafford
	5	Bellaire
	5	Sugarland
	5	Houston

PROJECT	PNAME	PNUMBER	PLOCATION	DNUM
	ProductX	1	Bellaire	5
	ProductY	2	Sugarland	5
	ProductZ	3	Houston	5
	Computerization	10	Stafford	4
	Reorganization	20	Houston	1
	Newbenefits	30	Stafford	4

WORKS_ON	ESSN	PNO	HOURS
	123456789	1	32.5
	123456789	2	7.5
	666884444	3	40.0
	453453453	1	20.0
	453453453	2	20.0
	333445555	2	10.0
	333445555	3	10.0
	333445555	10	10.0
	333445555	20	10.0
	999887777	30	30.0
	999887777	10	10.0
	987987987	10	35.0
	987987987	30	5.0
	987654321	30	20.0
	987654321	20	15.0
	888665555	20	null

DEPENDENT	ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
	333445555	Alice	F	1986-04-05	DAUGHTER
	333445555	Theodore	M	1983-10-25	SON
	333445555	Joy	F	1958-05-03	SPOUSE
	987654321	Abner	M	1942-02-28	SPOUSE
	123456789	Michael	M	1988-01-04	SON
	123456789	Alice	F	1988-12-30	DAUGHTER
	123456789	Elizabeth	F	1967-05-05	SPOUSE



Θ) Διαίρεση - division

$$R \div S$$

- $R(A) \div S(B), B \subseteq A$
- Το αποτέλεσμα είναι μια σχέση $Q(Y)$, όπου $Y = A \setminus B$ με πλειάδες $t \in Q(Y)$ αν και μόνο αν:
 - $\exists t_{R1} \in R, t_{R1}[Y] = t$ και
 - $\forall t_S \in S, \exists t_R \in R, t_R[X] = t_S, \text{ και } t_R[Y] = t$
- Για το Q ισχύει $Q(Y) \times S(B) \subseteq R(A)$
- Χρήσιμη όταν το ερώτημα περιέχει τη λέξη **κάθε** (ή όλα)

Παράδειγμα (1)

Βρες τους υπαλλήλους που έχουν εργαστεί σε κάθε έργο (σε όλα τα έργα) του τμήματος Administration

- **R** : Όλοι οι υπάλληλοι και τα έργα στα οποία έχουν εργαστεί
 $R \leftarrow \pi_{ESSN, PNO}(Works_on)$
- **S**: Οι κωδικοί για όλα τα έργα του τμήματος Administration
 $A \leftarrow \pi_{dnumber}(\sigma_{dname='Administration'}(Department))$
 $S \leftarrow \pi_{Pnumber}(A \bowtie_{A.Dnumber=Project.Dnum} Project)$
- **Q**: Οι κωδικοί υπαλλήλων που εμφανίζονται στη σχέση **R** με τα υπόλοιπα γνωρίσματα να παίρνουν **όλες τις τιμές** του **S**
- $Q \leftarrow R \div S$

Παράδειγμα (1) – συνέχεια

Βρες τους υπαλλήλους που έχουν εργαστεί σε κάθε έργο του τμήματος Administration

R

ESSN	PNO
123456789	1
123456789	2
...	...
333444555	3
333444555	10
333444555	20
999887777	30
999887777	10
987987987	10
987987987	30
987654321	30
987654321	20
888665555	20

123456789

999887777

$A \bowtie_{A.Dnumber=Project.Dnum} Project$

Dnumber	Pnumber	Dnum
4	10	4
4	30	4

Pnumber
10
30

S

ESSN
999887777
987987987

$Q \leftarrow R \div S$

Παράδειγμα (2)

R

A	B	C
α1	b1	c1
a1	b1	c2
a1	b2	c1
a2	b1	c1
a2	b2	c1
a2	b1	c3
a2	b2	c3

S

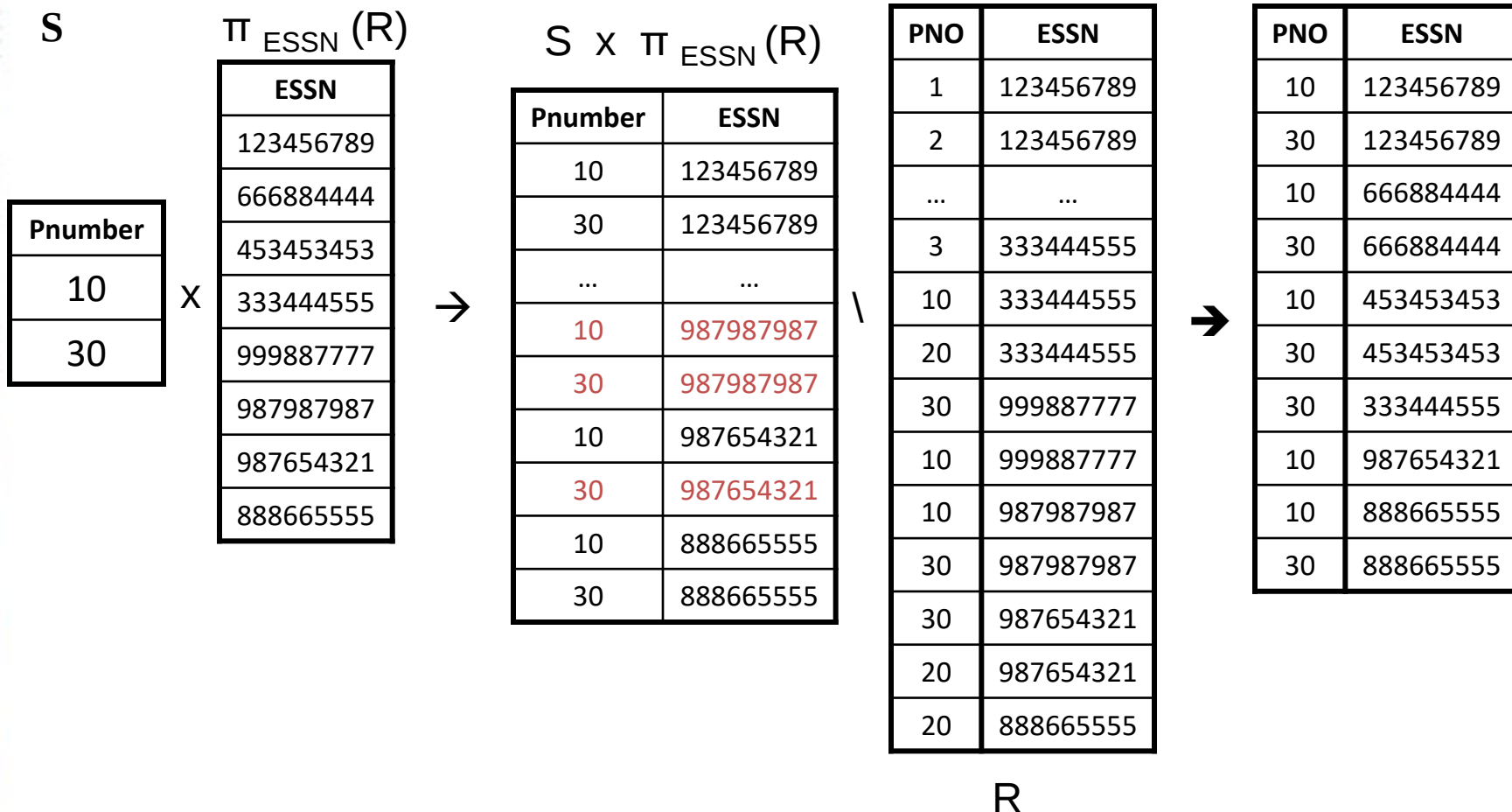
A	B
α1	b1
a2	b2

R ÷ S

C
c1

Αντί της διαίρεσης

1. Υπολογίζω τις πλειάδες που δεν πρέπει να είναι στο αποτέλεσμα: $T \leftarrow (S \times \pi_Y(R)) \setminus R$



Αντί της διαίρεσης

2. Τις αφαιρώ από την αντίστοιχη προβολή

$$Q \leftarrow \pi_Y(R) \setminus \pi_Y(T)$$

$\pi_{\text{ESSN}}(R)$

ESSN
123456789
666884444
453453453
333444555
999887777
987987987
987654321
888665555

$\pi_{\text{ESSN}}(T)$

ESSN
123456789
666884444
453453453
333444555
987654321
888665555

$\setminus$



Q

ESSN
999887777
987987987

Άρα βρίσκω τους υπαλλήλους που δε συμμετέχουν σε όλα τα έργα του Administration (όλοι οι υπάλληλοι και όλα τα έργα που εργάζονται μείον όσες αναθέσεις αφορούν το τμήμα Admininstration) και τους αφαιρώ από όλους τους υπαλλήλους που έχουν εργαστεί



Σύνθετες λειτουργίες



Συναθροιστικές συναρτήσεις

- **Συνάθροιση:** συνδυασμός των πλειάδων μιας σχέσης για τον υπολογισμό μιας συναθροιστικής τιμής
- συνήθεις συναρτήσεις: SUM, AVERAGE, MAX, MIN, COUNT (πλήθος πλειάδων)
- Παραδείγματα: πόσοι υπάλληλοι υπάρχουν συνολικά; ποιος υπάλληλος έχει το μεγαλύτερο μισθό;
- Οι συναρτήσεις παίρνουν ως παράμετρο μια συλλογή (όχι σύνολο) από τιμές
- αποτέλεσμα: μια σχέση και όχι μια τιμή

F <λίστα συναρτήσεων> (**<όνομα σχέσης>**)
όπου <λίστα συναρτήσεων> = {<συνάρτηση, γνώρισμα>}

Παραδείγματα

- Μέση διάρκεια ταινιών

$f_{\text{AVERAGE διάρκεια}}(\text{Ταινία})$

- παλιότερη και πιο πρόσφατη έγχρωμη ταινία

$f_{\text{MIN έτος, MAX έτος}}(\sigma_{\text{είδος} = \text{έγχρωμη}}(\text{Ταινία}))$

Τίτλος	Έτος	Διάρκεια	Είδος
Star Wars	1977	115	έγχρωμη
Casablanca	1942	102	ασπρόμαυρη
City of Gods	2002	130	έγχρωμη

Ομαδοποίηση

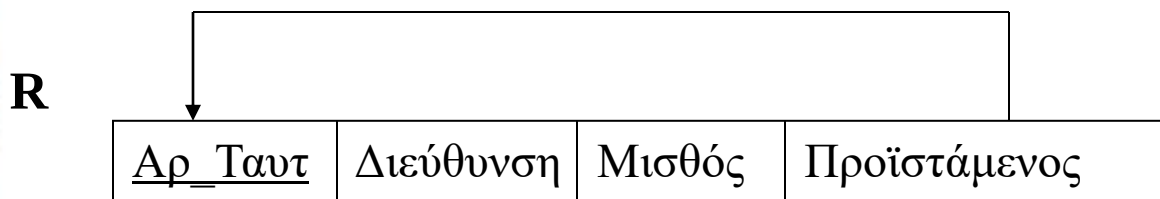
<γνωρίσματα ομαδοποίησης> $\mathcal{F}$ <λίστα συναρτήσεων> (<όνομα σχέσης>)

Παράδειγμα: πόσοι ηθοποιοί ανά ταινία;

τίτλος, έτος $\mathcal{F}$ COUNT Όνομα-Ηθοποιού (Παίζει)

*Ποιο θα ήταν το αποτέλεσμα αν δεν υπήρχαν
τα γνωρίσματα ομαδοποίησης;*

Αναδρομική κλειστότητα



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

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

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

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**) – είτε και των δύο (**πλήρης εξωτερική συνένωση**)

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