



EPL342 –Databases

Lecture 9: RA I

Relational Algebra

(Chapter 6.1-6.3 Elmasri-Navathe 5ED)

Διδάσκων: Παναγιώτης Ανδρέου

<http://www.cs.ucy.ac.cy/courses/EPL342>

Εισαγωγή στη Σχεσιακή Άλγεβρα



- Η Σχεσιακή Άλγεβρα παρέχει τους **τελεστές (operators)**:
 - **Μοναδιαίοι Σχεσιακοί Τελεστές (Unary Relational Ops)**
 - Επιλογή (Select, σ (sigma))
 - Προβολή (Project, π (pi))
 - Μετονομασία (Rename, ρ (rho))
 - **Σχεσιακοί Τελεστές από την Θεωρία Συνόλων**
 - Ένωση (UNION, \cup), Τομή (INTERSECTION, \cap), Διαφορά Συνόλων (DIFFERENCE ή MINUS, $-$)
 - Καρτεσιανό Γινόμενο (CARTESIAN PRODUCT, \times)
 - **Δυαδικοί Σχεσιακοί Τελεστές (Binary Relational Ops)**
 - Συνένωση (JOIN, \bowtie) (υπάρχουν πολλαπλές εκδοχές)
 - Διαίρεση (DIVISION, $/$)
 - **Επιπλέον Σχεσιακοί Τελεστές**
 - Συναρτήσεις Συνάθροισης AGGREGATE FUNCTIONS (π.χ., SUM, COUNT, AVG, MIN, MAX)
 - Εξωτερική Συνένωση, κ.α., (OUTER JOINS)

Τυπικές Γλώσσες Σχεσιακού Μοντέλου

(Formal Languages for the Relational Model)



- Για την εκτέλεση ερωτήσεων σε σχέσεις ορίζονται οι ακόλουθες τυπικές (θεωρητικές) γλώσσες:
 - Σχεσιακή Άλγεβρα (Relational Algebra)
 - Λογισμός Πλειάδων (Tuple Relational Calculus)
 - Λογισμός Πεδίων (Domain Relational Calculus)
- Αυτές οι γλώσσες είναι **εκφραστικά ισοδύναμες!**
 - Δηλαδή μπορούν να **εκφράσουν αντίστοιχες ερωτήσεις.**

– Οποιαδήποτε γλώσσα είναι εκφραστικά ισοδύναμη με την Σχεσιακή Άλγεβρα ονομάζεται **Σχεσιακά Πλήρης Γλώσσα (Relationally-Complete Language)**

- Η SQL είναι μια relationally complete γλώσσα

Τυπικές Γλώσσες Σχεσιακού Μοντέλου

(Formal Languages for the Relational Model)



- Πάνω σε αυτές τις **θεωρητικές γλώσσες** έχουν δημιουργηθεί **πραγματικές γλώσσες βάσεων δεδομένων** (π.χ., **SQL, QBE**)
 - **Σχεσιακή Άλγεβρα (Relational Algebra)**
 - **Προστακτική:** ορίζεται η σειρά εκτέλεσης των πράξεων.
 - Μια Επερώτηση του χρήστη σε SQL μεταφράζεται από την βάση σε μια έκφραση **σχεσιακής άλγεβρας**, το λεγόμενο **πλάνο εκτέλεσης (query plan)**.
 - **Λογισμός Πλειάδων (Tuple Calculus)**
 - **Δηλωτική:** δεν ορίζεται η σειρά εκτέλεσης των πράξεων απλά το επιθυμητό αποτέλεσμα,
 - Η **SQL** στηρίζεται πάνω στην δηλωτική φύση του **Λογισμού Πλειάδων** εάν SQL επερωτήσεις εκτελούνται σαν εκφράσεις **Σχεσιακής Άλγεβρας**
 - **Λογισμός Πεδίων (Domain Calculus)**
 - Όμοια με Λογισμό Πλειάδων (οι μεταβλητές είναι γνωρίσματα και όχι πλειάδες)
 - **Δηλωτική** και πάνω σ' αυτή στηρίζεται η **QBE**
 - Υλοποιείται IBM QMF, MS Access και Paradox, κ.α.

Τυπικές Γλώσσες Σχεσιακού Μοντέλου

(Formal Languages for the Relational Model)



- Παράδειγμα SQL (πραγματικής γλώσσας βάσης)

- `SELECT fname, lname, salary`
- `FROM EMPLOYEE`
- `WHERE dno=5`

Δηλωτική διατύπωση,
Προστακτική εκτέλεση

- Παράδειγμα Σχεσιακής Άλγεβρας:

- `TEMP ← σDNO=5(EMPLOYEE)`
- `RESULT ← πFNAME, LNAME, SALARY(TEMP)`

Προστακτική διατύπωση
& εκτέλεση

- Παράδειγμα Λογισμού Πλειάδων:

- `{ t.Fname, t.Lname, t.salary | EMPLOYEE(t) AND t.Dno=5 }`
- Η πιο πάνω έκφραση ονομάζεται **Set Builder Notation** (Σημειογραφία Δημιουργίας Συνόλων) {<γνωρίσματα αποτελέσματος> | <συνθήκες>}

Δηλωτική διατύπωση
& εκτέλεση

- Παράδειγμα Λογισμού Πεδίων Ορισμού:

- `{ (F,L,S,D) | (F,L,S,D) ∈ EMPLOYEE ^ D=5 }`
- Παρόμοια με Λογισμό Πλειάδων (οι μεταβλητές είναι γνωρίσματα – οι περιορισμοί εκφράζονται στο πεδίο ορισμού)
- Οι εκφράσεις σε σε **QBE** `{ flsd | EMPLOYEE(flsd) and d=5 }` βασίζονται στο πιο πάνω

Δηλωτική διατύπωση
& εκτέλεση

Εισαγωγή στη Σχεσιακή Άλγεβρα



- Η Σχεσιακή Άλγεβρα αποτελείται από ένα **βασικό σύνολο πράξεων (τελεστών)** για το σχεσιακό μοντέλο.
 - Είναι ουσιαστικά η πιο διαδεδομένη τυπική γλώσσα του Σχε. Μοντ.
- Αυτοί οι **τελεστές** επιτρέπουν σε κάποιον να διατυπώσει **βασικές πράξεις ανάκτησης**, δηλ., **επερωτήσεις (queries)**.
- Το **αποτέλεσμα** εκτέλεσης **ΟΛΩΝ** των τελεστών μας επιστρέφει πίσω μια **νέα σχέση**, η οποία μπορεί να δημιουργείται από 1 ή περισσότερες σχέσεις εισόδου.

– Αυτή η **ιδιότητα** κάνει την Σχεσιακή Άλγεβρα **«Κλειστή»** (“closed”), δηλ., ΌΛΑ τα αντικείμενα στη σχεσιακή άλγεβρα είναι σχέσεις (ακόμη και μια απλή αριθμητική τιμή)

- **Έκφραση σχεσιακής άλγεβρας (Relational Algebra Expression):**
Ακολουθία εφαρμογής τελεστών πάνω σε σχέσεις,
Π.χ., $\pi_{FNAME, LNAME, SALARY}(\sigma_{DNO=5}(EMPLOYEE))$

Εισαγωγή στη Σχεσιακή Άλγεβρα



- Η Σχεσιακή Άλγεβρα παρέχει τους **τελεστές (operators)**:
 - **Μοναδιαίοι Σχεσιακοί Τελεστές (Unary Relational Ops)**
 - Επιλογή (Select, σ (sigma))
 - Προβολή (Project, π (pi))
 - Μετονομασία (Rename, ρ (rho))
 - **Σχεσιακοί Τελεστές από την Θεωρία Συνόλων**
 - Ένωση (UNION, \cup), Τομή (INTERSECTION, \cap), Διαφορά Συνόλων (DIFFERENCE ή MINUS, $-$)
 - Καρτεσιανό Γινόμενο (CARTESIAN PRODUCT, \times)
 - **Διαδικοί Σχεσιακοί Τελεστές (Binary Relational Ops)**
 - Συνένωση (JOIN, \bowtie) (υπάρχουν πολλαπλές εκδοχές)
 - Εφόσον το \bowtie δεν υποστηρίζεται σαν σύμβολο θα χρησιμοποιείται το \otimes
 - Διαίρεση (DIVISION, $/$)
 - **Επιπλέον Σχεσιακοί Τελεστές**
 - Συναρτήσεις Συνάθροισης AGGREGATE FUNCTIONS (π.χ., SUM, COUNT, AVG, MIN, MAX)
 - Εξωτερική Συνένωση κ.α., (OUTER JOINS)

Μοναδιαίοι Τελεστές Επιλογή (σ)



- Ο **Τελεστής Επιλογής** (συμβολίζεται με σ (sigma)) χρησιμοποιείται για να **επιλέξουμε** ένα **υποσύνολο** των **πλειάδων** μιας σχέσης βάσει κάποιου **κριτηρίου επιλογής (selection condition)**

σ <Κριτήρια Επιλογής> (R)

– Το κριτήριο επιλογής λειτουργεί ως **φίλτρο** (δηλ., **επιστρέφει** μόνο τις πλειάδες που ικανοποιούν την συνθήκη)

– Π.χ., :

- Εύρεση των EMPLOYEE που δουλεύουν στο department 4:

$\sigma_{DNO = 4}$ (EMPLOYEE)

- Εύρεση των EMPLOYEE που έχουν μισθό πάνω από \$30,000:

$\sigma_{SALARY > 30,000}$ (EMPLOYEE)

Μοναδιαίοι Τελεστές Επιλογή (σ)



σ <Κριτήρια Επιλογής> (R)

- Το κριτήριο επιλογής αποτελείται από **μια (1)** ή περισσότερες **προτάσεις (clauses)** της μορφής:

Κριτήριο Επιλογής := <όνομα γνωρίσματος>
<τελεστής>
<σταθερή τιμή | όνομα γνωρίσματος>

- **Τελεστής:**
 - Διαδικός Τελεστής Σύγκρισης {<, >, =, !=, >=, <=}
- Οι προτάσεις μπορεί να συνδέονται με
 - **Λογικούς Τελεστές:** AND ή OR

– Π.χ., $\sigma_{DNO = 4 \text{ AND } SALARY > 30,000}(\text{EMPLOYEE})$

Μοναδιαίοι Τελεστές Επιλογή (σ)



• Ιδιότητες Τελεστή Επιλογής (σ)

1. Ο τελεστής $\sigma_{\langle \text{selection condition} \rangle}(R)$ παράγει μια νέα σχέση **S** ή οποία έχει το **ίδιο σχήμα** με την **R**, όπου $|S| == |R|$ και $|r(S)| \leq |r(R)|$
- Ισχύει η αντιμεταθετική ιδιότητα (**commutative**) :

$$\sigma_{\langle \text{condition1} \rangle} (\sigma_{\langle \text{condition2} \rangle} (R)) = \sigma_{\langle \text{condition2} \rangle} (\sigma_{\langle \text{condition1} \rangle} (R))$$

– Επίσης, μπορούμε να **διαδώσουμε (cascade)** μια επιλογή

$$\text{δηλ., } \sigma_{\langle \text{cond1} \rangle} (\sigma_{\langle \text{cond2} \rangle} (\sigma_{\langle \text{cond3} \rangle} (R))) = \sigma_{\langle \text{cond2} \rangle} (\sigma_{\langle \text{cond3} \rangle} (\sigma_{\langle \text{cond1} \rangle} (R)))$$

$$\text{Π.Χ., } \sigma_{\langle \text{dno}=5 \rangle} (\sigma_{\langle \text{fname}=\text{\"Chris\"} \rangle} (\sigma_{\langle \text{sex}=\text{\"M\"} \rangle} (\text{EMPLOYEE})))$$

$$= \sigma_{\langle \text{fname}=\text{\"Chris\"} \rangle} (\sigma_{\langle \text{sex}=\text{\"M\"} \rangle} (\sigma_{\langle \text{dno}=5 \rangle} (\text{EMPLOYEE})))$$

– Επίσης, μια **διάδοση τελεστών σ** μπορεί να **αντικατασταθεί** από μια και μόνο επιλογή η οποία θα είναι η **σύζευξη** όλων των συνθηκών:

$$\text{δηλ., } \sigma_{\langle \text{cond1} \rangle} (\sigma_{\langle \text{cond2} \rangle} (\sigma_{\langle \text{cond3} \rangle} (R))) = \sigma_{\langle \text{cond1} \rangle \text{ AND } \langle \text{cond2} \rangle \text{ AND } \langle \text{cond3} \rangle} (R)$$

$$\text{Π.Χ., } \sigma_{\langle \text{dno}=5 \rangle \text{ AND } \langle \text{fname}=\text{\"Chris\"} \rangle \text{ AND } \langle \text{sex}=\text{\"M\"} \rangle} (\text{EMPLOYEE})$$

Μοναδιαίοι Τελεστές Παραδείγματα Επιλογής (σ)



Δεδομένα

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

Επερώτηση

$\sigma_{(Dno=4 \text{ AND } Salary>25000) \text{ OR } (Dno=5 \text{ AND } Salary>30000)} (EMPLOYEE)$

Αποτέλεσμα

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
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

Μοναδιαίοι Τελεστές Προβολή (π)



- Ο **Τελεστής Προβολής** συμβολίζεται με π (pi)) χρησιμοποιείται για να **επιλέξουμε** ένα **υποσύνολο γνωρισμάτων** μιας σχέσης:

$\pi_{\text{Attribute1,Attribute2,...,AttributeN}}(R)$

- Συνεπώς, ο τελεστής αυτός μπορεί να ειπωθεί σαν μια **κατακόρυφη διαμέριση (vertical partitioning)** της αρχικής σχέσης R
- Π.χ., : Ανάκτησε όλα τα **Επίθετα, Ονόματα, Μισθούς** των EMPLOYEE
 $\pi_{\text{LNAME, FNAME, SALARY}}(\text{EMPLOYEE})$
- Ο τελεστής προβολής διαγράφει εξ' ορισμού τα **διπλότυπα (duplicate tuple elimination)**
 - Αυτό διότι το αποτέλεσμα ΠΡΕΠΕΙ να είναι ΣΥΝΟΛΟ πλειάδων και όπως αναφέραμε τα σύνολα δεν επιτρέπουν τα διπλότυπα.

Μοναδιαίοι Τελεστές Προβολή (π)



• Ιδιότητες Τελεστή Προβολής π

1. Ο τελεστής $\pi_{\langle \text{Attributes} \rangle}(\mathbf{R})$ παράγει μια νέα σχέση \mathbf{S} για την οποία ισχύει $|\mathbf{S}| \leq |\mathbf{R}|$ αλλά και $|\mathbf{r}(\mathbf{S})| \leq |\mathbf{r}(\mathbf{R})|$ (λόγω της ενδεχόμενης διαγραφής διπλοτύπων)
 - Εάν το Attributes περιλαμβάνει κάποιο κλειδί της R τότε $|\mathbf{r}(\mathbf{S})| = |\mathbf{r}(\mathbf{R})|$
 - Προφανώς, και στις δυο πιο πάνω περιπτώσεις ο αριθμός των γνωρισμάτων του \mathbf{S} είναι υποσύνολο του R.
2. Ο τελεστής π **ΔΕΝ** είναι **αντιμεταθετικός (not commutative)**:

$$\pi_{\langle \text{list1} \rangle} (\pi_{\langle \text{list2} \rangle} (\mathbf{R})) = \pi_{\langle \text{list2} \rangle} (\pi_{\langle \text{list1} \rangle} (\mathbf{R}))$$

Το πιο πάνω ισχύει ΜΟΝΟ εάν $\text{list1} \subseteq \text{list2}$

Π.Χ., $\pi_{\langle \text{ssn, fname, dno} \rangle} (\pi_{\langle \text{ssn, fname, lname, dno} \rangle} (\text{EMPLOYEE}))$ **ΣΩΣΤΟ**

Π.Χ., $\pi_{\langle \text{ssn, fname, lname, dno} \rangle} (\pi_{\langle \text{ssn, fname, dno} \rangle} (\text{EMPLOYEE}))$ **ΛΑΘΟΣ**

3. Η Προβολή μπορεί να **αντιμετατεθεί** με την **Επιλογή** σε ορισμένες περιπτώσεις

$$\pi_{a_1, \dots, a_n} (\sigma_A(\mathbf{R})) = \sigma_A (\pi_{a_1, \dots, a_n} (\mathbf{R})) \quad \text{where fields in } A \subseteq \{a_1, \dots, a_n\}$$

Μοναδιαίοι Τελεστές Παραδείγματα Προβολής (π)



Δεδομένα

Διπλότυπα

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

Επερώτηση

$\pi_{\text{Sex, Salary}}(\text{EMPLOYEE})$.

Αποτέλεσμα

Sex	Salary
M	30000
M	40000
F	25000
F	43000
M	38000
M	25000
M	55000

7 εγγραφές
αντί 8
εγγραφές
(λόγω
duplicates)

9-14

Εκφράσεις Σχεσιακής Άλγεβρας (Relational Algebra Expressions)



- Οι εκφράσεις σχεσιακής άλγεβρας μπορούν να **διατυπωθούν** κατά τους ακόλουθους τρόπους:
 - A) **Μια έκφραση** η οποία θα αποτελείται από **εμφωλευμένες (nested) εκφράσεις**:
 - Π.χ., $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$
 - B) **Πολλαπλές εκφράσεις** με επί μέρους αποτελέσματα
 - Π.χ., $\text{TEMP} \leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$
 $\text{RESULT} \leftarrow \pi_{\text{FNAME, LNAME, SALARY}}(\text{TEMP})$
- Στην B περίπτωση πρέπει να δηλώσουμε ονόματα για τα ενδιάμεσα αποτελέσματα.

Μοναδιαίοι Σχεσιακοί Τελεστές

Τελεστής RENAME



- Σε ορισμένες περιπτώσεις θέλουμε να **μετονομάσουμε** τα **γνωρίσματα** μιας **σχέσης** ή τη **σχέση** (ή και τα δυο)
 - Χρήσιμο όταν μια επερώτηση χρειάζεται **πολλαπλές πράξεις**
 - Απαραίτητα σε ορισμένες **περιπτώσεις** (π.χ., joins).

- Ο τελεστής **Μετονομασίας (RENAME)** μπορεί να χρησιμοποιείται καθ' οποιοδήποτε απ' τους πιο κάτω τρόπους:

- $\rho_{S(B_1, B_2, \dots, B_n)}(R)$:
 - Αλλαγή Ονόματος Σχέσης από **R** σε **S**
 - **B1, B2, ..., Bn**: Νέα Ονόματα Γνωρισμάτων
- $\rho_S(R)$
 - Αλλαγή Ονόματος Σχέσης από **R** σε **S**
- $\rho_{(B_1, B_2, \dots, B_n)}(R)$: Για επερωτήσεις δεν χρησιμοποιείται γενικά
 - **B1, B2, ..., Bn**: Νέα Ονόματα Γνωρισμάτων Αποτελέσματος
- $\rho_{R'(a \rightarrow b)}(R)$
 - Αλλαγή Ονόματος Γνωρίσματος **R.a** σε **R'.b**

Δεδομένα



EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

Fname	Lname	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

No Renaming

(A) $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$

TEMP

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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
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

R

First_name	Last_name	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

(B) With Renaming

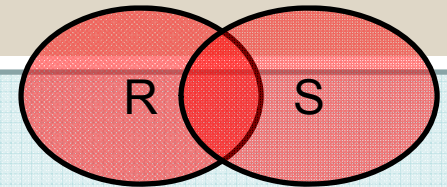
TEMP $\leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$

R(First_Name, Last_Name, Salary)

$\leftarrow \pi_{\text{FNAME, LNAME, SALARY}}(\text{TEMP})$

Τελεστές Θεωρίας Συνόλων

Ένωση (UNION)



- **Τελεστής Ένωσης (UNION)**

- Είναι δυαδικός τελεστές και συμβολίζεται με \cup
- Το αποτέλεσμα του $R \cup S$, είναι μια σχέση η οποία περιλαμβάνει όλες τις πλειάδες οι οποίες ανήκουν **i)** στο R, **ii)** το S ή **iii)** στο R και S.
- $R \cup S = \{x \mid x \in R \vee x \in S\}$

- Θυμηθείτε ότι οι **τελεστές** της **σχεσιακής άλγεβρας** είναι **κλειστοί** (δηλαδή το αποτέλεσμα ενός τελεστή είναι μια νέα σχέση).
- Τα διπλότυπα (duplicate tuples) απαλείφονται

Τελεστές Θεωρίας Συνόλων

Ένωση (UNION)



- Οι δυο τελεσταίοι (operands), R and S, πρέπει να είναι **Συμβατοί-προς-τον-τύπο** (“**type compatible**” ή **UNION-compatible**)
 - R and S πρέπει να έχουν τον ίδιο αριθμό γνωρισμάτων (π.χ., ~~Department~~ \cup ~~Employee~~)
 - Κάθε ζεύγος αντίστοιχων γνωρισμάτων πρέπει να είναι συμβατό-προς-τον-τύπο (πεδίο ορισμού) (π.χ., **Emp(ssn:int) \cup Mgr(ssn:int)**)
- Για την Ένωση ισχύουν οι ακόλουθες ιδιότητες:
 - **$R \cup S = S \cup R$** Αντιμεταθετική (Commutative)
 - **$R \cup (S \cup T) = (R \cup S) \cup T$** Προσεταιριστική (Associative)

Τελεστές Θεωρίας Συνόλων

Ένωση (UNION)



- Παράδειγμα:
 - Βρες το SSN όλων των υπαλλήλων που δουλεύουν στο τμήμα (dno) 5 ή επιβλέπουν κάποιο υπάλληλο που δουλεύει στο τμήμα 5.
 - *Με ποια σειρά σχεσιακών τελεστών θα απαντήσουμε το πιο πάνω ερώτημα;*

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

Τελεστές Θεωρίας Συνόλων Ένωση (UNION)



- Διατύπωση σε Σχεσιακή Άλγεβρα

Επερώτηση: $TEMP \leftarrow \sigma_{DNO=5}(EMPLOYEE)$
 $RESULT1 \leftarrow \pi_{SSN}(TEMP)$
 $RESULT2(SSN) \leftarrow \pi_{SUPERSSN}(TEMP)$
 $RESULT \leftarrow RESULT1 \cup RESULT2$

Σημειώστε την μετονομασία

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

Αποτέλεσμα:

RESULT1

Ssn
123456789
333445555
666884444
453453453

RESULT2

Ssn
333445555
888665555

RESULT

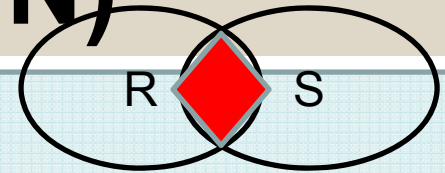
Ssn
123456789
333445555
666884444
453453453
888665555

EPL342:

University of Cyprus)

Τελεστές Θεωρίας Συνόλων

Τομή (INTERSECTION)



- **Τελεστής Τομής (Intersection)**

- Είναι δυαδικός τελεστές και συμβολίζεται με \cap
- Το αποτέλεσμα του $R \cap S$, είναι μια σχέση η οποία περιλαμβάνει **ΟΛΕΣ** τις πλειάδες οι οποίες ανήκουν τόσο στο **R** όσο **ΚΑΙ** στο **S**.

- $R \cap S = \{x \mid x \in R \wedge x \in S\}$

- Τα **διπλότυπα (duplicate tuples)** απαλείφονται και ισχύει και πάλι ο περιορισμός **συμβατότητας-προς-τον-τύπου** (“type compatibility”)
- Για την **Τομή** ισχύουν οι ακόλουθες **ιδιότητες**:
 - $R \cap S = S \cap R$ Αντιμεταθετική (Commutative)
 - $R \cap (S \cap T) = (R \cap S) \cap T$ Προσεταιριστική (Associative)
 - $\overline{(R \cap S)} = \overline{R} \cup \overline{S}$ Κανόνας De Morgan μετατροπής.

Τελεστές Θεωρίας Συνόλων

Τομή (INTERSECTION)



- Παράδειγμα:
 - Βρες το SSN όλων των υπαλλήλων που δουλεύουν στο τμήμα (dno) 5 ΚΑΙ έχουν εξαρτώμενο (dependent)

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

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

Απάντηση:
SSN
333445555

Τελεστές Θεωρίας Συνόλων Τομή (INTERSECTION)



- Διατύπωση σε Σχεσιακή Άλγεβρα

$R1 \leftarrow \pi_{SSN}(\sigma_{DNO=5}(\text{EMPLOYEE}))$

$R2(SSN) \leftarrow \pi_{ESSN}(\text{DEPENDENT})$

$\text{RESULT} \leftarrow R1 \cap R2$

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

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

Το ίδιο ερώτημα
μπορεί να
διατυπωθεί και με
συνένωση (που θα
δούμε αργότερα) 9-24

Επανάληψη Ταυτοτήτων Συνόλων (Revision of Set Identities)



A, B and C subset of a universe D. (σύμπαν)

	$\overline{}$ → Complement (συμπλήρωμα συνόλου)	
1. $\overline{\overline{A}} = A$	Complementation law	Συμπληρωματική ιδιότητα
2. $A \cup \overline{A} = D$	Idempotent 1	Ταυτοδύναμη ιδιότητα
3. $A \cap \overline{A} = \emptyset$	Idempotent 2	
4. $A \cup B = B \cup A$	Commutativity 1	Αντιμεταθετική ιδιότητα
5. $A \cap B = B \cap A$	Commutativity 2	
6. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Associativity 1	Προσεταιριστική Ιδιότητα
7. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	Associativity 2	
8. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributivity 1	Επιμεριστική Ιδιότητα
9. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	Distributivity 2	
10. $\overline{A \cup B} = \overline{A} \cap \overline{B}$	De Morgan's law 1	Ιδιότητες (Κανόνες) De Morgan
11. $\overline{A \cap B} = \overline{A} \cup \overline{B}$	De Morgan's law 2	

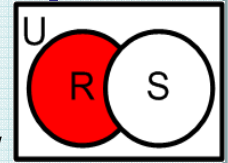
Τελεστές Θεωρίας Συνόλων

Διαφορά Συνόλων (Minus)



- Τελεστής **Διαφοράς (Minus ή Set Difference)**

- Είναι δυαδικός τελεστές και συμβολίζεται με $-$
- Το αποτέλεσμα του $R - S$, είναι μια σχέση η οποία περιλαμβάνει όλες τις πλειάδες της R η οποίες δεν ανήκουν στο S .



- $R - S = \{ x \mid x \in R \wedge x \notin S \} = R \cap \overline{S}$

- Οι δυο τελεσταίοι (operands), R and S, πρέπει και πάλι να είναι **Συμβατοί-προς-τον-τύπο** (“**type compatible**”)
- **ΔΕΝ** ισχύει η **αντιμεταθετική ιδιότητα** :
 $R - S \neq S - R$
- Ισχύει όμως η προσεταιριστική $R - (S - T) = (R - S) - T$

Τελεστές Θεωρίας Συνόλων

Διαφορά Συνόλων (Minus)



Συμβατές-προς-Τύπο

(παρόλο που έχουν διαφορετικά ονόματα γνωρισμάτων)

(a) STUDENT

Fn	Ln
Susan *	Yao
Ramesh*	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

|Student|=7

INSTRUCTOR

Fname	Lname
John	Smith
Ricardo	Browne
Susan *	Yao
Francis	Johnson
Ramesh *	Shah

|Instructor|=5

|Student – Instructor| = 5

(d)

Fn	Ln
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

Αφαίρεση
Susan και
Ramesh
from
Student

|Instructor – Student| = 3

(e)

Fname	Lname
John	Smith
Ricardo	Browne
Francis	Johnson

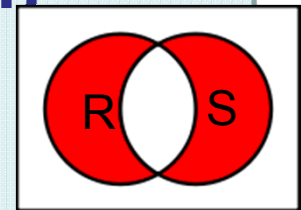
Αφαίρεση
Susan και
Ramesh
from
Instructor

Τελεστές Θεωρίας Συνόλων

Συμμετρική Διαφορά Συνόλων



- Όπως θα θυμάστε, υπάρχει και η **Συμμετρική Διαφορά (Symmetric Difference)**, η οποία συμβολίζεται με το σύμβολο: \oplus



- Αναπαριστάται ουσιαστικά από τον **λογικό τελεστή $\underline{\vee}$ XOR (Exclusive-OR)**
 - Δηλ., $R \oplus S = \{ x \mid x \in R \underline{\vee} x \in S \}$
- Ισχύουν τα ακόλουθα:

$R \oplus S = S \oplus R$ αντιμεταθετική ιδιότητα

$R \oplus S = (R - S) \cup (S - R)$

- Γενικά, υπάρχουν πολλές άλλες ταυτότητες οι οποίες μπορούν να αποδειχθούν με **λογικές ισοδυναμίες, πίνακες συμμετοχής** ή με **χρήση γνωστών ισοδυναμιών**

Τελεστές Θεωρίας Συνόλων

Συμμετρική Διαφορά Συνόλων



|Student|=7

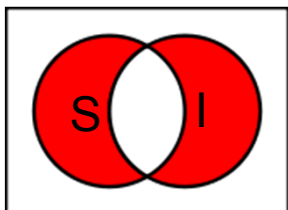
(a) STUDENT

Fn	Ln
Susan *	Yao
Ramesh*	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

|Instructor|=5

INSTRUCTOR

Fname	Lname
John	Smith
Ricardo	Browne
Susan	Yao *
Francis	Johnson
Ramesh	Shah *



$$|S \oplus I| =$$

$$|(S - I) \cup (I - S)| = 5 + 3 = 8$$

Fn	Ln
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert
John	Smith
Ricardo	Browne
Francis	Johnson

Τελεστές Θεωρίας Συνόλων

Καρτεσιανό Γινόμενο (Cartesian Product)



- **Καρτεσιανό Γινόμενο (\times)**
 - **CARTESIAN (or CROSS) PRODUCT**
 - Το $\mathbf{R} \times \mathbf{S}$ είναι το σύνολο όλων των **διατεταγμένων ζευγών** (r,s) , όπου $r \in \mathbf{R}$ και $s \in \mathbf{S}$, δηλ.,
 - $\mathbf{R} \times \mathbf{S} = \{ (r,s) \mid r \in \mathbf{R} \wedge s \in \mathbf{S} \}$
 - Πρακτικά, το \times χρησιμοποιείται για να **συνδυαστούν** οι **πλειάδες δυο σχέσεων** (και θα χρησιμοποιηθεί για να εξηγηθεί ο σημαντικός τελεστής της συνένωσης – join)
 - Το \times μεταξύ δυο σχέσεων βαθμού μεγαλύτερου του 2, π.χ., $\mathbf{R}(A_1, A_2, \dots, A_n) \times \mathbf{S}(B_1, B_2, \dots, B_m)$, μας επιστρέφει μια σχέση \mathbf{Q} βαθμού $n+m$
 - $\mathbf{Q}(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$ (Η \mathbf{Q} έχει $n*m$ πλειάδες)
 - Οι τελεσταίοι \mathbf{R}, \mathbf{S} δεν χρειάζεται να συμφωνούν στον τύπο.

Τελεστές Θεωρίας Συνόλων

Καρτεσιανό Γινόμενο (Cartesian Product)



- Στην πραγματικότητα, το **Καρτεσιανό Γινόμενο** δεν είναι τόσο **πρακτικός** τελεστής.
- Χρησιμοποιείται μόνο όταν θέλουμε να **δημιουργήσουμε** όλες τις **δυνατές διατάξεις** μεταξύ δυο σχέσεων.
 - Π.χ., **STUDENT(ssn) x COURSES(course_id)**
 - Δημιούργησε όλους τους **συνδυασμούς (ssn, course_id)** ... το οποίο δεν είναι και τόσο χρήσιμο.
- Γίνεται ωστόσο πρακτικό εάν χρησιμοποιείται ως **μέρος ακολουθίας άλλων τελεστών**,
 - π.χ., Φτιάχνουμε το **καρτεσιανό γινόμενο** και επιλεγούμε ένα **υποσύνολο** αυτού του συνόλου ως **απάντηση** (το οποίο μας οδηγεί στην έννοια της Συνένωσης (**JOIN**))
- Το επόμενο **παράδειγμα** δείχνει ότι το καρτεσιανό γινόμενο **ΔΕΝ** είναι και τόσο **πρακτικό** από μόνο του.

Τελεστές Θεωρίας Συνόλων

Καρτεσιανό Γινόμενο (Cartesian Product)



Θεωρήστε τη σχέση **Reserves** η οποία μας δείχνει ποια βάρκα (bid) νοίκιασε ο κάθε **Ναυτικός, Sailor (sid)**.

|Reserves|=6

<u>sid</u>	<u>bid</u>	<u>day</u>
28	103	12/4/06
28	103	11/3/06
31	101	10/10/06
31	102	10/12/06
31	101	10/11/06
58	103	11/12/06

×

|Sailors|=5

<u>sid</u>	sname	rating	age
22	dustin	7	45.0
28	yuppy	9	35.0
31	lubber	8	55.5
44	guppy	5	35.0
58	rusty	10	35.0

Εάν εκτελέσουμε το **Καρτεσιανό Γινόμενο** μεταξύ των δυο αυτών σχέσεων τότε μας επιστρέφεται ένα **αποτέλεσμα** το οποίο **δεν έχει και ιδιαίτερο νόημα**.

Συγκεκριμένα, επιστρέφονται τριάντα (30) 8-άδες:

28, 103, 12/4/06, 22, dustin, 7, 45.0

28, 103, 12/4/06, 28, yuppy, 9, 35.0

28, 103, 12/4/06, 31, lubber, 8, 55.5

28, 103, 12/4/06, 44, guppy, 5, 35.0

28, 103, 12/4/06, 58, rusty, 10, 35.0

....

28, 103, 11/3/06, yuppy, 22, dustin, 7, 45.0

...