

Βάσεις Δεδομένων

Ενότητα # 6: Σχεσιακή Άλγεβρα
Παραδείγματα

Γεώργιος Ευαγγελίδης, Καθηγητής
Τμήμα Εφαρμοσμένης Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Μακεδονίας» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Μέρος 1

Σχεσιακή Άλγεβρα

Γλώσσες αιτημάτων

- Χρειάζονται για τη διαχείριση και ανάκτηση δεδομένων από μια βάση δεδομένων.
- Το σχεσιακό μοντέλο έχει να κάνει με σύνολα.
- Χρειαζόμαστε μια μαθηματική γλώσσα χειρισμού συνόλων = **Σχεσιακή Άλγεβρα**.
- Απλό μοντέλο => απλή αλλά ισχυρή γλώσσα => επιδέχεται βελτιστοποιήσεων.
- Απαραίτητη για την κατανόηση της SQL.

Εισαγωγικά

- Οι πράξεις της σχεσιακής άλγεβρας έχουν ως είσοδο σχεσιακούς πίνακες και ως έξοδο σχεσιακό πίνακα.
- Τα σχήματα των συμμετεχόντων σχέσεων (πινάκων) καθώς και του αποτελέσματος παραμένουν σταθερά.
- Δεν μας ενδιαφέρει το περιεχόμενο των στιγμιοτύπων τους.
- Μπορεί το αποτέλεσμα να είναι ο κενός πίνακας.

Βασικές πράξεις

- **Προβολή** (π) φιλτράρει στήλες.
- **Επιλογή** (σ) φιλτράρει γραμμές.
- **Γινόμενο** (\times) συνδυάζει δυο πίνακες.
- **Διαφορά** ($-$) όπως στα σύνολα.
- **Ένωση** (\cup) όπως στα σύνολα.

Προβολή (π)

Ποια είναι όλα τα καταχωρισμένα έτη κυκλοφορίας cd;

$\pi_{\text{year}}(\text{CD})$

- Γενική μορφή $\pi_{\text{λίστα-πεδίων}}$ (πίνακας).
- Ποιο είναι το κλειδί του αποτελέσματος;
- Οι διπλοεγγραφές αφαιρούνται – το αποτέλεσμα πρέπει να είναι σχεσιακός πίνακας.
- Σχήμα και πληθυσμός αποτελέσματος;

Επιλογή (σ)

Ποια cd κυκλοφόρησαν το 1965;

$\sigma_{\text{year}=1965}$ (CD)

- Γενική μορφή $\sigma_{\text{συνθήκες}}$ (πίνακας).
- Μπορεί να έχουμε σύνθετες συνθήκες με τη χρήση AND, OR, NOT.
- Σχήμα και πληθυσμός αποτελέσματος;

Γινόμενο (x)

Συνδυάζει όλες τις γραμμές ενός πίνακα με όλες τις γραμμές ενός άλλου πίνακα.

A	
π1	π2
κ1	λ1
κ2	λ2

x

A x B

Σχήμα και πληθυσμός αποτελέσματος;

B		
π3	π4	π5
μ1	ν1	ξ1
μ2	ν2	ξ2
μ3	ν3	ξ3
μ4	ν4	ξ4

=

$A \times B = B \times A$;

Έχει σημασία η σειρά των γραμμών;

... η σειρά των στηλών;

A x B				
π1	π2	π3	π4	π5
κ1	λ1	μ1	ν1	ξ1
κ1	λ1	μ2	ν2	ξ2
κ1	λ1	μ3	ν3	ξ3
κ1	λ1	μ4	ν4	ξ4
κ2	λ2	μ1	ν1	ξ1
κ2	λ2	μ2	ν2	ξ2
κ2	λ2	μ3	ν3	ξ3
κ2	λ2	μ4	ν4	ξ4

Διαφορά και Ένωση

- Αν CD1 είναι ο πίνακας με τα δικά μου cd και CD2 με τα cd ενός φίλου μου:

Ποια cd έχω εγώ αλλά όχι ο φίλος μου; **CD1 – CD2**

Ποια cd έχουμε και οι δυο μαζί; **CD1 U CD2**

- Οι πίνακες συμβατοί ως προς την ένωση.
- Τί γίνεται με τις διπλοεγγραφές;
- Σχήμα και πληθυσμός αποτελέσματος;

Σύνθεση πράξεων

- Όπως στις μαθηματικές πράξεις $f(g(h(x)))$.

Ποιοι είναι οι τίτλοι των cd που έχω εγώ αλλά όχι ο φίλος μου που κυκλοφόρησαν το 2010;

$\pi_{\text{title}} (\sigma_{\text{year}=2010} (\text{CD1} - \text{CD2}))$

Μετονομασία (ρ)

- Πρόκειται για βοηθητική πράξη.
- Όταν συνδυάζω πίνακες (π.χ., με γινόμενο), τι κάνω αν έχω πεδία με την ίδια ονομασία στους αρχικούς πίνακες;

CD1 x (ρ_{id2/id, title2/title} (CD2))

- Γενική μορφή:

ρ_{νέο_όνομα_πεδίου/παλιό_όνομα_πεδίου} (πίνακας)

Επιπλέον σύνθετες πράξεις (1)

- **Τομή** (\cap) όπως στα σύνολα.

$$A \cap B = A - (A - B) = B - (B - A)$$

$$A \cap B = A \cup B - (A - B) - (B - A)$$

- **Σύζευξη** = επιλογή πάνω σε γινόμενο:

$$A \times_{\text{συνθήκη}} B = \sigma_{\text{συνθήκη}} (A \times B)$$

- Αποκτά νόημα το γινόμενο καθώς επιλέγουμε ποιες γραμμές θα ανήκουν στο αποτέλεσμα του.

Επιπλέον σύνθετες πράξεις (2)

- **Φυσική σύζευξη** = όταν η επιλογή είναι συνθήκη ισότητας ανάμεσα στα κοινά πεδία των πινάκων, πρακτικά κύριο κλειδί με ξένο κλειδί.

π.χ., αν $A(\underline{a1}, a2)$ και $B(\underline{b1}, b2, b3)$ και $b3$ είναι ξένο κλειδί και αναφέρεται στο $a1$:

$$A \bowtie B = \sigma_{A.a1=B.b3} (A \times B)$$

- Μόνο ένα αντίγραφο της κοινής στήλης μένει στο αποτέλεσμα της σύζευξης.

Επιπλέον σύνθετες πράξεις (3)

- **Φυσική σύζευξη:** πολύ χρήσιμη πράξη καθώς συνδυάζει γραμμές πινάκων μεταξύ τους μέσω των τιμών κλειδιού/ξένου κλειδιού.
- Έστω: CD(id, title, year, comid) COMPANY(id, name)
- Ποιες εταιρίες κυκλοφόρησαν cd το 2013;

$\pi_{\text{company.name}} ((\sigma_{\text{year}=2013} (\text{CD})) \times \text{COMPANY})$

ή

$\pi_{\text{company.name}} (\sigma_{\text{cd.comid}=\text{company.id AND year}=2013} (\text{CD} \times \text{COMPANY}))$

Επιπλέον σύνθετες πράξεις (4)

- Διαίρεση (/) δυαδική πράξη.

$$A / B$$

Προϋπόθεση: σχήμα(B) \subset σχήμα(A).

$$\text{σχήμα}(A / B) = \text{σχήμα}(A) - \text{σχήμα}(B).$$

Ερμηνεία αποτελέσματος: όλες οι διακριτές εγγραφές του A αφού κρατήσουμε μόνο τα μη κοινά πεδία του σε σχέση με τον B **που συνδυάζονται με όλες** τις εγγραφές του B.

Επιπλέον σύνθετες πράξεις (5)

- Παράδειγμα διαίρεσης A / B .
- Μόνο οι εγγραφές (χ, ψ) και (κ, λ) του πίνακα $A(\pi_1, \pi_3)$ συνδυάζονται με **όλες** τις εγγραφές του B . Οι (α, β) και (γ, δ) συνδυάζονται με μερικές.

A					B			A / B	
π_1	π_2	π_3	π_4		π_2	π_4		π_1	π_3
χ	1	ψ	2	/	1	2	=	χ	ψ
κ	1	λ	2		3	4		κ	λ
χ	3	ψ	4						
α	3	β	4						
γ	1	δ	2						
κ	3	λ	4						

Μέρος 2

Παραδείγματα

Στοιχεία Πινάκων

- Έστω το σχήμα βάσης δεδομένων:

RECORDING(rid, title, year) ή R

PERFORMER(pid, name) ή P

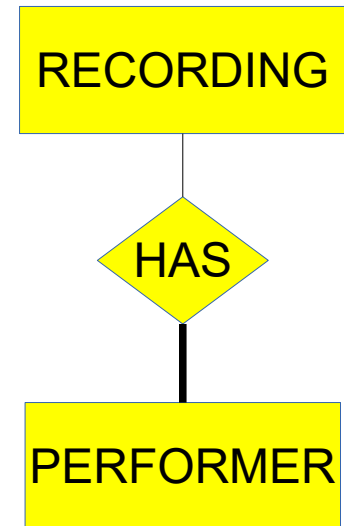
HAS(rid, pid) ή H

Όπου:

RECORDING έχει 100 εγγραφές

PERFORMER έχει 10 εγγραφές

HAS έχει 200 εγγραφές



Προβολές/Επιλογές

$\Pi_{\text{rid, title, year}} (R)$ μας δίνει τον πίνακα R.

$\sigma_{1=1} (R)$ μας δίνει τον πίνακα R.

$\sigma_{1=2} (R)$ μας δίνει τον κενό πίνακα (\emptyset).

$\sigma_{\text{name}='Lemmy'} (P)$ έχει 0 ως 10 εγγραφές.

$\sigma_{\text{pid}='777'} (P)$ έχει 0 ως 1 εγγραφές.

$\Pi_{\text{name}} (P)$ έχει 1 ως 10 εγγραφές.

$\Pi_{\text{pid}} (P)$ έχει 10 εγγραφές.

R 100

P 10

H 200

Γινόμενο/Σύζευξη/Διαίρεση

$R \times H$ έχει 20000 εγγραφές.

$R \bowtie x H$ έχει 200 εγγραφές.

$H \bowtie x P$ έχει 200 εγγραφές.

$T = R \bowtie x (H \bowtie x P)$ έχει 200 εγγραφές.

$\pi_{pid, name, title} (\sigma_{name='Lemmy' \text{ AND } year=2013} (T))$.

$H \div \pi_{pid} (P)$ = οι ηχογραφήσεις που έχουν ερμηνευτεί από **όλους** του ερμηνευτές!

R 100
P 10
H 200

Ένωση

Ερμηνευτές (pid) που τραγούδησαν το 1970 ή το 2010:

$\Pi_{pid} (\sigma_{year=1970} (H |x| R)) \cup \Pi_{pid} (\sigma_{year=2010} (H |x| R))$

ή με χρήση σύζευξης:

$\Pi_{pid} (\sigma_{year=1970 \text{ OR } year=2010} (H |x| R))$

R 100
P 10
H 200

Τομή (1)

Ερμηνευτές (pid) που τραγούδησαν και το 1970 και το 2010:

$$\pi_{pid} (\sigma_{year=1970} (H \mid x \mid R)) \cap \pi_{pid} (\sigma_{year=2010} (H \mid x \mid R))$$

ή με χρήση σύζευξης θα πρέπει να συνδυάσω όλες τις ερμηνείες ενός συγκεκριμένου ερμηνευτή με τις ερμηνείες του ίδιου ερμηνευτή:

$$H2 = \rho_{pid2/pid, rid2/rid} (H)$$

$$\sigma_{pid=pid2} (H \times H2)$$

R 100

P 10

H 200

Τομή (2)

- Έστω στιγμιότυπο του H:

a 1
b 1
c 1
a 2

- Τότε $\sigma_{pid=pid2} (H \times H2)$:

a 1 a 1
a 1 b 1
a 1 c 1
b 1 a 1
b 1 b 1
b 1 c 1
c 1 a 1
c 1 b 1
c 1 c 1
a 2 a 2

R 100
P 10
H 200

Τομή (3)

- ... και για να ελέγξω το έτος ηχογράφησης:

$$H2 = \rho_{pid2/pid, rid2/rid} (H)$$

$$R2 = \rho_{rid2/rid, title2/title, year2/year} (R)$$

$$T1 = H \bowtie R$$

$$T2 = H2 \bowtie R2$$

$$ANSWER = \pi_{pid} (\sigma_{pid=pid2 \text{ AND } year=1970 \text{ AND } year2=2010} (T1 \times T2))$$

R 100

P 10

H 200