

Βάσεις Δεδομένων - 4

The Relational Model

Δρ. Χαράλαμπος Καρανίκας

Περιεχόμενα

- **Εισαγωγή** στο Σχεσιακό Μοντέλο
- **Ορισμοί:** Σχέση, Γνώρισμα, Κλειδί, Σχήμα Σχέσης, Πεδίο Ορισμού, Πλειάδα, Κατάσταση Σχέσης, Πληθυκός Αριθμός Σχέσης, Παραδείγματα
- **Χαρακτηριστικά** Σχεσιακού Μοντέλου: Διάταξη Πλειάδων, Διάταξη Γνωρισμάτων, Τιμές & Κενές Τιμές
- **Περιορισμοί** Σχεσιακού Μοντέλου και **Σχήματα**
- **Πράξεις Ενημερώσεων** και Αντιμετώπιση Παραβιάσεων των **Περιορισμών**

Οι διαφάνειες βασίζονται:

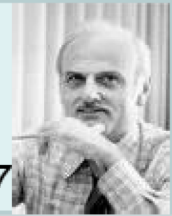
- Βασικές Αρχές Συστημάτων Διαχείρισης Βάσεων Δεδομένων, Gillenson Mark, 2018
- Θεμελιώδεις αρχές συστημάτων βάσεων δεδομένων, Elmasri Ramez, Navathe Shamkant B.
- Εισαγωγή στις Βάσεων Δεδομένων, Δημήτρης Ζεϊναλιπούρ, , Τμήμα Πληροφορικής Πανεπιστήμιο Κύπρου

Εισαγωγή στο Σχεσιακό Μοντέλο

- Μέχρι τώρα είδαμε **Μοντέλο Οντοτήτων-Συσχετίσεων (ER Model)**, το οποίο χρησιμοποιείται για την **εννοιολογική αναπαράσταση** δεδομένων με χρήση αφηρημένων εννοιών όπως **Οντότητες, Γνωρίσματα και Συσχετίσεις**
- Παρόλο που ένα τέτοιο μοντέλο αναπαριστάται **διαγραμματικά** και γίνεται **εύκολα αντιληπτό από τους σχεδιαστές και πελάτες**, δεν είναι ακόμη προφανές πως γίνεται «αντιληπτό» από ένα υπολογιστικό σύστημα.
- Σε αυτή την ενότητα θα δούμε πως το **Σχεσιακό Μοντέλο (Relational Model)**, το οποίο κάνει χρήση του μαθηματικού υποβάθρου των Σχέσεων για να απαντήσει το πιο πάνω ερώτημα.

Ιστορικά

- Το **Σχεσιακό Μοντέλο** ορίζεται το 1970 από τον Βρετανό ερευνητή Edgar F. Codd στο IBM Research
 - "A Relational Model of Data for Large Shared Data Banks".
"Codd, E.F.(1970), Communications of the ACM 13(6):377-387"
- Σε αυτό το μοντέλο, τα δεδομένα αναπαριστώνται σε **πίνακες** πάνω στους οποίους ορίζονται **διάφοροι περιορισμοί**. (Θα οριστεί αυστηρότερα στη συνέχεια)
- Το μοντέλο προκάλεσε μια **επανάσταση*** στο χώρο των βάσεων δεδομένων λόγω της **απλότητας** και του **μαθηματικού** του **υπόβαθρου**
 - **1969**: Το Σχεσιακό Μοντέλο υλοποιείται από τη βάση **IBM System R**
 - **1970**: Η IBM δημιουργεί την **SEQUEL** (προπομπό της **SQL**)
 - **1981**: Ο **Codd** παίρνει το **Turing Award** στη πληροφορική
 - **1985**: Η **IBM** κάνει την **SQL** Πατέντα (US Pat. 4,506,326).
 - **Σήμερα**: Το Σχεσιακό Μοντέλο υλοποιείται από τις περισσότερες σύγχρονες βάσεις δεδομένων (Oracle, IBM DB2, SQL Server, PostgreSQL, MySQL, κτλ).



Εισαγωγή στο Σχεσιακό Μοντέλο

- **Στόχος του Σχεσιακού Μοντέλου** είναι να προσφέρει μια **δηλωτική (declarative)** μέθοδο για τον ορισμό **δεδομένων** και **επερωτήσεων**
- Πάνω στο σχεσιακό μοντέλο στηρίζεται σήμερα η **Σχεσιακή Άλγεβρα (Relational Algebra)** (και κατ' επέκταση η SQL)

Σχεσιακό Μοντέλο: Άτυποι ορισμοί

Στο **Σχεσιακό Μοντέλο (Relational Model)** κάνει χρήση της μαθηματικής έννοιας της **Σχέσης (Relation)**

Τυπικά, μια **Σχέση r** είναι ένα σύνολο **πλειάδων (tuples)** $r = \{t_1, t_2, \dots, t_M\}$, όπου κάθε πλειάδα είναι μια **διατεταγμένη λίστα τιμών** $t = \langle v_1, v_2, \dots, v_N \rangle$.

Συνεπώς, μια σχέση είναι ουσιαστικά ένας **πίνακας $M \times N$** (**M γραμμών και N στηλών**)!

Κάθε γραμμή αναπαριστά συγκεκριμένα γεγονότα τα οποία αναφέρονται σε κάποια οντότητα (**entity**) ή συσχέτιση (**relationship**) του πραγματικού κόσμου.

- Κάθε γραμμή row ονομάζεται και πλειάδα (tuple) ή εγγραφή (record)
- Κάθε στήλη (column) ονομάζεται και χαρακτηριστικό (characteristic) ή γνώρισμα (attribute)

Παράδειγμα

Όνομα Σχέσης

Επικεφαλίδα

Γνωρίσματα

Εγγραφές

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25

Τυπικά, η **επικεφαλίδα** ΔΕΝ είναι μέρος της **σχέσης** αλλά περιλαμβάνεται στις **Μέτα-πληροφορίες** της.

Που αποθηκεύονται οι τέτοιες Μέτα-πληροφορίες σε με σχεσιακή βάση δεδομένων;

Σχεσιακό Μοντέλο: Άτυποι ορισμοί

Κλειδί (Key) μιας Σχέσης: Η ανάθεση τιμής ενός ή περισσότερων γνωρισμάτων τα οποία προσδιορίζουν μοναδικά την γραμμή στον πίνακα.

- Ο ορισμός είναι αντίστοιχος με αυτόν που δώσαμε στο πλαίσιο του ER όπου ορίσαμε ως το **Κλειδί μιας Οντότητας (Key)** το πεδίο(α) που αναγνωρίζουν μοναδικά μια Οντότητα
 - Στο **Σχεσιακό Μοντέλο** μια **σχέση** είναι είτε μια **Οντότητα** ή μια **Συσχέτιση** του ER Μοντέλου

Στο παράδειγμα του πίνακα STUDENT, το SSN είναι προφανώς το κλειδί.

Κάποιες φορές η θέση της γραμμής (row-id) ή σειριακοί αριθμοί χρησιμοποιούνται ως κλειδιά

- Ονομάζονται **Τεχνητά Κλειδιά (Artificial keys)** ή **Υποκατάστατα Κλειδιά (surrogate keys)**

Ορισμοί Σχεσιακού Μοντέλου

Το **Σχήμα (Schema)** ή περιγραφή μιας **Σχέσης**:

- Θα αναφέρεται ως **$R(A_1, A_2, \dots, A_n)$**
- **R** είναι το **όνομα** της σχέσης
- Τα **γνωρίσματα** της σχέσης είναι τα **$A_1, A_2, \dots, A_n$**

Παράδειγμα:

CUSTOMER (Cust-id, Cust-name, Address, Phone#)

- CUSTOMER δηλώνει το **όνομα** της σχέσης
- Η σχέση ορίζεται από τα **γνωρίσματα**: Cust-id, Cust-name, Address, Phone#

Κάθε γνώρισμα ορίζεται στο **πεδίο ορισμού (domain)**

- π.χ., το domain του Cust-id είναι 6 ψηφία (ακολουθεί ορισμός)

Ο **βαθμός (degree ή arity)** μιας σχέσης είναι το πλήθος **n** των γνωρισμάτων του σχήματος.

7-4

Ορισμοί Σχεσιακού Μοντέλου (n-άδα ή Πλειάδα)

Μια **πλειάδα (tuple)** είναι ένα **διατεταγμένο*** σύνολο τιμών $v_1, v_2, \dots, v_n$ το οποίο εμπερικλείεται στον ακόλουθο συμβολισμό ' $\langle \dots \rangle$ '

- Κάθε τιμή v_j ($1 \leq j \leq n$) της πλειάδας **ορίζεται** στο αντίστοιχο πεδίο **ορισμού** του συγκεκριμένου γνωρίσματος A_j ($1 \leq j \leq n$) ή είναι **κενό**.
- * εάν και θα χαλαρώσουμε αυτό τον περιορισμό αργότερα

Παράδειγμα: CUSTOMER (Cust-id, Cust-name, Address, Phone#)

Μια γραμμή της σχέσης CUSTOMER είναι μια 4-άδα η οποία αποτελείται από 4 τιμές, για παράδειγμα:

- $\langle 632895, \text{"John Smith"}, \text{"101 Main St. Atlanta, GA 30332"}, \text{"(404) 894-2000"} \rangle$

Μια σχέση είναι ένα **σύνολο (set)** τέτοιων πλειάδων.

Ορισμοί Σχεσιακού Μοντέλου (Πεδίο Ορισμού)

Το **Πεδίο Ορισμού (Domain – dom(Ai))** είναι το σύνολο των δυνατών **ατομικών** τιμών ενός γνωρίσματος (δηλ., δεν κάνει νόημα να διασπαστούν οι τιμές περαιτέρω), π.χ.,

Το Πεδίου Ορισμού έχει ένα **Λογικό Ορισμό (εξήγηση)**:

- Π.χ.,: “nicosia_phone_numbers”: *το σύνολο των αριθμών που είναι έξι ψηφία και ξεκινούν με 22.*

Επιπλέον, έχει **Όνομα**, **Τύπο** και **Μορφοποίηση**:

- **Τύπος Δεδομένων nicosia_phone_numbers** : Συμβολοσειρά
 - “22-892700”
- **Μορφοποίηση nicosia_phone_numbers** : Οι αριθμοί πρέπει να έχουν την μορφή 22-dddddd, όπου d είναι ένα δεκαδικό ψηφίο.

Τέλος, ίσως να απαιτείται και επιπλέον πληροφορία (π.χ., μονάδα μέτρησης σε ένα πεδίο **persons_weight: kg, g**)

Ορισμοί Σχεσιακού Μοντέλου (Κατάσταση Σχέσης)

Η **Κατάσταση ή Στιγμιότυπο Σχέσης (Relation State – $r(R)$)** είναι ένα υποσύνολο (ακόμη και κενό) του Καρτεσιανού Γινομένου των πεδίων ορισμού των γνωρισμάτων

- Συνεπώς, το $r(R)$ είναι ένα σύνολο πλειάδων (tuples) $r=\{t_1, t_2, \dots, t_M\}$, όπου κάθε πλειάδα είναι μια διατεταγμένη ακολουθία τιμών $t=\langle v_1, v_2, \dots, v_N \rangle$.
- Κάθε τιμή v_j είναι ένα στοιχείο του $\text{dom}(A_j)$ ή NULL

Το $r(R)$ ονομάζεται συχνά και **έκταση σχέσης (relation extension)** και το σχήμα R μιας σχέσης ως **πρόθεση σχέσης (relation intension)**

Πληθυκός Αριθμός Σχέσης (Cardinality) ($|r(R)|$): Ο αριθμός των πλειάδων μιας σχέσης.

- δηλ., ο αριθμός γραμμών του πίνακα της σχέσης

Παράδειγμα

Στιγμιότυπο $r(\text{STUDENT})$

Όνομα Σχέσης
STUDENT

Γνωρίσματα

Πλειάδες

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25

$r(\text{STUDENT})$

Βαθμός (Degree) = $|R| = 7$ (γνωρίσματα)

Πληθυκός Αριθμός (Cardinality) = $|r(\text{STUDENT})| = 5$ (πλειάδες)

Ορισμοί Σχεσιακού Μοντέλου (Σύνοψη)

$R(A_1, A_2, \dots, A_n)$ είναι το **σχήμα** της σχέσης.

R είναι το **όνομα** της σχέσης

$A_1, A_2, \dots, A_n$ είναι τα **γνωρίσματα** της σχέσης

– $\text{dom}(A_j)$: Το πεδίο ορισμού του A_j

$r(R)$: είναι μια συγκεκριμένη **κατάσταση (state)** της σχέσης R

– Αυτό είναι ένα σύνολο από πλειάδες (έγγραφες)

– $r(R) = \{t_1, t_2, \dots, t_m\}$ όπου κάθε t_i είναι μια n -άδα

• $t_i = \langle v_1, v_2, \dots, v_n \rangle$ όπου κάθε τιμή v_j ανήκει στο $\text{dom}(A_j)$

– Πληθυκός Αριθμός Σχέσης: $|r(R)| = m$

– Βαθμός Σχέσης: $|R| = n$

Ορισμοί Σχεσιακού Μοντέλου (Σύνοψη)

Έστω ότι $R(A_1, A_2)$ είναι ένα σχήμα σχέσης:

- Υποθέστε ότι $\text{dom}(A_1) = \{0,1\}$
- Υποθέστε επίσης ότι $\text{dom}(A_2) = \{a,b,c\}$

Τότε: $\text{dom}(A_1) \times \text{dom}(A_2)$ είναι όλοι δυνατοί συνδυασμοί:

$\{ \langle 0,a \rangle, \langle 0,b \rangle, \langle 0,c \rangle, \langle 1,a \rangle, \langle 1,b \rangle, \langle 1,c \rangle \}$

Η κατάσταση της σχέσης είναι υποσύνολο του καρτεσιανού γινομένου των πεδίων ορισμού $r(R) \subseteq \text{dom}(A_1) \times \text{dom}(A_2)$

- Π.χ., $r(R)$ μπορεί να είναι $\{\}$ ή $\{\langle 0,b \rangle\}$ ή $\{\langle 0,a \rangle, \langle 0,b \rangle, \langle 1,c \rangle\}$ ή οποιοδήποτε άλλο στοιχείο του Δυναμοσυνόλου (PowerSet) του $\text{dom}(A_1) \times \text{dom}(A_2)$
- Το **Δυναμοσύνολο (powerset) $P(A)$** ενός σύνολο A είναι το σύνολο όλων των υποσυνόλων του A .
 - Αν $A=\{a,b\}$ τότε $P(A)=\{\{\}, \{a\}, \{b\}, \{a,b\}\}$
- Ο **πληθυκός αριθμός του δυναμοσυνόλου $|P(A)|=2^{|A|}$**
- Συνεπώς υπάρχουν $2^6=64$ διαφορετικά στιγμιότυπα του $r(R)$

7-15

Ορισμοί Σχεσιακού Μοντέλου (Σύνοψη)

Πιο κάτω φαίνεται η αντιστοίχιση της **άτυπης** ορολογίας προς την **επίσημη** ορολογία

<u>Άτυποι Όροι</u>	<u>Επίσημοι Όροι</u>
Πίνακας	Σχέση (Relation)
(Κεφαλίδα) Στήλης	Γνώρισμα (Attribute)
Όλες οι δυνατές τιμές μιας στήλης	Πεδίο Ορισμού (Domain)
Γραμμή	Πλειάδα (Tuple)
Δήλωση Πίνακα	Σχήμα Σχέσης (Schema)
Δεδομένα Πίνακα	Στιγμιότυπο Σχέσης (State)

Χαρακτηριστικά των Σχέσεων - Διάταξη Πλειάδων

1. Διάταξη των πλειάδων σε μια σχέση $r(R)$:

- Οι πλειάδες ΔΕΝ είναι **διατεταγμένες** ή ταξινομημένες, παρόλο ίσως να παρουσιάζονται έτσι.
- Μια σχέση είναι ουσιαστικά ένα **σύνολο πλειάδων**.
- Τα σύνολα εξ' ορισμού δεν είναι διατεταγμένα.
 - **Ορισμός Συνόλου (Cantor 1895)**: Μια συλλογή ή ομάδα αντικειμένων ή στοιχείων ή μελών
- Η **διάταξη** ΔΕΝ αποτελεί μέρος του ορισμού μιας **σχέσης** αφού μια σχέση προσπαθεί να αναπαραστήσει κάποια γεγονότα σε **λογικό επίπεδο**.

STUDENT1

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25

$$r(\text{Student1}) = r(\text{Student2})$$

STUDENT2

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21

Χαρακτηριστικά των Σχέσεων – Διάταξη Γνωρισμάτων

2. Διάταξη Γνωρισμάτων σε μια Πλειάδα:

- Ο ορισμός της πλειάδας που χρησιμοποιήσαμε θεωρεί ότι τα γνωρίσματα είναι διατεταγμένα.
 - «Πλειάδα είναι ένα διατεταγμένο σύνολο τιμών $v_1, v_2, \dots, v_m \dots$ »)
- Τώρα θα **χαλαρώσουμε** αυτό τον περιορισμό αφού σε λογικό επίπεδο η **διάταξη των πλειάδων** δεν παίζει **σημαντικό** ρόλο, π.χ.,
 - Student (Name, **SSN**, Home_Phone, Address,...)
 - Student (**SSN**, Name, Home_Phone, Address,...)
- Συνεπώς, θα θεωρούμε ότι μια **πλειάδα** είναι ένα σύνολο από ζεύγη της μορφής (<γνώρισμα>, <τιμή>) όπου κάθε τιμή ορίζεται στο **dom(γνώρισμα)**.

Χαρακτηριστικά των Σχέσεων - Τιμές και NULL Τιμές

3. Τιμές και NULL Τιμές στις Πλειάδες:

- Κάθε **τιμή** σε μια **πλειάδα** είναι μια **ατομική τιμή** η οποία ορίζεται στο Πεδίο Ορισμού του εν λόγω γνωρίσματος
- Κάθε **Πλειότιμη Τιμή (Multi-valued Attribute)** αναπαριστάται από μια **επί μέρους σχέση**
 - Π.χ., Γνώρισμα «Locations» σε μια σχέση Department
- Κάθε **Σύνθετη Τιμή (Composite attribute)** αναπαριστάται από τα **απλά συστατικά τους**.
 - Π.χ., Γνώρισμα «Name» αναπαριστάται στη σχέση από “Fname”, “Minit” και “Lname”

Χαρακτηριστικά των Σχέσεων - Τιμές και NULL Τιμές

- **NULL Τιμές**
 - Τιμές που δεν είναι **γνωστές**
 - Π.χ., δεν καταχωρήσαμε το Office_phone κάποιου Employee
 - Τιμές που μπορεί να **μην ισχύουν** για κάποια πλειάδα
 - Π.χ., Ένας Υπάλληλος δεν δουλεύει σε γραφείο αλλά σε κάποιο άλλο χώρο (και δεν έχει γραφείο)

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21

- Γιατί μας απασχολούν οι NULL τιμές;
 - Γιατί **σπαταλούν χώρο**.
 - Γιατί η **σύγκριση** τους μπορεί να οδηγήσει σε **ασάφειες**.
 - Π.χ., Βρες όλους τους υπαλλήλους που έχουν το ίδιο τηλέφωνο.
 - Οι σχεσιακές βάσεις υλοποιούν μια τριαδική λογική: {**TRUE**, **FALSE**, **UNKNOWN**} αντί δυαδικής {**TRUE**, **FALSE**} στους τελεστές σύγκρισης για να λύσουν αυτό το πρόβλημα,
- Κατά τον σχεδιασμό βάσεων προσπαθούμε να **αποφεύγουμε** τις NULL τιμές όσο το δυνατό περισσότερο.

Σχεσιακοί Κανόνες Ακεραιότητας (Relational Integrity Constraints)

Μέχρι τώρα είδαμε διάφορους **ορισμούς** και τα **χαρακτηριστικά** τα οποία αφορούν μια **Σχέση**.

Σε μια σχεσιακή βάση ωστόσο υπάρχουν **πολλές Σχέσεις**

- Θυμηθείτε τις **οντότητες** και τις **συσχετίσεις** του **ER Μοντέλου**, οι οποίες αναπαριστώνται και οι δυο σαν Σχέσεις (Πίνακες) στο Σχεσιακό Μοντέλο.

Σε μια σχεσιακή βάση υπάρχουν επίσης **πολλοί τύποι περιορισμών**, οι οποίοι αναφέρονται

- Σε μια σχέση (π.χ., περιορισμός κλειδιού, πεδίου ορισμού και οντότητας)
- Σε πολλαπλές σχέσεις (π.χ., περιορισμός αναφορικής ακεραιότητας)

Σχεσιακοί Κανόνες Ακεραιότητας (Relational Integrity Constraints)

- **Σχεσιακοί Περιορισμοί (Relational Constraints)** είναι συνθήκες οι οποίες πρέπει να ισχύουν για κάθε έγκυρη κατάσταση σχέσης (state).
- Στο σχεσιακό μοντέλο υπάρχουν τρεις τύποι περιορισμών:
 - A. Κλειδιού (Key constraints)**
 - Άτυπα: Κάθε σχέση έχει ένα πρωτεύων κλειδί.
 - B. Ακεραιότητας Οντοτήτων (Entity integrity constraints)**
 - Άτυπα: Το πρωτεύων Κλειδί δεν μπορεί να είναι NULL
 - C. Αναφορικής Ακεραιότητας (Referential integrity constraints)**
 - Άτυπα: Εάν μια πλειάδα A αναφέρεται σε άλλη πλειάδα B τότε η B πρέπει να υπάρχει.
- Τέλος, υπάρχει και ο **Περιορισμός Πεδίου Ορισμού (Domain Constraint)** ο οποίος εξυπακούεται.
 - Η τιμή κάθε πλειάδας πρέπει να ορίζεται στο αντίστοιχο της πεδίο ορισμού (ή μπορεί να είναι NULL, εάν επιτρέπεται για το εν λόγω ννώρισμα).

Σχεσιακοί Κανόνες Ακεραιότητας

Περιορισμοί Κλειδιού (Key Constraints)

- **Κλειδί Σχέσης (Relation Key):** Ο ελάχιστος αριθμός γνωρισμάτων που προσδιορίζει μοναδικά μια πλειάδα κάποιας σχέσης R.

- Π.χ., SSN ή StudentID για Student.

Εάν μια σχέση έχει πολλαπλά **Εναλλακτικά Κλειδιά (candidate keys)**, τότε ένα από αυτά επιλέγεται από τον σχεδιαστή ως το **Πρωτεύων Κλειδί (Primary key)**.

- Τα γνωρίσματα ενός πρωτεύων κλειδιού υπογραμμίζονται.

A) Κανόνας Πρωτεύοντος Κλειδιού: Κάθε Σχέση έχει ένα (1) Πρωτεύων Κλειδί.

Παράδειγμα: Θεωρήστε την Σχήμα Σχέσης CAR :

- **CAR(State, Reg#, SerialNo, Make, Model, Year)**
- Το πρωτεύων κλειδί SerialNo χρησιμοποιείται
 - A) Για να ξεχωρίζει μοναδικά κάθε πλειάδα στη σχέση (η ταυτότητα της)
 - B) Για να αναφερόμαστε σε άλλες σχέσεις από μια πλειάδα.

Γενικός Κανόνας: Επιλέξτε ως πρωτεύων κλειδί το **μικρότερο** εκ' των εναλλακτικών κλειδιών (σε άποψη μεγέθους).

Σχεσιακοί Κανόνες Ακεραιότητας

Ακεραιότητα Οντοτήτων (Entity Integrity)

B) Κανόνας Ακεραιότητας Οντοτήτων (Entity Integrity) ορίζει ότι τα γνωρίσματα ενός πρωτεύοντος κλειδιού **PK** ΔΕΝ μπορεί να είναι **NULL**

- δηλ., τιμή πρωτεύοντος κλειδιού **t[PK] ≠ NULL** για οποιαδήποτε πλειάδα **t** που ανήκει στο **r(R)**
- Εάν το **PK** είναι σύνθετο, τότε το **NULL** δεν επιτρέπεται σε **KANENA** από τα γνωρίσματα του **PK**
- Π.χ., **WORKS_ON(ESSN, PNO, Hours)** // PK είναι σύνθετο
 - (123, 4, 5) OK
 - (123, 4, NULL) OK
 - (NULL, 4, 5) ERROR

2-2

Σχεσιακοί Κανόνες Ακεραιότητας

Αναφορική Ακεραιότητα (Referential Integrity)

- Θεωρήστε την **N:1** συσχέτιση μεταξύ **Employee-Department** (χωρίς περιορισμούς συμμετοχής)



- Σε **N:1** συσχετίσεις το κλειδί της οντότητας στη πλευρά του **1** γίνεται **γνώρισμα** στη πλευρά του **N**, συνεπώς:
 - Σχέση που **Αναφέρει (Referencing Relation)** → **Ξένο Κλειδί (Foreign Key)**
 - Π.χ., EMPLOYEE(ssn, name, **dno**)
 - Σχέση που **Αναφέρεται (Referenced Relation)** → **Πρωτεύων Κλειδί**
 - Π.χ., DEPARTMENT(**dno**, name, location)

Το **dno** της σχέσης **EMPLOYEE** ονομάζεται **ξένο κλειδί (foreign key)** και έχει τα ακόλουθα χαρακτηριστικά

- Α) Έχει το **ίδιο πεδίο ορισμού** με το αντίστοιχο πεδίο του **DEPARTMENT**
- Β) Έχει **τιμή** ή είναι **NULL** (εάν επιτρέπεται κάτι τέτοιο από το σχήμα)

Example of Foreign Key

EMPLOYEE

EMPNO	ENAME	JOB	MGR	HIREDATE	SAL	DEPTNO
7369	SMITH	CLERK	7902	17-DEC-80	800	20
7521	ALLEN	SALESMAN	7698	20-FEB-81	1600	30
7566	WARD	SALESMAN	7698	22-FEB-81	1250	30
7654	JONES	MANAGER	7839	02-APR-81	2975	20
7787	MARTIN	SALESMAN	7696	28-SEP-8	1250	30

DEPARTMENT

DEPTNO	DNAME	LOCATION
10	ACCOUNTING	NEW YORK
20	RESEARCH	DALLAS
30	SALES	CHICAGO

Foreign
Key



Σχεσιακοί Κανόνες Ακεραιότητας

Αναφορική Ακεραιότητα (Referential Integrity)

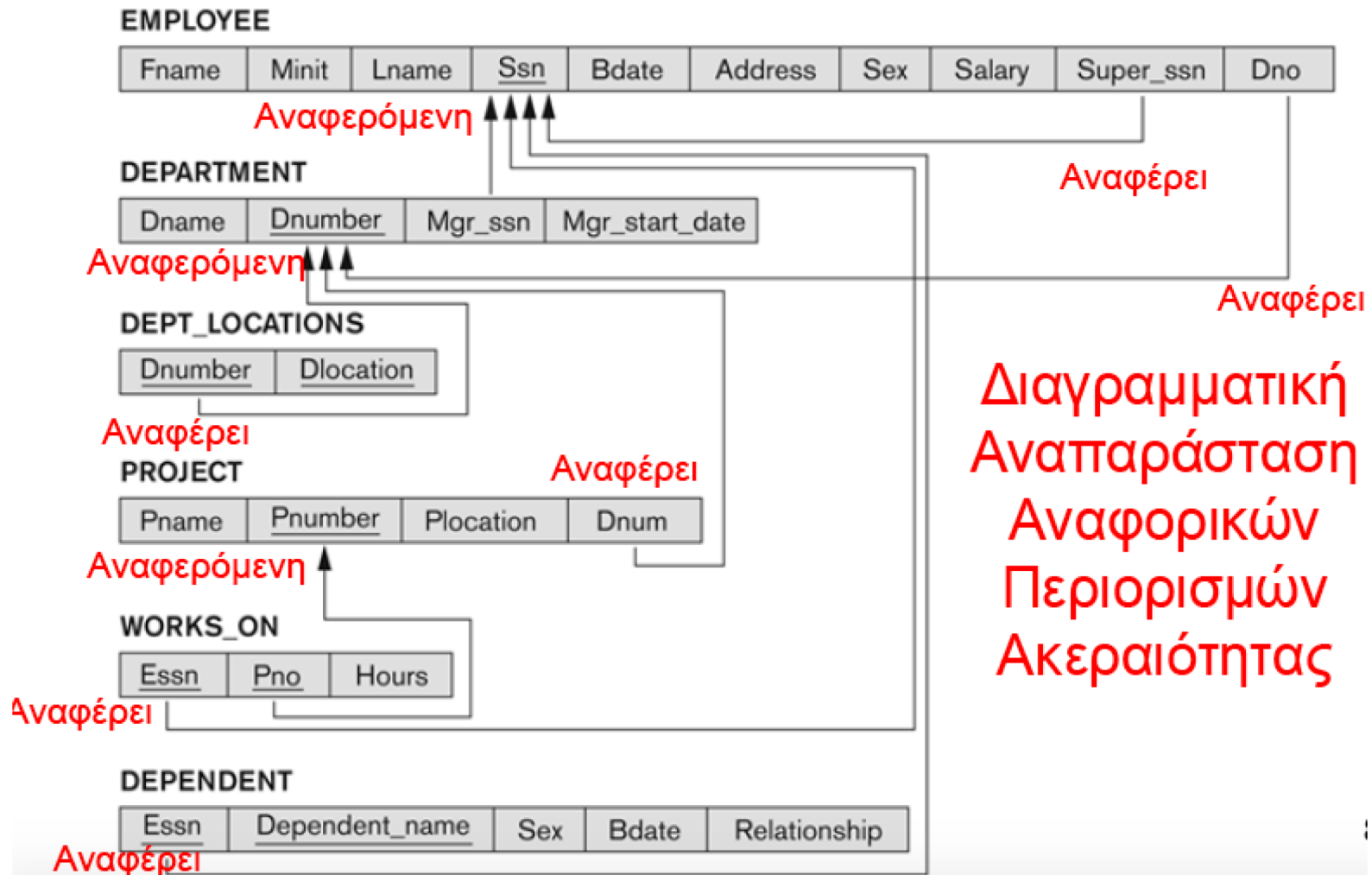
Γ) Κανόνας Αναφορικής Ακεραιότητας (Referential Integrity) ορίζει ότι τα ξένα κλειδιά (**FK**) πρέπει να είναι είτε υπαρκτά ή **NULL**

- Π.χ., EMPLOYEE(SSN, Fname, Lname, **Dno**)
 - (123, Costas, Ioannou, 5) **OK** (υπάρχει το 5)
 - (123, Costas, Ioannou, NULL) **OK** (εφόσον επιτρέπεται NULL)
 - (123, Costas, Ioannou, 1000) **ERROR** (ΔΕΝ υπάρχει το 1000)

Νοείται, ότι στη περίπτωση που το **FK** είναι μέρος του **PK** μιας σχέσης ($FK \subseteq PK$), τότε το **FK** δεν μπορεί να είναι **NULL** (κανόνας πρωτεύοντος κλειδιού)

- Π.χ., DEPENDENT(ESSN, Dname, Sex, Bdate, Relation),

Αναφορική Ακεραιότητα (Referential Integrity) Διαγραμματικά



Εισαγωγή στη Σχεσιακή Άλγεβρα

- Η θεωρία πίσω από τις πράξεις του σχεσιακού μοντέλου ορίζεται επισήμως και είναι ισοδύναμη με first-order logic.
- Η Σχεσιακή Άλγεβρα αποτελείται από ένα **βασικό σύνολο πράξεων (τελεστών)** για το σχεσιακό μοντέλο.
- Αυτοί οι **τελεστές** επιτρέπουν σε κάποιο 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.

Τύποι Τελεστών (Types of Operators)

Basic Operators - Relational Algebra

Selection
Projection
Union
Intersection
Difference
Cartesian Product
Division

Manipulative Operators

Insert
Delete
Update

Additional Operators

Outer Join
Outer Union
Recursive Join

Πράξεις Ενημερώσεων στο Σχεσιακό Μοντέλο (Manipulative Operators)

Οι **βασικές πράξεις** για αλλαγή της κατάστασης μιας βάσης δεδομένων είναι:

- ✓ **Insert, Εισαγωγή** νέας πλειάδας στη σχέση.
- ✓ **Delete, Διαγραφή** Υφιστάμενης(ων) Πλειάδας(ων)
- ✓ **Update, Τροποποίηση** πεδίων συγκεκριμένης(ων) πλειάδων

Τι γίνεται σε περιπτώσεις παραβίασης των κανόνων ακεραιότητας από μια πράξη;

A) (Default Πράξη) Ακύρωση πράξης. **NO ACTION**

B) Επιτρέπεται η πράξη (σπάνια) αλλά ενημερώνεται ο χρήστης

- Μπορεί να συμβεί σε άπλες περιπτώσεις μόνο (π.χ., εάν υπάρχει περιορισμός **TELEPHONE NOT NULL**).
- Δεν μπορεί να συμβεί σε περιπτώσεις κανόνων ακεραιότητας (κλειδιού, οντότητας, αναφοράς) και γενικά δεν υποστηρίζεται.

Γ) Εκτέλεση επιπλέον λειτουργιών. Π.χ., Στην SQL υπάρχει η έννοια του CASCADE, SET NULL, SET DEFAULT.

- Π.χ., **CREATE TABLE EMPLOYEE (....**
CONSTRAINT MGR
FOREIGN KEY(Super_ssn) REFERENCES EMPLOYEE(ssn)
ON DELETE SET NULL ON UPDATE CASCADE
- Εάν σβηστεί ο Supervisor τότε ο Emp.Super_ssn γίνεται NULL
- Εάν ενημερωθεί το SSN του Supervisor τότε ο Emp.Super_ssn (σχέσεων που την αναφέρουν) ενημερώνεται επίσης.

Insert Operator

DEPARTMENT

DEPTNO	DNAME	LOCATION
10	ACCOUNTING	NEW YORK
20	RESEARCH	DALLAS
30	SALES	CHICAGO
40	OPERATIONS	BOSTON
50	MARKETING	LOS ANGELES

INSERT (50, MARKETING, LOS ANGELES) into DEPARTMENT

INSERT - Παραβίαση Περιορισμών

Το **INSERT** μπορεί να παραβιάζει όλους τους περιορισμούς:

- Πεδίου Ορισμού:
 - π.χ., **EMPLOYEE(SSN:integer)**: εισαγωγή (123.3)
(πραγματικού αριθμού)
- Κλειδιού:
 - π.χ., **EMPLOYEE(SSN)**: εισαγωγή (123), όπου η σχέση **EMPLOYEE** περιέχει ήδη το 123, π.χ., {1,2,4,123,300}
- Αναφορικής Ακεραιότητας:
 - π.χ., **EMPLOYEE(SSN, DNO)**: εισαγωγή (123, 1000), θεωρώντας ότι το **DNO 1000** δεν υπάρχει.
- Οντότητας:
 - π.χ., **EMPLOYEE(SSN, Name)**: εισαγωγή (NULL, "Costas")

Delete Operator - Παραβίαση Περιορισμών

Το **DELETE** μπορεί να παραβιάζει τους ακόλουθους περιορισμούς:

- Πεδίου Ορισμού: **ΟΧΙ**
- Κλειδιού: **ΟΧΙ**
- Αναφορικής Ακεραιότητας:
 - Π.χ., εάν η εγγραφή που διαγράφεται αναφέρεται από άλλη σχέση.
 - Όπως αναφέραμε ήδη, η αναφορική ακεραιότητα μπορεί να επιβληθεί στην SQL με χρήση των επιπλέον ενεργειών: **CASCADE, SET NULL, SET DEFAULT**
- Οντότητας: **ΟΧΙ**

Delete Operator - Παραβίαση Περιορισμών (Παράδειγμα)

DEPARTMENT

DEPTNO	DNAME
10	ACCOUNTING
20	RESEARCH

EMPLOYEE

EMPLNO	ENAME	DEPTNO
1001	JOHN	10
1002	BILL	20

Delete (DEPTNO = 20)

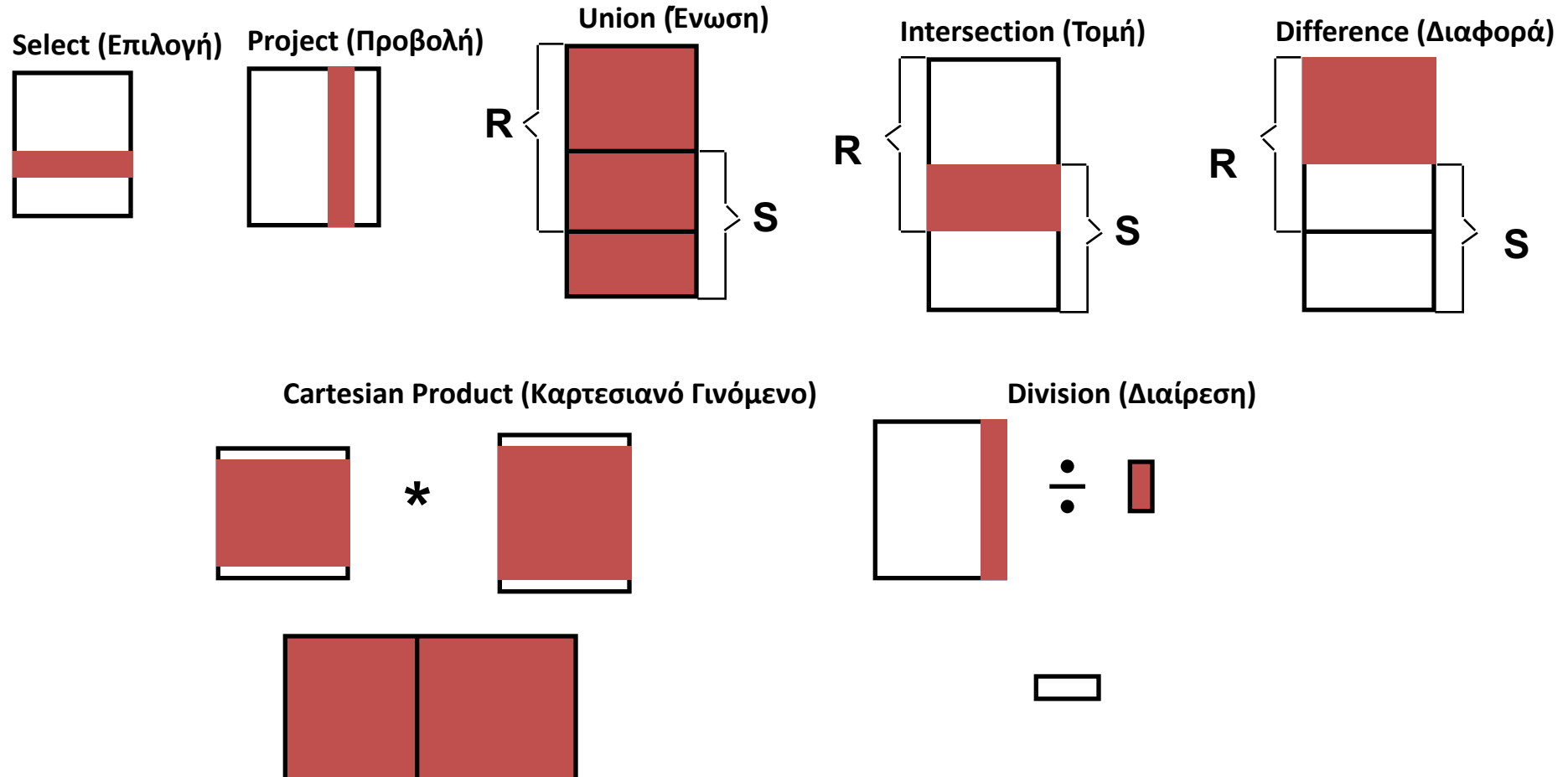
Παραβιάζει referential integrity

Update Operator - Παραβίαση Περιορισμών

Το **UPDATE** από→προς μπορεί να παραβιάζει τους ακόλουθους περιορισμούς:

- Πεδίου Ορισμού:
 - π.χ., **EMPLOYEE(SSN:integer)**: ενημέρωση 121→123.0 (πραγματικού αριθμού)
- Κλειδιού:
 - Π.χ., **EMPLOYEE(SSN)**: ενημέρωση 121→123, όπου η σχέση περιέχει ήδη το 123, π.χ., {1,2,4,123,300}
- Αναφορικής Ακεραιότητας:
 - π.χ., **EMPLOYEE(SSN, DNO)**: ενημέρωση (1,5)→(1,1000), θεωρώντας ότι το **DNO 1000** δεν υπάρχει ενώ το **5** υπάρχει.
- Οντότητας:
 - Π.χ., **EMPLOYEE(SSN)**: ενημέρωση 1→NULL

Βασικοί τελεστές - Σχεσιακή Άλγεβρα



SELECT Operator - Τελεστής Επιλογής

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

$$\sigma_{\langle \text{Κριτήρια Επιλογής} \rangle} (R)$$

- 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:

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

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

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

SELECT Operator - Τελεστής Επιλογής

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:**
 - Δυαδικός Τελεστής Σύγκρισης {<, >, =, !=, >=, <=}
- Οι προτάσεις μπορεί να συνδέονται με
 - **Λογικούς Τελεστές:** AND ή OR
 - Π.χ., $\sigma_{DNO = 4 \text{ AND } SALARY > 30,000}$ (EMPLOYEE)

Παράδειγμα SELECT Operator

EMPLOYEE

σ (**SALARY > £2,000**)

EMPNO	ENAME	JOB	MGR	SAL
7782	CLARK	MANAGER	7839	2450
7839	KING	PRESIDENT	NULL	5000
7934	MILLER	CLERK	7782	1300

EMPLOYEE

EMPNO	ENAME	JOB	MGR	SAL
7782	CLARK	MANAGER	7839	2450
7839	KING	PRESIDENT	NULL	5000

Ιδιότητες SELECT Operator

- **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

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

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:

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

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

Παράδειγμα SELECT Operator

EMPLOYEE

Δεδομένα

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

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

Αποτέλεσμα

Fname	Minit	Lname	Ssn	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

PROJECT Operator - Τελεστής Προβολής

- 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.

Ιδιότητες PROJECT Operator

• Ιδιότητες Τελεστή Προβολής π

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

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

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

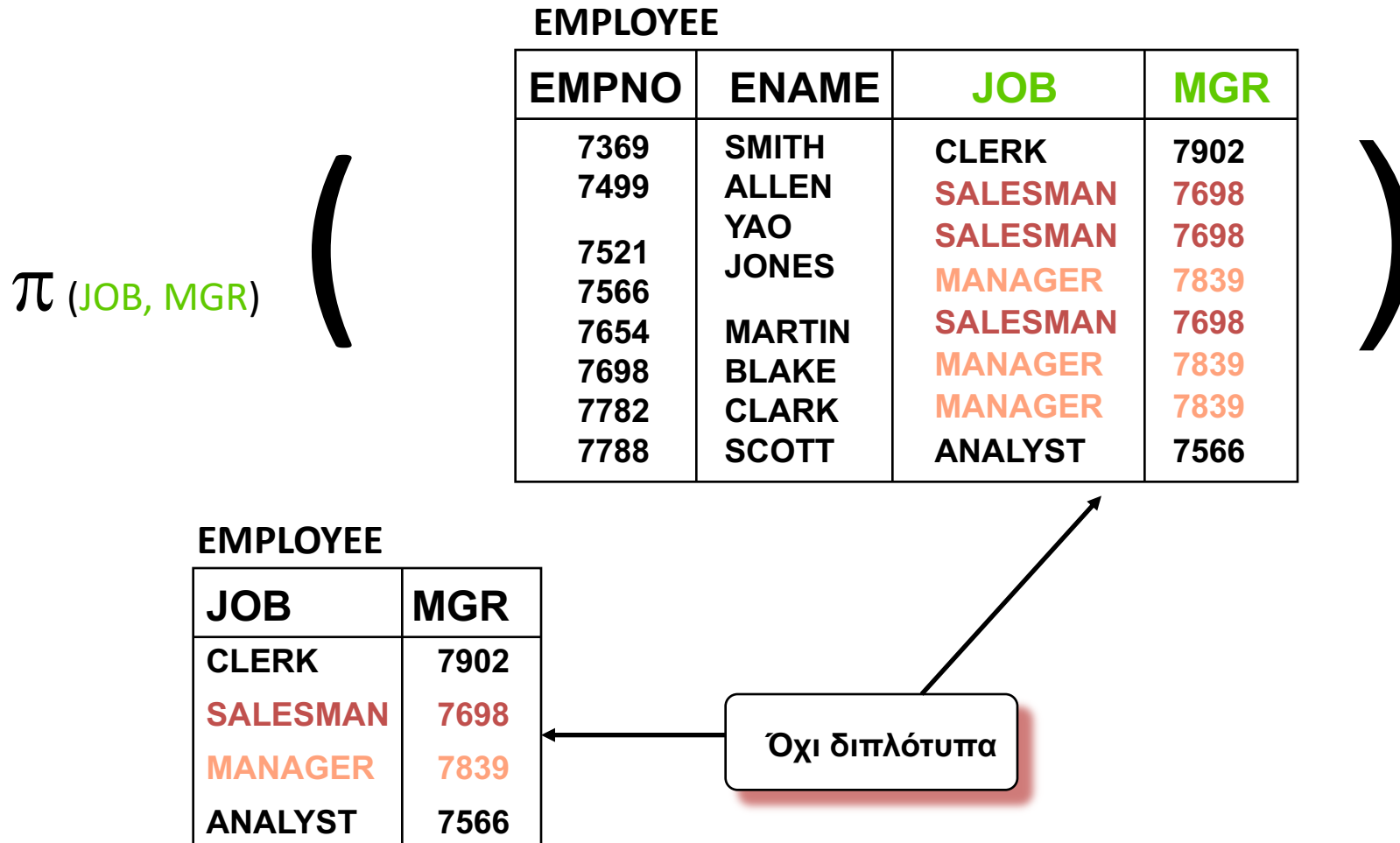
Π.Χ., $\pi_{\langle \text{ssn}, \text{fname}, \text{dno} \rangle}(\pi_{\langle \text{ssn}, \text{fname}, \text{lname}, \text{dno} \rangle}(\text{EMPLOYEE}))$ **ΣΩΣΤΟ**

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

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

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

Παράδειγμα PROJECT Operator



Εκφράσεις Σχεσιακής Άλγεβρας

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

Set Theoretic Operators - Τελεστές Θεωρίας συνόλων

- Union, Intersection, Difference
- Οι δύο Σχέσεις πρέπει να είναι συμβατές Συνένωσης
 - $|R| = |S|$ και
 - Ανταποκρινόμενα χαρακτηριστικά να έχουν το ίδιο πεδίο ορισμού

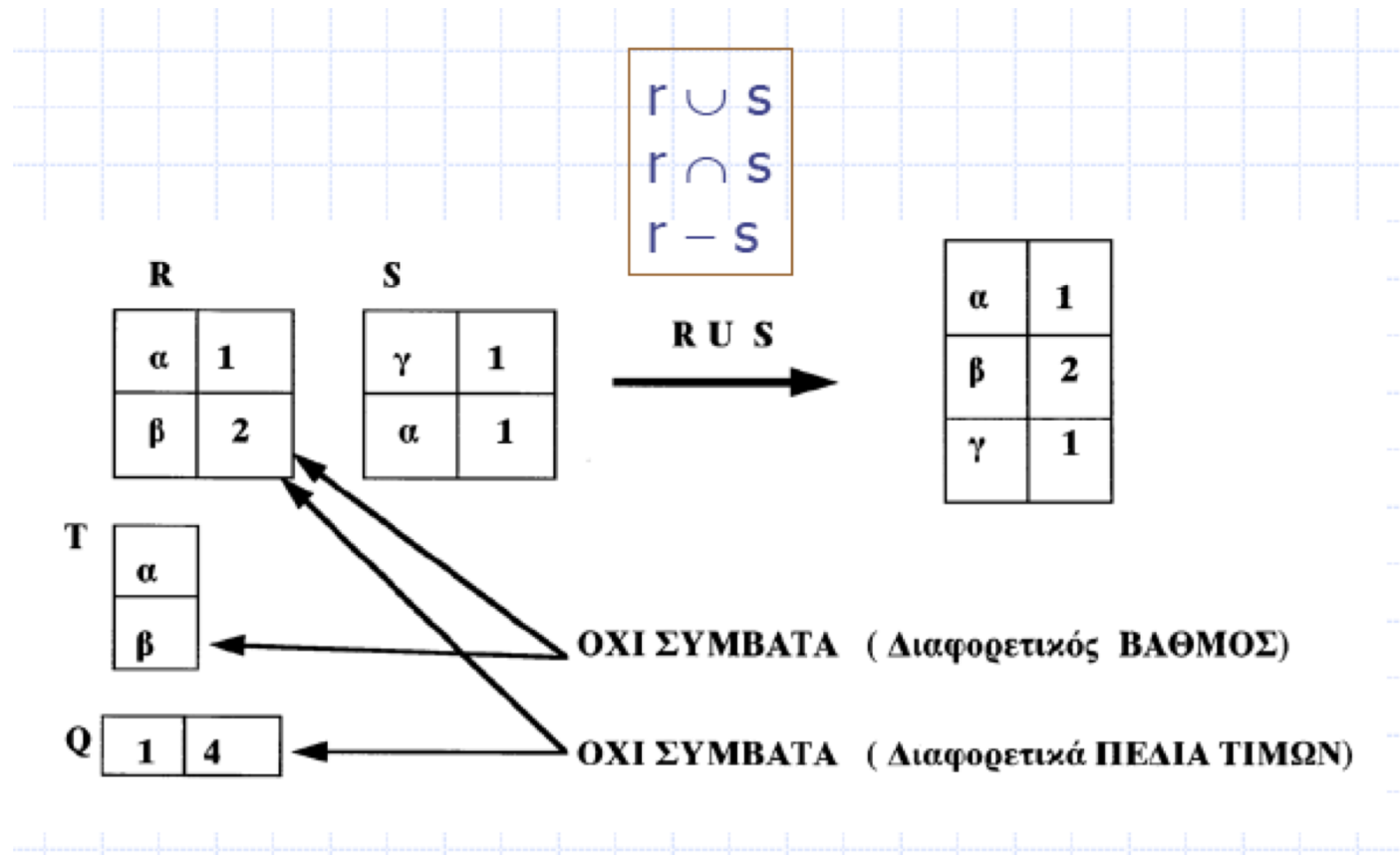
Δύο Σχέσεις $R_1(A_1, A_2, \dots, A_n)$ and $R_2(B_1, B_2, \dots, B_m)$ είναι συμβατές συνένωσης εάν:

$$n = m$$

and,

$$\text{dom}(A_i) = \text{dom}(B_i) \text{ for } 1 \leq i \leq n$$

Παράδειγμα συμβατότητας



UNION Operator - Τελεστής Ένωση

- **Τελεστής Ένωσης (UNION)**



- Είναι δυαδικός τελεστής 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

Παράδειγμα UNION Operator

EMPLOYEE

EMPNO	ENAME	SAL
7782	CLARK	2450
7839	KING	5000



MANAGER

EMPNO	ENAME	SAL
8034	SMITH	7000
8044	YAO	6000

STAFF

EMPNO	ENAME	SAL
7782	CLARK	2450
7839	KING	5000
8034	SMITH	7000
8044	YAO	6000

Παράδειγμα UNION Operator

- 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

Παράδειγμα UNION Operator

- 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

INTERSECTION Operator - Τελεστής Τομή

- **Τελεστής Τομής (Intersection)**



- Είναι δυαδικός τελεστής 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)
 - $(R \cap S) = \overline{(\overline{R} \cup \overline{S})}$ Κανόνας De Morgan μετατροπής.

Παράδειγμα INTERSECTION Operator

EMPLOYEE

EMPNO	ENAME	SAL
8034	SMITH	7000
8044	YAO	6000
7782	CLARK	2450
7839	KING	5000



MANAGER

EMPNO	ENAME	SAL
8034	SMITH	7000
8044	YAO	6000

MANAGER-EMPLOYEE

EMPNO	ENAME	SAL
8034	SMITH	7000
8044	YAO	6000

Παράδειγμα INTERSECTION Operator

- 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

Παράδειγμα INTERSECTION Operator

- Διατύπωση σε Σχεσιακή Άλγεβρα

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

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

$RESULT \leftarrow R1 \cap R2$

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

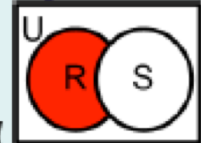
DEPENDENT

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

DIFFERENCE Operator - Τελεστής Διαφοράς

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

- Είναι δυαδικός τελεστής 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 $R - (S - T) = (R - S) - T$

Παράδειγμα DIFFERENCE Operator

EMPLOYEE

EMPNO	ENAM	SAL
7782	CLARK	2450
7839	KING	5000
8034	SMITH	7000
8044	YAO	6000

—

MANAGER

EMPNO	ENAME	SAL
8034	SMITH	7000
8044	YAO	6000

NON-MANAGER-EMPLOYEE

EMPNO	ENAME	SAL
7782	CLARK	2450
7839	KING	5000

CARTESIAN PRODUCT Operator - Καρτεσιανό Γινόμενο

- **Καρτεσιανό Γινόμενο ($\times$)**
 - **CARTESIAN (or CROSS) PRODUCT**
 - Το $R \times S$ είναι το σύνολο όλων των διατεταγμένων ζευγών (r,s) , όπου $r \in R$ και $s \in S$, δηλ.,
 - $R \times S = \{ (r,s) \mid r \in R \wedge s \in S \}$
 - 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., $R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$, μας επιστρέφει μια σχέση Q βαθμού $n+m$
 - $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$ (Η Q έχει $n*m$ πλειάδες)
 - Οι τελεσταίοι R, S δεν χρειάζεται να συμφωνούν στον τύπο. . .

Είναι όλοι οι δυνατοί συνδυασμοί των στοιχείων των δύο συνόλων

CARTESIAN PRODUCT Operator - Καρτεσιανό Γινόμενο

- 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**))

Παράδειγμα CARTESIAN PRODUCT

DEPT
DEPTNO
10
20
30

EMP
EMPNO
73
75
79

DEPT x EMP

DEPTNO	EMPNO
10	73
10	75
10	79
20	73
20	75
20	79
30	73
30	75
30	79

Παράδειγμα CARTESIAN PRODUCT - έννοια Συνένωσης Join

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>	<u>sname</u>	<u>rating</u>	<u>age</u>
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)

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

Παράδειγμα CARTESIAN PRODUCT - έννοια Συνένωσης Join

- Αυτό που θα θέλαμε στη πραγματικότητα, είναι να συνδυάσουμε τα Reserves με τους Sailors κατά τέτοιο τρόπο που να δημιουργείται **κάποια χρήσιμη πληροφορία**
- Χρήσιμο Ερώτημα: Βρες σε ποιο όνομα ανήκει κάθε κράτηση Reserve**

|Reserves|=6

sid	bid	day
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

sid	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 εκτελεστεί το **Reserves x Sailors** ακολουθούμενο από $\sigma_{\text{Reserves.sid} = \text{Sailors.sid}}$ τότε βρίσκουμε το επιθυμητό αποτέλεσμα

Το αποτέλεσμα θα αυτή την φορά μόνο έξι πλειάδες:

(28, 103, 12/4/06, yuppy, 9, 35.0)	(28, 103, 11/3/06, yuppy, 9, 35.0)
(31, 101, 10/10/06, lubber, 8, 55.5)	(31, 102, 10/12/06, lubber, 8, 55.5)
(31, 101, 10/11/06, lubber, 8, 55.5)	(58, 103, 11/12/06, rusty, 10, 35.0)

Συνένωση (JOIN)

Τελεστής Συνένωσης (JOIN $\bowtie$, $\otimes^*$)

- Είναι ένας **δυναμικός τελεστής** ο οποίος **συνδυάζει σχετιζόμενες** πλειάδες σχέσεων σε μία **ενιαία πλειάδα**.
- Τυπικά, ορίζεται ως το καρτεσιανό γινόμενο μεταξύ δυο σχέσεων ακολουθούμενο από επιλογή.

$$R \otimes_{\langle \text{Κριτήρια} \rangle} S = \sigma_{\langle \text{Κριτήρια} \rangle} (R \times S)$$

Παράδειγμα Συνένωσης (JOIN)

- **Επερώτηση:** Βρείτε το όνομα του manager κάθε department.

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

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	1941-08-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fine Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmed	Y	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

- a) Πρέπει να συνδυάσουμε κάθε πλειάδα του πίνακα **Department** με τις πλειάδες του πίνακα **Employee**.

R = (DEPARTMENT × EMPLOYEE)

- b) Μετά πρέπει να επιλέξουμε τις πλειάδες για τις οποίες ισχύει το **Mgr_ssn=SSN**, δηλ.,

$\sigma_{Mgr_ssn = SSN}(R)$

Το **a-b** ισοδυναμεί με την έκφραση:

DEPARTMENT ⋈_{Mgr_ssn = SSN} **EMPLOYEE**

DEPT_MGR **Αποτέλεσμα**

Dname	Dnumber	Mgr_ssn	...	Fname	Minit	Lname	Ssn	...
Research	5	333445555	...	Franklin	T	Wong	333445555	...
Administration	4	987654321	...	Jennifer	S	Wallace	987654321	...
Headquarters	1	888665555	...	James	E	Borg	888665555	...

Division - Διαίρεση

Ο τελεστής της **Διαίρεσης (Division, /)**, παρόλο που **ΔΕΝ** υποστηρίζεται σε **πραγματικές γλώσσες** βάσεων δεδομένων*, είναι χρήσιμος στο πλαίσιο της Σχεσιακής Άλγεβρας για να διατυπώσει επερωτήσεις της μορφής:

Βρες τους Suppliers που προσφέρουν ΌΛΑ τα Parts (όπως προσδιορίζονται τα Parts).

Π.χ.,

sno	pno
s1	p1
s1	p2
s1	p3
s1	p4
s2	p1
s2	p2
s3	p2
s4	p2
s4	p4

PARTS			
pno			sno
p2		=	s1
p4			s4

* Στην **συνέχεια** θα δούμε πως ο τελεστής **μπορεί να υλοποιηθεί** με χρήση υπαρκτών τελεστών της **SQL** (δηλ., αντίστοιχους τελεστές των **σ-π-χ**)

Παράδειγμα Division - Διαίρεση

Διαιρετέος (nominator)

sno	pno
s1	p1
s1	p2
s1	p3
s1	p4
s2	p1
s2	p2
s3	p2
s4	p2
s4	p4

A

Διαιρέτες (denominators)

pno
p2

B1

Απάντηση;

sno
s1
s2
s3
s4

A/B1

pno
p2
p4

B2

Απάντηση;

sno
s1
s4

A/B2

pno
p1
p2
p4

B3

Απάντηση;

sno
s1

A/B3

Αναπαράσταση Διαίρεσης με Ποσοδείκτες

$$A / B = \{ x \mid \exists(x,y) \in A \ \forall y \in B \}$$

Πληρότητα των Σχεσιακής Άλγεβρας

- SELECT, PROJECT, UNION, DIFFERENCE and CARTESIAN PRODUCT είναι οι βασικοί τελεστές της σχεσιακής άλγεβρας
- Οι επιπλέον τελεστές ορίζονται ως συνδυασμός δύο ή περισσότερων από τους βασικούς. Για παράδειγμα:
JOIN = CARTESIAN PRODUCT + SELECT
DIVISION = PROJECT + CARTESIAN PRODUCT + DIFFERENCE
A INTERSECT B = A DIFFERENCE (A DIFFERENCE B)