



Εισαγωγή στις Βάσεις Δεδομένων

)))))))))

Η Πληροφορία ως έννοια

Ορισμός εργασίας : η επικοινωνία είναι η ανταλλαγή πληροφορίας. Η διαδικασία της επικοινωνίας προϋποθέτει την αποστολή και λήψη ενός μηνύματος μεταξύ ενός πομπού και ενός δέκτη, όπου και οι δύο χρησιμοποιούν ένα **κοινό κώδικα**.

Οι πληροφορίες μπορεί να είναι οποιασδήποτε μορφής και **μεταφέρουν κάποιο περιεχόμενο**.

Οι πληροφορίες μπορεί να παραμένουν **στατικές** δηλ. σταθερές στην πάροδο του χρόνου ή να είναι **δυναμικές**, δηλ. να αλλάζουν με την πάροδο του χρόνου.

Η πληροφορία ορίζεται ως το **στοιχείο γνώσης** που μετουσιώνει και δίνει αξία στα πράγματα/γεγονότα, τα οποία αποκτούν σημασία με αυτόν τον τρόπο. Πρέπει πάντως να γίνει ο διαχωρισμός μεταξύ πληροφορίας και γνώσης: η πληροφορία είναι το μέγεθος που, καθώς αυξάνει και επεξεργάζεται, προσδίδει σημασία στα πράγματα και παράγει γνώση.

Πληροφορίες και δεδομένα

Τα δεδομένα δεν είναι πληροφορία, όταν δεν έχουν λάβει από τη νόηση συγκεκριμένη σημασία, ενώ η πληροφορία είναι δεδομένα με σημασία, με «ουσιαστικό» περιεχόμενο.

Τα δεδομένα είναι **ποιοτικά** ή **ποσοτικά** χαρακτηριστικά μιας μεταβλητής ή ενός συνόλου μεταβλητών.

Σε αντίθεση με τα δεδομένα η **πληροφορία** έχει ορισμένο νόημα και είναι οργανωμένη για συγκεκριμένο σκοπό. Με την προσθήκη αξιών στα δεδομένα με διαμόρφωση, οργάνωση, μαθηματική ή στατιστική ανάλυση, διόρθωση λαθών ή συμπίεση παράγεται η πληροφορία. Τη διαδικασία αυτή ονομάζουμε **επεξεργασία δεδομένων**.

Δεδομένα στην Πληροφορική

Στην Πληροφορική συχνά δεν γίνεται διάκριση ανάμεσα στα δεδομένα και την πληροφορία.

Δεδομένα : τρόποι αναπαράστασης εννοιών και γεγονότων που δύνανται να υποστούν **διαχείριση** και **επεξεργασία** από τον άνθρωπο, ή από ένα αυτοματοποιημένο υπολογιστικό σύστημα.

Αυτοί οι τρόποι αναπαράστασης σχετίζονται με τη **φύση της πληροφορίας** που μεταφέρουν.

Μοντελοποίηση δεδομένων : η επιλογή του κατάλληλου είδους αναπαράστασης ώστε να μπορεί να αποθηκευτεί η επιθυμητή πληροφορία χωρίς να υπάρξουν απώλειες και χωρίς να γίνει σπατάλη πόρων

Αδόμητα Δεδομένα

Αδόμητα ή ακατέργαστα δεδομένα (raw data) : Συνήθως πρόκειται για δεδομένα που συλλέγονται από κάποια πηγή χωρίς επεξεργασία και αναφέρονται και ως **πρωτεύοντα δεδομένα (primary data)**.

Τα αδόμητα δεδομένα μπορεί να προέρχονται, π.χ., από χειρονακτική εισαγωγή στοιχείων σε ένα πρόγραμμα.

Έτσι, τα αδόμητα δεδομένα εν γένει περιέχουν σφάλματα και διαφορετική ή καθόλου **μορφοποίηση (format)**: μια εισαγωγή δεδομένων ημερομηνίας μπορεί να περιέχει τις παρακάτω μορφές: "**31st January 1999**", "**31/01/1999**", "**31/1/99**", "**31 Jan**".

Άπαξ και αντληθούν τα αδόμητα δεδομένα, πρέπει να υποστούν **επεξεργασία** για να αποκτήσουν ενιαία μορφή, απαλλαγμένη από σφάλματα, ώστε να παρέχουν πραγματική πληροφορία.

Τύποι Δεδομένων

Στον προγραμματισμό τα δεδομένα χαρακτηρίζονται από τον **τύπο τους (data type)**. Ο τύπος δεδομένων είναι μια ταξινόμηση που καθορίζει :

- τις πιθανές **τιμές** που μπορεί να έχουν τα δεδομένα μας (π.χ. ακέραιος αριθμός ή πραγματικός, συμβολοσειρά κλπ)
- τις **πράξεις** που μπορούν να γίνουν με τις τιμές αυτού του τύπου
- το **νόημα** των δεδομένων, και
- τον **τρόπο αποθήκευσης** των δεδομένων.

Ο τύπος των δεδομένων είναι μια τυπική περιγραφή των τιμών που μπορεί να αποθηκευτούν σε μια σταθερά ή μεταβλητή καθώς και των πράξεων που μπορεί να εκτελεστούν με τις τιμές αυτές.

Συστήματα Τύπων Δεδομένων

Οι τύποι δεδομένων χρησιμοποιούνται εντός **συστημάτων τύπων**, τα οποία προσφέρουν διαφορετικούς τρόπους για τον ορισμό, την ανάπτυξη και τη χρήση τους και διαφοροποιούνται ως προς τους βαθμούς ασφάλειας των τύπων.

Ο κάθε τύπος δεδομένων αναπαριστά και έναν **περιορισμό** για το σύστημα τύπων δεδομένων, περιγράφοντας μια αναπαράσταση που μεταφράζεται σε μια δομή τιμών ή αντικειμένων τα οποία αποθηκεύονται στη μνήμη του υπολογιστή.

Το σύστημα τύπων δεδομένων χρησιμοποιεί την πληροφορία των τύπων για να **ελέγξει την ορθότητα** των προγραμμάτων που προσπελούν ή χειρίζονται τα δεδομένα.

Τύποι Μηχανής

Όλα τα δεδομένα στους υπολογιστές που βασίζονται στην ψηφιακή ηλεκτρονική αναπαρίστανται στο κατώτερο επίπεδο ως **bits**, δηλ. εναλλακτικές καταστάσεις (0 ή 1).

Η μικρότερη μονάδα μνήμης δεδομένων είναι συνήθως μια ομάδα από bits που ονομάζονται **byte** και συνήθως πρόκειται για μια οκτάδα από bits (1 byte = 8 bits).

Η μονάδα που μπορεί να επεξεργαστεί ένας υπολογιστής σε επίπεδο κώδικα μηχανής λέγεται **λέξη (word)** και σήμερα (2013) είναι μήκους 32 ή 64bits.

Λογικός Τύπος (Boolean)

Τα δεδομένα αληθείας αποτυπώνονται δύο καταστάσεις :

- **ΑΛΗΘΕΣ** - **TRUE** και
- **ΨΕΥΔΕΣ** - **FLASE**.

Τα δεδομένα αυτά αποτελούν το **λογικό τύπο ή τύπο Boole (Boolean type)**.

Παρότι μόνο δύο τιμές είναι πιθανές, σπάνια αναπαρίστανται στους υπολογιστές με ένα μόνο δυαδικό ψηφίο για λόγους αποδοτικότητας.

Πολλές γλώσσες προγραμματισμού δεν είναι εφοδιασμένες με ρητό τύπο αληθείας. Αντί αυτού ερμηνεύουν το 0 (για παράδειγμα) ως ψευδές και τις υπόλοιπες τιμές ως αληθείς.

Αριθμητικοί τύποι

ΟΝΟΜΑΣΙΑ ΤΥΠΟΥ	ΘΕΣΕΙΣ ΑΠΟΘΗΚΕΥΣΗΣ	ΕΥΡΟΣ ΤΙΜΩΝ
TINYINT	1 byte	-128 ΕΩΣ 127 [0 ΕΩΣ 255 ΧΩΡΙΣ ΠΡΟΣΗΜΟ - UNSIGNED]
SMALLINT	2 bytes	-32,768 ΕΩΣ 32,767 [0 ΕΩΣ 65,535]
MEDIUMINT	3 bytes	-8,388,608 ΕΩΣ 8,388,607 [0 ΕΩΣ 16,777,215]
INT	4 bytes	-/+2.147E+9 [0 ΕΩΣ 4.294E+9]
BIGINT	8 bytes	-/+9.223E+18 [0 ΕΩΣ 18.45E+18]
FLOAT	4 bytes	Min=+/-1.15E-38 Max=+/-3.403E+38
DOUBLE	8 bytes	Min=+/-2.225E-308 Max=+/-1.798E+308
DECIMAL (M, [D]) <i>αποθηκεύονται ως συμβολοσειρά</i>	M+2	Max εύρος = το εύρος DOUBLE

Τύποι Κειμένου/Συμβολοσειρών

CHAR (size)	Αποθηκεύει ένα σταθερό μήκος συμβολοσειράς Το μέγεθος (μήκος) ορίζεται σε παρένθεση. Μπορεί να αποθηκεύσει μέχρι 255 χαρακτήρες.
VARCAR (size)	Αποθηκεύει ένα μεταβλητό μήκος συμβολοσειράς Το μέγιστο μέγεθος (μήκος) ορίζεται σε παρένθεση. Μπορεί να αποθηκεύσει μέχρι 255 χαρακτήρες. Σημείωση: Αν το μήκος της συμβολοσειράς είναι μεγαλύτερο από 255 χαρακτήρες, μετατρέπεται σε τύπο TEXT
TINYTEXT	Αποθηκεύει το πολύ 255 χαρακτήρες
TEXT	Αποθηκεύει το πολύ 65,535 χαρακτήρες
BLOB	Binary Large Objects. Αποθηκεύει το πολύ 65,535 bytes δεδομένων
MEDIUMTEXT	Αποθηκεύει το πολύ 16,777,215 χαρακτήρες
MEDIUMBLOB	Binary Large Objects. Αποθηκεύει το πολύ 16,777,215 bytes δεδομένων
LONGTEXT	Αποθηκεύει το πολύ 4,294,967,295 χαρακτήρες
LOBLOB	Binary Large Objects. Αποθηκεύει το πολύ 4,294,967,295 bytes δεδομένων
ENUM (x, y, z, etc.)	Επιτρέπει την εισαγωγή λίστας πιθανών τιμών (μέχρι 65535 τιμές). Αν η τιμή που εισάγεται δεν είναι μεταξύ των πιθανών τιμών, αποθηκεύεται η κενή τιμή.
SET	Παρόμοιος τύπος με το ENUM εκτός του ότι το SET μπορεί να περιέχει μέχρι 64 τιμές και μπορεί να αποθηκεύσει περισσότερες από μία

Τύποι Χρόνου

DATE ()	Ημερομηνία με μορφή οργάνωσης: ΕΕΕΕ-ΜΜ-ΗΗ. Υποστηρίζονται ημερομηνίες από '1000-01-01' έως '9999-12-31'
DATETIME ()	Ημερομηνία και ώρα με μορφή οργάνωσης: ΕΕΕΕ-ΜΜ-ΗΗ ΩΩ:ΛΛ:ΔΔ. Υποστηρίζονται ημερομηνίες από '1000-01-01 00:00:00' έως '9999-12-31 23:59:59'
TIMESTAMP ()	Τρόπος αποτύπωσης της τρέχουσας ώρας που προέρχεται από το λειτουργικό σύστημα UNIX με μορφή οργάνωσης: ΕΕΕΕ-ΜΜ-ΗΗ ΩΩ:ΛΛ:ΔΔ. Υποστηρίζονται χρόνοι από '1970-01-01 00:00:01' έως '2038-01-09 03:14:07'
TIME ()	Χρόνος με μορφή οργάνωσης: ΩΩ:ΛΛ:ΔΔ. Υποστηρίζονται χρόνοι από '-838:59:59' έως '838:59:59'
YEAR ()	Έτος με 2 ή 4 ψηφία. Για 4 ψηφία υποστηρίζονται τα έτη : 1901 έως 2155 ενώ για τα 2 ψηφία τα έτη από 70 έως 69, αναπαριστώντας τα έτη από 1970 έως 2069.

Τύπος Απαρίθμησης

Ο τύπος της **απαρίθμησης (enumeration)** αφορά σε μεταβλητές που παίρνουν δεδομένες τιμές που διαφέρουν μεταξύ τους, οι οποίες μπορούν να συγκριθούν, αλλά δεν έχουν απαραίτητα κάποια συγκεκριμένη αναπαράσταση στην μνήμη του υπολογιστή.

Οι μεταγλωττιστές και οι διερμηνευτές τις αναπαριστούν αυθαίρετα. Στη MySQL αποτελούν μέρος του τύπου κειμένου.

Για παράδειγμα, τα τέσσερα χαρτιά σε μία τράπουλα θα μπορούσαν να είναι τέσσερις απαριθμητές με ονόματα ΣΠΑΘΙ, ΚΑΡΩ, ΜΠΑΣΤΟΥΝΙ, ΚΟΥΠΑ, που θα ανήκουν σε μία απαρίθμηση με όνομα τράπουλα. Αν η μεταβλητή V δηλωθεί να έχει τράπουλα για τύπο, μπορούμε να αναθέσουμε οποιαδήποτε από αυτές τις τιμές σε αυτή την μεταβλητή.

Σύνθετοι τύποι

Η ομαδοποίηση απλών τύπων σε μια πιο σύνθετη δομή αποθήκευσης δημιουργεί ένα **σύνθετο τύπο δεδομένων**.

ΟΝΟΜΑ
ΕΠΩΝΥΜΟ
ΗΛΙΚΙΑ
ΚΩΔΙΚΟΣ

EMPLOYEE
ΟΝΟΜΑ
ΕΠΩΝΥΜΟ
ΗΛΙΚΙΑ
ΚΩΔΙΚΟΣ

Στην τεχνολογία των βάσεων δεδομένων, οι δομές αυτές ονομάζονται **εγγραφές (records)**.

Δομές Δεδομένων

Δομή δεδομένων (data structure): διαφορετικοί τρόποι οργάνωσης και αποθήκευσης δεδομένων, ώστε αυτά να μπορούν να χρησιμοποιηθούν αποδοτικά στον προγραμματισμό.

Σε συγκεκριμένες εφαρμογές χρησιμοποιούνται συγκεκριμένες δομές δεδομένων, δηλαδή αυτές που είναι οι πιο αποδοτικές αλγοριθμικά για το κάθε είδος εφαρμογής.

Χρησιμοποιούνται σχεδόν σε κάθε πρόγραμμα ή σύστημα λογισμικού. Παρέχουν έναν τρόπο αποδοτικής διαχείρισης τεράστιου όγκου δεδομένων, όπως μεγάλες βάσεις δεδομένων και υπηρεσίες ευρετηρίου στο διαδίκτυο.

Οι δομές δεδομένων διακρίνονται επίσης σε **στατικές** και σε **δυναμικές**:

- Στις στατικές δομές το πλήθος των στοιχείων είναι σταθερό και καθορίζεται στον ορισμό του τύπου στο τμήμα δηλώσεων πχ ο πίνακας.
- Στις δυναμικές δομές το πλήθος των στοιχείων της δομής καθορίζεται κατά τη διάρκεια εκτέλεσης του προγράμματος.

Οι Βάσεις Δεδομένων

- Βάση δεδομένων : ένα σύνολο δεδομένων που χαρακτηρίζονται από κάποια λογική οργάνωση και ομαδοποίηση έτσι ώστε να είναι εύκολη και αποτελεσματική η διαχείρισή τους.
- Οι βάσεις δεδομένων κατέχουν κεντρική θέση στην επιστήμη της Πληροφορικής, διότι επιτυγχάνουν τη διαχείριση της πληροφορίας.
- Τα τελευταία χρόνια παρατηρείται μια ραγδαία ανάπτυξη της τεχνολογίας των βάσεων δεδομένων τόσο σε θεωρητικό επίπεδο όσο και σε επίπεδο εφαρμογών:
 - Επιχειρήσεις - Τραπεζικός τομέας
 - Δημόσια Διοίκηση
 - Εκπαίδευση
 - Δίκτυα Μεταφορών
 - Δίκτυα Επικοινωνιών κλπ

Συστήματα διαχείρισης ΒΔ

- Η διαχείριση μιας βάσης δεδομένων πραγματοποιείται με τη βοήθεια ειδικών προγραμμάτων, που ονομάζονται **συστήματα διαχείρισης βάσεων δεδομένων (ΣΔΒΔ) - Database Management Systems (DBMS)**.
- Ένα ΣΔΒΔ είναι **λογισμικό** που χρησιμοποιείται για:
 - τη δημιουργία,
 - τη συντήρηση,
 - την επεξεργασία
 - τη συνολική διαχείριση μιας ΒΔ.

Παραδείγματα ΣΔΒΔ

- Microsoft Access,
- MS SQL Server
- Oracle,
- Sybase
- MySQL
- Postgress,

Microsoft Access : σε προσωπικούς υπολογιστές και σε επιχειρήσεις μικρού μεσαίου /μεγέθους.

MS SQL Server, Oracle, Sybase : σε μεγάλες επιχειρήσεις, υποστηρίζουν πολλούς χρήστες, πολλές ταυτόχρονες προσπελάσεις πάνω στα ίδια δεδομένα, καθώς και απομακρυσμένη πρόσβαση μέσω δικτύου.

MySQL και Postgress : ανοικτού λογισμικού, κυρίως σε εφαρμογές διαδικτύου



PostgreSQL



ORACLE

Εισαγωγή στις Βάσεις Δεδομένων II

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΕΒΔΟ100)

Αρχές καλής σχεδίασης

- **Ακεραιότητα**
- **Αξιοπιστία**
- **Αποτελεσματικότητα**
- **Ταχύτητα**
- **Ασφάλεια**

Χαρακτηριστικά σχεδίασης 1

- Δεν θα πρέπει να περιλαμβάνει **περιττά δεδομένα (redundancy)**. Αυτό σημαίνει ότι τα ίδια δεδομένα δεν θα πρέπει να καταχωρούνται στη βάση δύο ή περισσότερες φορές.
- **Διαφορετικά :**
 - **σπαταλούμε άσκοπα αποθηκευτικό χώρο**
 - **υπάρχει ο κίνδυνος δημιουργίας ασυνεπών δεδομένων (inconsistency)**
- Για το λόγο αυτό, ένας από τους πρώτους ελέγχους που πραγματοποιούμε στη βάση αμέσως μετά το σχεδιασμό της, είναι ο **έλεγχος παρουσίας επαναλαμβανόμενων πεδίων, και η απομάκρυνσή τους, εφ' όσον υπάρχουν.**

Χαρακτηριστικά σχεδίασης 2

- Η βάση θα πρέπει να είναι σχεδιασμένη έτσι ώστε τα δεδομένα που περιλαμβάνει να ανακτώνται **εύκολα** και **γρήγορα**.
- **Διαφορετικά** : η βάση είναι εξαιρετικά δυσκίνητη και αναποτελεσματική.
- Η σωστή σχεδίαση επιτυγχάνεται με εφαρμογή επί της δομής της βάσης, μιας τεχνικής, η οποία ονομάζεται **κανονικοποίηση (normalization)**.

Χαρακτηριστικά σχεδίασης 3

- Το ΣΔΒΔ θα πρέπει να χαρακτηρίζεται από **σύστημα ασφάλειας (security system)** που να απαγορεύει την πρόσβαση στα δεδομένα μη εξουσιοδοτημένων ατόμων.
- Η τεχνική που συνήθως εφαρμόζεται, συνίσταται στον καθορισμό **ομάδων χρηστών (user groups)** με διαφορετικά δικαιώματα πρόσβασης στον καθένα από αυτούς.
- Ο κάθε χρήστης λαμβάνει ένα **κωδικό πρόσβασης (password)** και τα καθήκοντα που μπορεί να επιτελέσει είναι εντελώς συγκεκριμένα και καθορισμένα εκ των προτέρων.

Χαρακτηριστικά σχεδίασης 4

- Το ΣΔΒΔ θα πρέπει να μπορεί να διαχειρίζεται **ταυτόχρονες προσπελάσεις** πάνω στα ίδια δεδομένα (**concurrency control**), ώστε να αποτρέψει την καταχώρηση πληροφοριών στην ίδια θέση αποθήκευσης.
- Αυτό επιτυγχάνεται με το «κλείδωμα» των δεδομένων τη στιγμή που δοθεί μια εντολή εγγραφής/αλλαγής σε αυτά.

Χαρακτηριστικά σχεδίασης 5

- Το ΣΔΒΔ θα πρέπει να διαθέτει σύστημα δημιουργίας **αντιγράφων ασφαλείας (backups)** των δεδομένων που είναι καταχωρημένα σε αυτή.
- Η ταυτόχρονη αποθήκευση των δεδομένων σε περισσότερα από ένα σημεία, είναι μια εργασία επιβεβλημένη, προκειμένου να είναι δυνατή η ανάκτησή τους σε περιπτώσεις κατάρρευσης της βάσης για οποιοδήποτε λόγο.

Αρχιτεκτονικές Σ.Δ.Β.Δ.

Οι κοινοί τρόποι σχεδιασμού του λογισμικού σε ό,τι αφορά τις επιμέρους μονάδες (*modules*) και τη μεταξύ τους επικοινωνία χαρακτηρίζονται ως **αρχιτεκτονικά μοντέλα** ή **αρχιτεκτονικές**.

Οι βασικές αρχιτεκτονικές των σύγχρονων ΣΔΒΔ κατηγοριοποιούνται ανάλογα με τις **βαθμίδες (tiers)** που εμπλέκονται.

Ο όρος των βαθμίδων ανάγεται στη λογική **αρχιτεκτονική P-A-D** (*Presentation-Application-Data*)

Βασίζεται στην αλληλεπίδραση των τριών λογικών διαδικασιών:

- **Παρουσίαση – Presentation** : η διεπαφή χρήστη (user interface)
- **Εφαρμογή – Application** : η επεξεργασία ή λογική εφαρμογής
- **Δεδομένα – Data** : η διαχείριση

Εξυπηρετητές και Πελάτες

Η ανάπτυξη μεγάλων εφαρμογών, επικεντρωμένων στην ανταλλαγή δεδομένων μέσω δικτύου, οδήγησε στο **περιβάλλον Πελάτη/Εξυπηρετητή (Client/Server)** στη δεκαετία του 1980.

Εξυπηρετητής (Server) : μια κεντρική εφαρμογή με την οποία συνδέονται και ανταλλάσσουν πληροφορίες άλλες εφαρμογές που βρίσκονται στους υπολογιστές των χρηστών και χαρακτηρίζονται ως **Πελάτες (Clients)** .

Η ανταλλαγή της πληροφορίας γίνεται μέσω *συναλλαγών (transactions)*, δηλαδή μέσω καλά ορισμένων *αιτημάτων (requests)* και *αποκρίσεων (responses)*.



Αδύνατοι και Παχείς

Το περιβάλλον Πελάτη/Εξυπηρετητή αφορά στη φυσική αρχιτεκτονική και εμφανίζει επιμέρους αρχιτεκτονικές.

Όταν το μεγαλύτερο μέρος της επεξεργασίας (λογική της εφαρμογής) γίνεται στον Εξυπηρετητή και ο Πελάτης αναλαμβάνει κυρίως την παρουσίαση αποτελεσμάτων της διαδικασίας, τότε έχουμε την περίπτωση **Αδύνατου Πελάτη (Thin Client) – Παχύ Εξυπηρετητή (Fat Server)**.

Όταν ο Πελάτης αναλαμβάνει το μεγαλύτερο μέρος της επεξεργασίας μιλάμε για **Παχύ Πελάτη (Fat Client) – Αδύνατο Εξυπηρετητή (Thin Server)**.

Στην ειδική περίπτωση του Διαδικτύου, όπου αποκλειστικά και μόνο το επίπεδο της παρουσίασης είναι στο φυλλομετρητή (browser), ενώ το σύνολο της επεξεργασίας είναι στον Εξυπηρετητή, μιλάμε για **Υπέρ-Παχύ Εξυπηρετητή (Super-Fat Server)**.

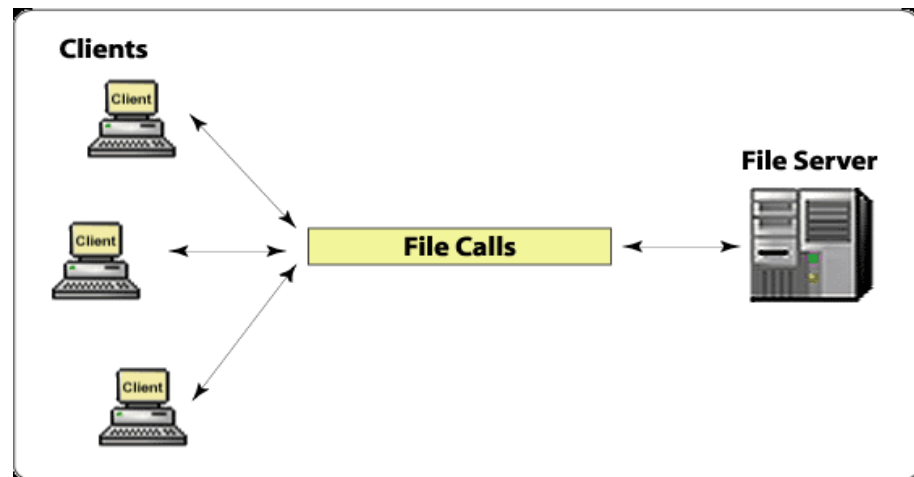


Αρχιτεκτονική 1-tier

Ο υπολογιστής που φιλοξενεί την εφαρμογή είναι ο ίδιος με το οποίο ο χρήστης επικοινωνεί με την εφαρμογή.

Τυπική περίπτωση είναι η ανάπτυξη μιας βάσης στο ΣΔΒΔ MS Access.

Μια επέκταση είναι όταν ένας file server φιλοξενεί τα αρχεία αποθήκευσης του ΣΔΒΔ (αλλά όχι κατ' ανάγκη το ίδιο) και βρίσκεται σε τοπικό δίκτυο με άλλους υπολογιστές.



Τα + και τα – του 1-tier

Πλεονεκτήματα:

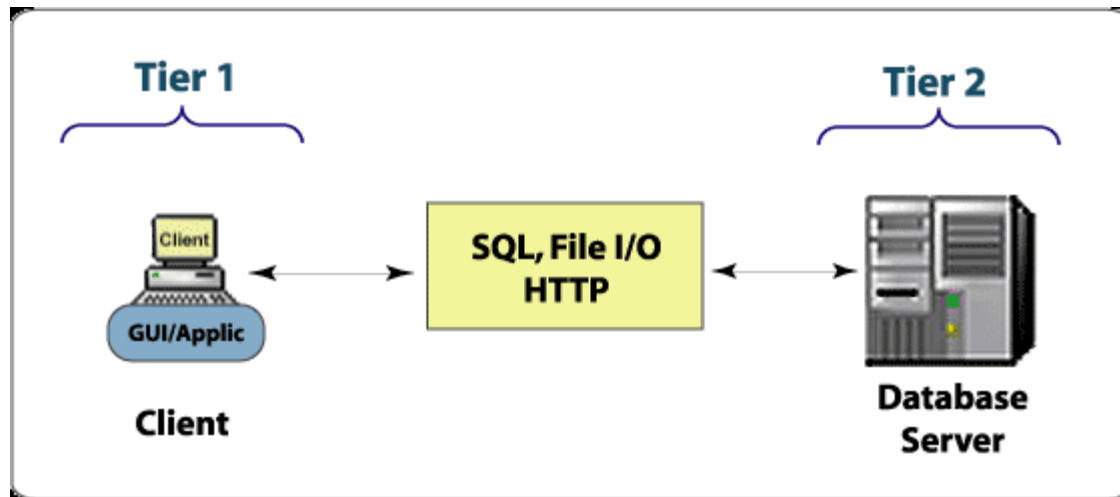
- Ο διαμοιρασμός δεδομένων σε πολλαπλούς χρήστες γίνεται με απλό τρόπο, αν και οι δυνατότητες είναι περιορισμένες.
- Το κόστος αποθήκευσης είναι χαμηλό και μοιράζεται στους διασυνδεδεμένους χρήστες.
- Επειδή η ίδια εφαρμογή ΣΔΒΔ εγκαθίσταται πολλές φορές, ενδεχομένως επιτυγχάνεται χαμηλότερη τιμή αγοράς ανά εγκατάσταση (site license).

Μειονεκτήματα:

- Μικρή δυνατότητα κλιμάκωσης.
- Αδυναμία ταυτόχρονης χρήσης.

Αρχιτεκτονική 2-tier

Ένα κεντρικό σύστημα, πολλές φορές με αποκλειστική αρμοδιότητα (server), παρέχει τη λειτουργικότητα για πολλά συστήματα χρηστών (clients). Το ΣΔΒΔ είναι στον εξυπηρετητή, με τον οποίο επικοινωνούν οι εφαρμογές – πελάτες.



Τα + και τα – του 2-tier

Πλεονεκτήματα:

- Κλιμάκωση, οριζόντια (αύξηση πελατών) και κάθετη (αύξηση εξυπηρετητών).
- Αποτελεσματική χρήση των υπολογιστικών πόρων με καλό επίπεδο ασφάλειας.
- Εύκολη ενσωμάτωση νέας τεχνολογίας (π.χ. κινητές συσκευές).
- Μαζί με τον client μπορεί να υπάρχουν και άλλες εφαρμογές.

Μειονεκτήματα:

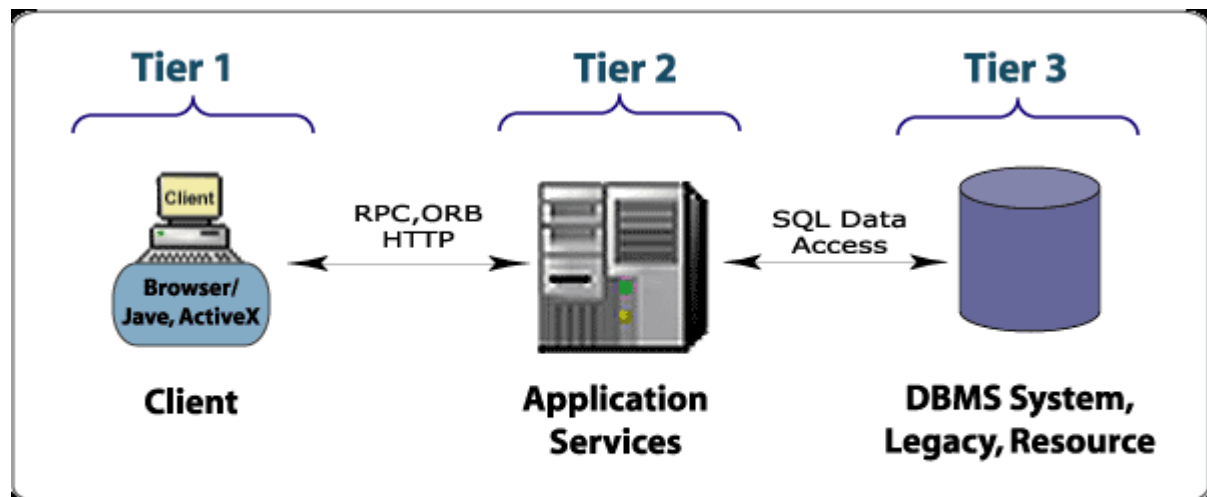
- Η client εφαρμογή εγκαθίσταται σε πολλά συστήματα, πράγμα που προκαλεί δυσκολία αλλαγής της (π.χ. αναβάθμισης) και, για το λόγο αυτό, το κόστος συντήρησης αυξάνεται σημαντικά.

Αρχιτεκτονική N-tier

Η λειτουργικότητα της εφαρμογής του ΣΔΒΔ χωρίζεται σε διαφορετικά επίπεδα που διασυνδέονται λογικά και μοιράζονται πληροφορίες.

Κάθε ενδιάμεσο επίπεδο αποτελεί ένα μοντέλο client/server.

Η πιο συνήθης υλοποίηση είναι αυτή των τριών βαθμίδων (3-tier), όπου υπάρχει σαφής φυσικός διαχωρισμός της επεξεργασίας από τη βάση δεδομένων και την παρουσίαση.



Τα + και τα – του N-tier

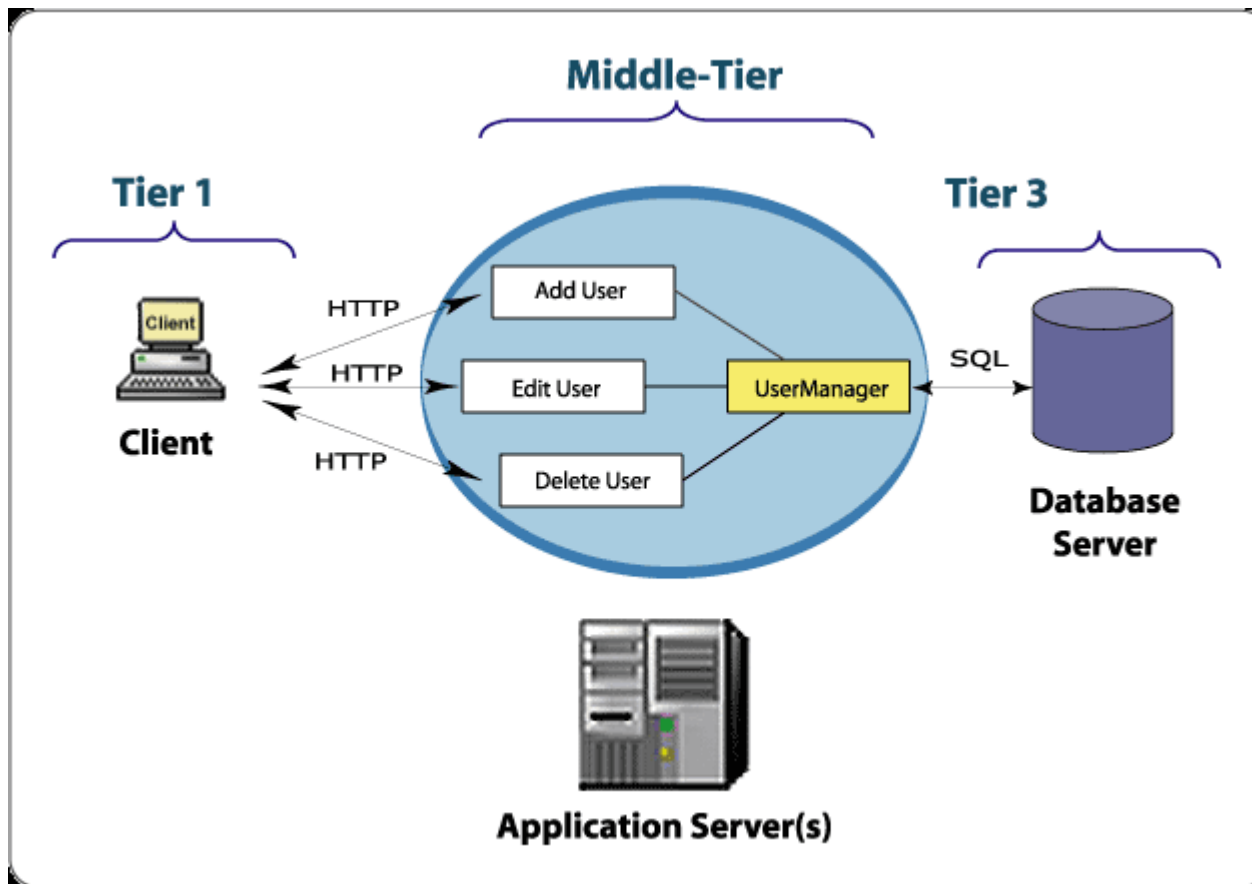
Πλεονεκτήματα:

- Αυξημένη ασφάλεια, αφού ο χρήστης δεν προσπελαίνει άμεσα τη βάση δεδομένων
- Αυξημένη ευελιξία και έλεγχος λόγω των ενδιάμεσων επιπέδων και, επομένως, καλύτερη απόδοση
- Ελαχιστοποιημένη διαχείριση σε σχέση με το μοντέλο client/server

Μειονεκτήματα:

- Πολυπλοκότητα της σχεδίασης
- Αυξημένο αρχικό κόστος υλοποίησης

Αρχιτεκτονική διαδικτύου



Τα + και τα – του διαδικτύου

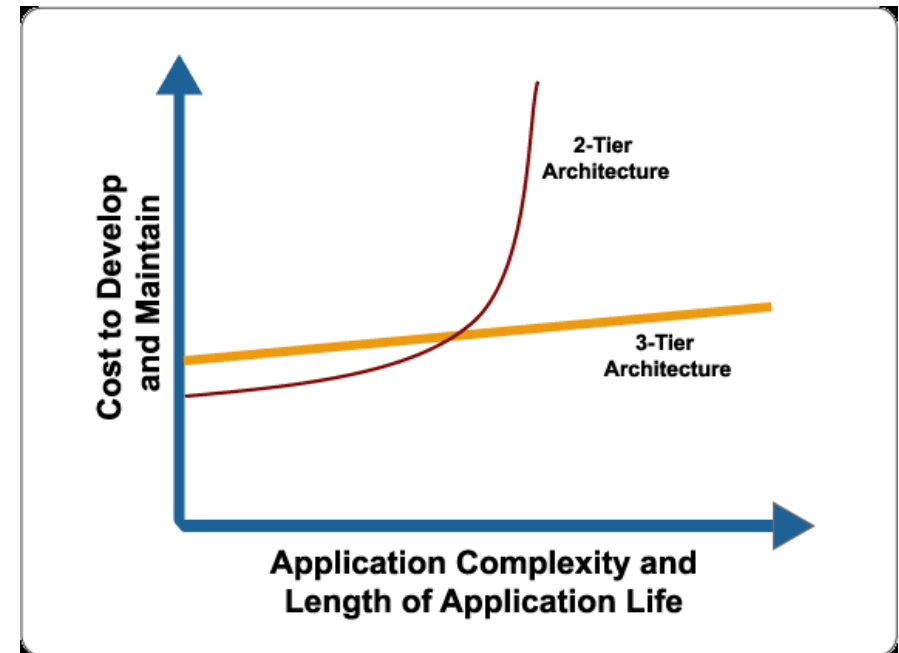
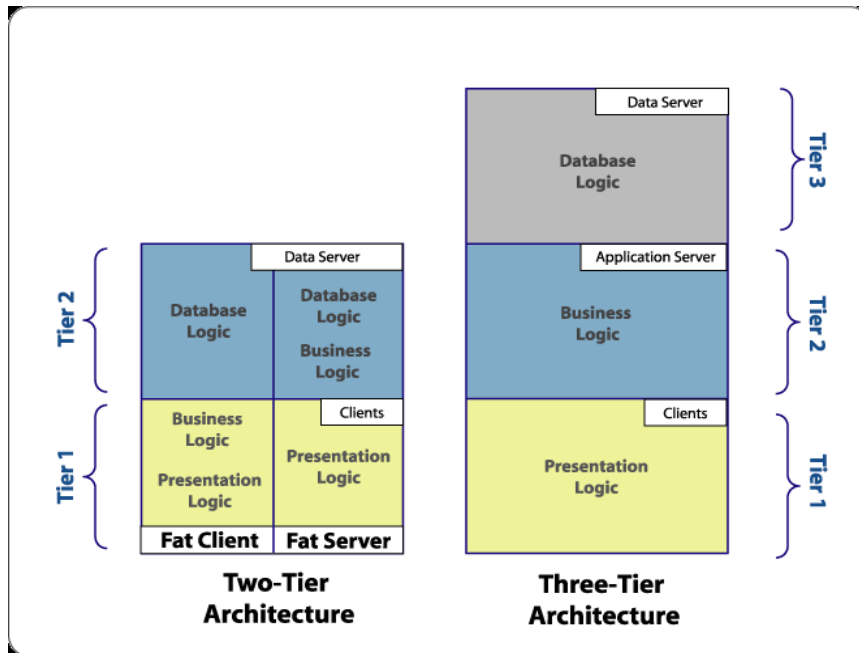
Πλεονεκτήματα

- Αυξημένη ασφάλεια, αφού ο χρήστης δεν προσπελαύνει άμεσα τη βάση δεδομένων
- Αυξημένη ευελιξία και έλεγχος και, επομένως, καλύτερη απόδοση
- Ελαχιστοποιημένη διαχείριση σε σχέση με το μοντέλο client/server
- Χαμηλές απαιτήσεις για την πλευρά του χρήστη

Μειονεκτήματα

- Αυξημένοι κίνδυνοι που αντιμετωπίζονται μέσω ισχυρών πολιτικών ασφάλειας (firewalls, λογισμικό ανίχνευσης εισβολέων, κ.ά.).

Σύγκριση 2-tier και 3-tier



Στάδια Ζωής Βάσεων Δεδομένων

- **Στάδιο 1:** Σχεδίαση και υλοποίηση της βάσης
- **Στάδιο 2:** Καταχώρηση των δεδομένων στη βάση (data entry)
- **Στάδιο 3:** Διαχείριση της βάσης δεδομένων και επεξεργασία των δεδομένων

Επίπεδα Σχεδίασης Β.Δ.

- **Φυσικό επίπεδο (physical level):** σχεδιάζεται και αναλύεται ο τρόπος με τον οποίο τα δεδομένα της βάσης θα αποθηκεύονται στο ΣΔΒΔ. Στο επίπεδο αυτό, καθορίζονται, μεταξύ άλλων, οι τύποι δεδομένων που θα χρησιμοποιηθούν.
- **Εννοιολογικό ή λογικό επίπεδο (conceptual / logical level):** καθορίζεται ο τρόπος που τα δεδομένα και οι σχέσεις που υφίστανται ανάμεσά τους θα μοντελοποιηθούν στη βάση του συστήματος.
- **Επίπεδο όψης (view level):** καθορίζεται ο τρόπος που ένα μέρος των δεδομένων θα προβάλλεται στο χρήστη σύμφωνα με καθορισμένα κριτήρια που εκείνος θέτει με άμεσο ή έμμεσο τρόπο.

Εμείς επικεντρωνόμαστε στο *λογικό επίπεδο* γιατί εκεί γίνεται κυρίως η μοντελοποίηση της βάσης δεδομένων.

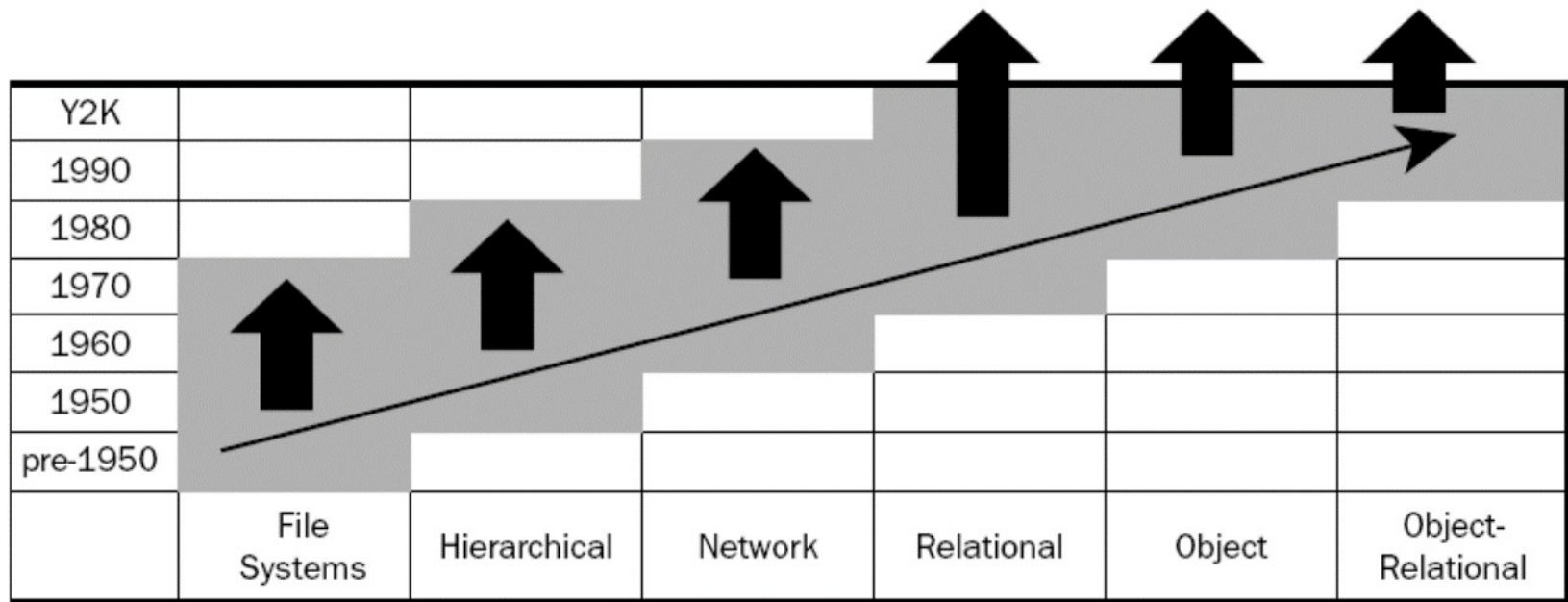
Μοντέλα λογικού επιπέδου

Τα μοντέλα σχεδίασης εννοιολογικού-λογικού επιπέδου δεν αφορούν απλά στους τρόπους οργάνωσης δεδομένων, αλλά προχωρούν στο να ορίσουν και τις πράξεις που τελούνται στα δεδομένα.

Η ιστορική εξέλιξη των μοντέλων είναι:

- το ιεραρχικό μοντέλο (hierarchical model),
- το δικτυωτό μοντέλο (network model),
- το σχεσιακό μοντέλο (relational model) και
- το αντικειμενοστραφές μοντέλο (object-oriented model).

Μοντέλα λογικού επιπέδου



Μοντέλο Επίπεδου Αρχείου

Αρχείο κειμένου (text file) που περιέχει συλλογή ακολουθιών δεδομένων, διαχωρισμένα με ειδικούς χαρακτήρες, που μπορούν να διερευνηθούν **σειριακά** για τον εντοπισμό τους.

Η σειριακή αναζήτηση δεδομένων είναι αναποτελεσματική, αλλά μπορεί να είναι χρήσιμη για μικρούς καταλόγους και για απλές δομές δεδομένων.

Η διαχείριση γίνεται μέσω προγραμμάτων που διαβάζουν και γράφουν στα αρχεία.

Προβλήματα :

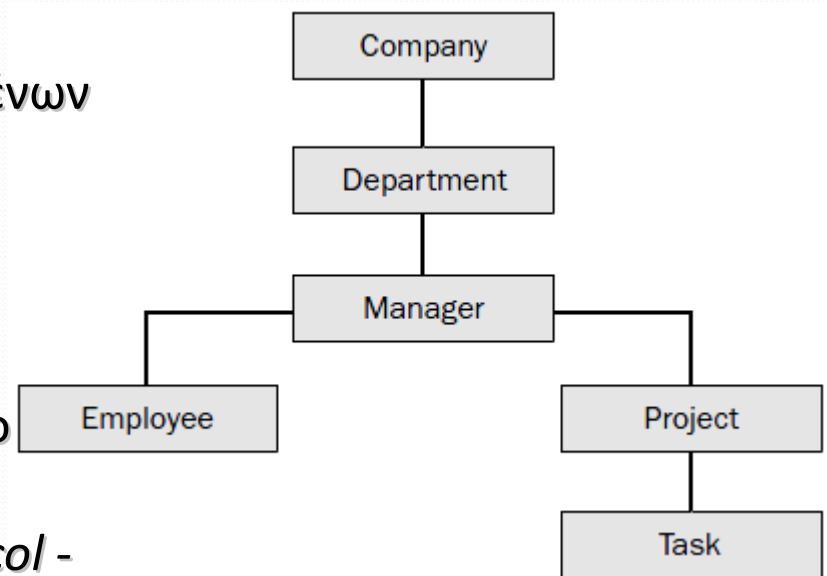
- ταυτόχρονη εγγραφή από περισσότερες από μία εφαρμογές,
- απώλεια δεδομένων λόγω κατάρρευσης μιας εφαρμογής κατά τη διάρκεια προσπέλασης του αρχείου.

Ιεραρχικό μοντέλο (1)

- Τα δεδομένα οργανώνονται σε δομές που ακολουθούν τη λογική του **ανεστραμμένου δένδρου**.
- Τις δομές αυτές μπορούμε να τις θεωρήσουμε **πίνακες καταχωρήσεων - record tables**, οι οποίοι χαρακτηρίζονται από **δείκτες - pointers**.
- Οι μεταξύ των πινάκων σχέσεις είναι του τύπου **πρόγονος/απόγονος - parent/child** δηλ. δημιουργούνται σχέσεις «ένα προς πολλά». Αυτό καθίσταται δυνατό μέσω των δεικτών που φέρουν οι πίνακες.
- Ο πρώτος πίνακας-πρόγονος καλείται **ρίζα-root**, και έχει έναν ή περισσότερους πίνακες – απογόνους.
- Κάθε πίνακας, μπορεί να έχει έναν ή περισσότερους απογόνους, αλλά **μόνο έναν πρόγονο**.

Ιεραρχικό μοντέλο (2)

- Η αναζήτηση είναι **μονοσήμαντη** μέσω των «κλαδιών», ξεκινώντας από τη ρίζα.
- Η ακεραιότητα αναφοράς των δεδομένων είναι **εγγενής** και δεν χρειάζεται ιδιαίτερος μηχανισμός για να επιτευχθεί.
- Ένα τυπικό παράδειγμα οργάνωσης σύμφωνα με την ιεραρχικό μοντέλο είναι το πρότυπο καταλόγου *Lightweight Directory Access Protocol - LDAP*.



Ιεραρχικό μοντέλο (3)

Τα σοβαρότερα μειονεκτήματα είναι :

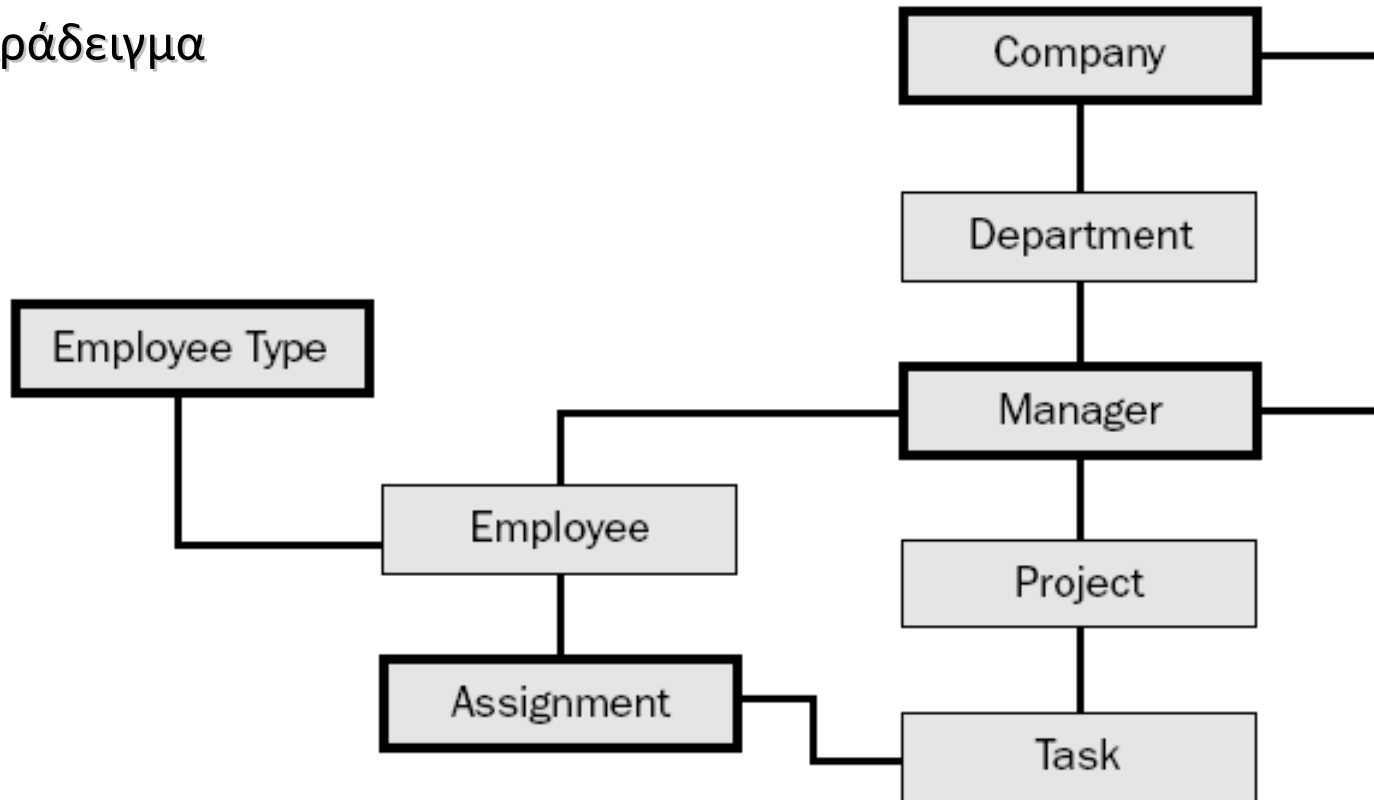
- η ανάγκη εκκίνησης της αναζήτησης από τη ρίζα
- η μεγάλη δυσκολία αλλαγής της δομής του «δέντρου», η οποία επηρεάζει, όχι μόνο την ίδια τη βάση δεδομένων, αλλά και τις εφαρμογές που τη χειρίζονται.
- λόγω της στενής σχέσης των πινάκων, όταν διαγράφεται ένας πρόγονος, πρέπει να διαγραφούν και οι απόγονοί του.

Δικτυακό μοντέλο (1)

- Αποτελεί βελτίωση του ιεραρχικού μοντέλου
- Επιτρέπει σε κάθε απόγονο να έχει πολλούς προγόνους, δηλ. επιτρέπει τη δημιουργία σχέσεων «πολλά προς πολλά».
- Οι πίνακες συνδέονται μεταξύ τους δημιουργώντας ένα **δίκτυο συνδέσεων**, και ουσιαστικά συνδυάζονται μεμονωμένες ιεραρχικές βάσεις δεδομένων σε μια μεγαλύτερη.
- Η πλοήγηση δεν απαιτεί την έναρξη από κάποια ρίζα, αλλά **μπορεί να αρχίσει από οποιαδήποτε πίνακα.**
-

Δικτυακό μοντέλο (2)

Παράδειγμα



Σχεσιακό Μοντέλο

- Βασίζεται στις σχέσεις που αναπτύσσονται μεταξύ των πινάκων, ενώ αλλάζει η έννοια της ιεραρχίας και δεν αποτελεί το κυρίαρχο χαρακτηριστικό.
- Τα δεδομένα οργανώνονται σε πίνακες που καλούνται **οντότητες (entities)**.
- Οι **συσχετίσεις (relations)** μεταξύ των οντοτήτων αποδίδονται επίσης με πίνακες.
- Οι οντότητες και οι συσχετίσεις καλούνται συνολικά **ΣΧΕΣΕΙΣ**
- Η αναζήτηση στις σχεσιακές βάσεις δεδομένων γίνεται **απ' ευθείας στον πίνακα** που έχει τα δεδομένα που μας ενδιαφέρουν, χωρίς να πρέπει να ακολουθηθεί μια δενδρική διαδρομή.

Αντικειμενοστραφές μοντέλο

- Ενώ οι σχεσιακές ΒΔ προορίζονται για την ανεύρεση δεδομένων σε δύο διαστάσεις (πίνακες σειρών), οι **αντικειμενοστραφείς** είναι αποδοτικές στην εξεύρεση ενός μοναδικού στοιχείου. Αυτό τις καθιστά **αποδοτικότερες σε νέους τύπους δεδομένων** όπως ήχου και εικόνας (σταθερής ή video).
- Οι αντικειμενοστραφείς ΒΔ **προσπελαύνονται άμεσα από γλώσσες** όπως η Java και η C++, πράγμα που τις κάνει **10 έως 1000 φορές ταχύτερες** από τις αντίστοιχες σχεσιακές.
- Αν και σχεδιαστικά οι αντικειμενοστραφείς ΒΔ ξεπερνούν πολλές αδυναμίες των σχεσιακών, πρακτικά παρουσιάζουν το μειονέκτημα ότι η διείσδυσή τους στην αγορά είναι μικρή σε σύγκριση με τις σχεσιακές βάσεις. Αυτό σημαίνει **λιγότερους εξειδικευμένους προγραμματιστές και μη διασφάλιση εκτέλεσης των εφαρμογών σε διαφορετικές πλατφόρμες.**

Μοντέλο Συσχέτισης Αντικειμένων

- **Υβριδικό μοντέλο** Συσχέτισης Αντικειμένων – *Object Related* της Oracle8, που εισήχθησαν στο τέλος της δεκαετίας του 1990.
- Ενσωματώνει αντικειμενοστραφή χαρακτηριστικά σε μια σχεσιακή βάση δεδομένων.
- Επιτρέπει τη δημιουργία νέων τύπων δεδομένων, αν και η βασική αρχιτεκτονική παραμένει σχεσιακή.
- Μια τέτοια βάση δεδομένων χρησιμοποιεί το FBI για καταχώρηση και ανάλυση δακτυλικών αποτυπωμάτων

Είδη λειτουργιών και γλωσσών

- Λειτουργίες που τροποποιούν τη **δομή** : γλώσσα ορισμού δεδομένων (**Data Definition Language, DDL**),
- Λειτουργίες που τροποποιούν το **περιεχόμενο** : γλώσσα διαχείρισης δεδομένων (**Data Manipulation Language, DML**)

Η γλώσσα SQL

- **Δομημένη Γλώσσα Ερωτοαποκρίσεων - Structured Query Language, SQL**
- Απαντάται σε όλα ανεξαιρέτως τα συστήματα διαχείρισης σχεσιακών βάσεων δεδομένων
- Εμφανίζεται σε πολλές παραλλαγές αλλά η βασική ιδέα που περιγράφει τη λειτουργία της, είναι ουσιαστικά η ίδια.
- Οι βασικές λειτουργίες της διαχωρίζονται σε 4 κατηγορίες:
 - η εισαγωγή δεδομένων (data insertion),
 - η διαγραφή δεδομένων (data deletion),
 - η ανάκτηση δεδομένων (data retrieval) και
 - η τροποποίηση δεδομένων (data update).

<http://w3schools.com/sql/default.asp>



Μοντέλο Συσχετίσεων Οντοτήτων (ERM)

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Στόχοι και Αποτελέσματα

Στόχος: να περιγραφούν οι βασικές έννοιες που χρησιμοποιούνται για τη διαγραμματική περιγραφή μιας βάσης δεδομένων.

Προσδοκώμενα αποτελέσματα:

Όταν ολοκληρώσετε τη μελέτη αυτού του κεφαλαίου, θα έχετε κατανοήσει:

- Ότι για την περιγραφή μιας βάσης δεδομένων χρησιμοποιούνται διαγράμματα που είναι γνωστά ως λογικά μοντέλα δεδομένων.
- Τις έννοιες που χρησιμοποιούνται για την ανάπτυξη ενός λογικού μοντέλου δεδομένων.
- Να αναπτύσσετε ένα λογικό μοντέλο δεδομένων χρησιμοποιώντας τα κατάλληλα σύμβολα.

Βιβλιογραφική παραπομπή :

Κεχρής, Ε., “Σχεσιακές βάσεις δεδομένων – Θεωρία & εργαστηριακές ασκήσεις”, **Κεφάλαιο 2**

Μάργαρης, Α.Ι., “Εισαγωγή στις Βάσεις Δεδομένων” - **Κεφάλαιο 2**

Πηγουνάκης, Κ., Σημειώσεις μαθήματος, **Κεφάλαιο 3**

Ανάλυση απαιτήσεων

- Καθορισμός των απαιτήσεων δεδομένων για τη ΒΔ σε επίπεδο απλών αντικειμένων
- Κατηγοριοποίηση και περιγραφή της πληροφορίας για τα απλά αντικείμενα
- Ταυτοποίηση και κατηγοριοποίηση των σχέσεων μεταξύ των αντικειμένων
- Καθορισμός των τύπων συναλλαγών (transactions) που θα εκτελούνται στη ΒΔ και των αλληλεπιδράσεων μεταξύ δεδομένων και συναλλαγών
- Καθορισμός των κανόνων που διασφαλίζουν την ακεραιότητα των δεδομένων

Πορεία σχεδίασης

Ανάλυση απαιτήσεων (Requirements analysis): Καταγραφή της φύσης των δεδομένων, των χαρακτηριστικών της ΒΔ και της αναμενόμενης εξόδου (output)

Εννοιολογικός Σχεδιασμός (Conceptual design): Δημιουργία της δομής και διασφάλιση τήρησης των κανόνων καλής σχεδίασης (κανονικοποίηση)

Λογικός Σχεδιασμός (Logical design): Δημιουργία της ΒΔ μέσω εντολών γλώσσας DDL (*Data Definition Language*)

ER Modelling

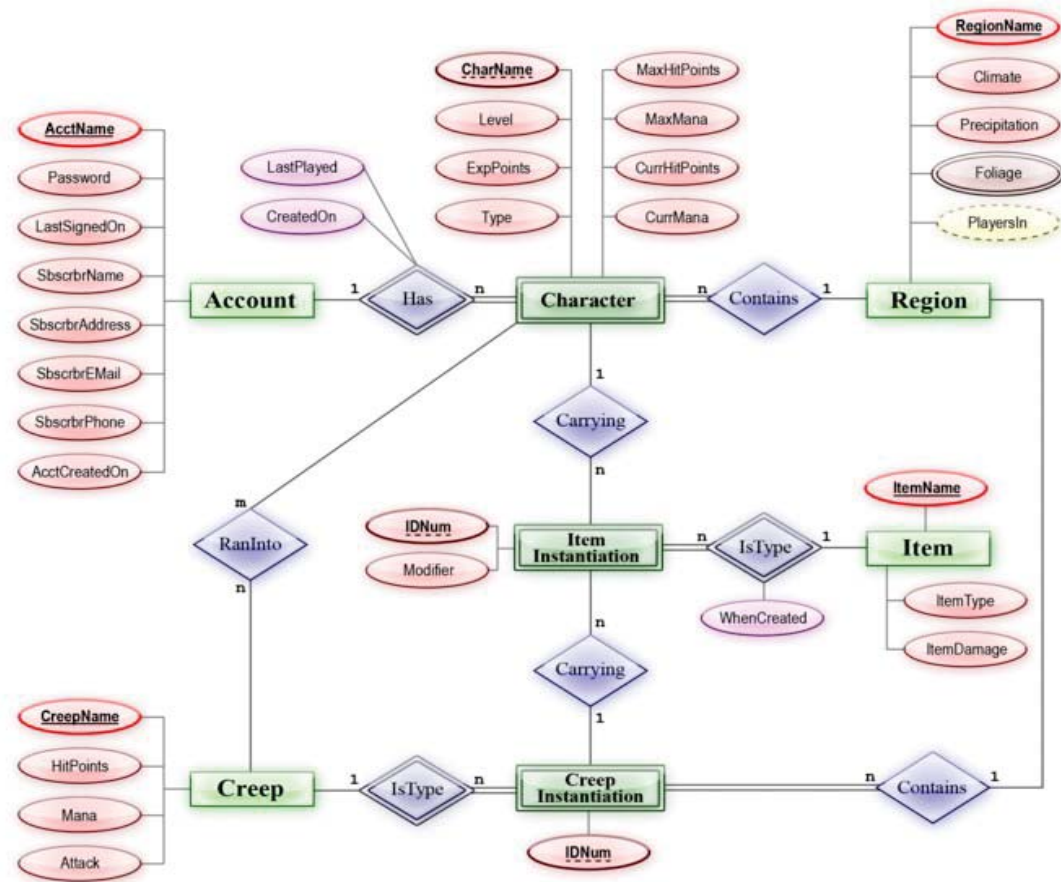
Μοντέλο οντοτήτων συσχετίσεων (Entity Relationship model, ER model): αφηρημένη και εννοιολογική αναπαράσταση των δεδομένων που χρησιμοποιείται από τους μηχανικούς λογισμικού.

Προτάθηκε από τον **Peter Chen το 1976**

Χρησιμοποιείται κατά το στάδιο του **λογικού και εννοιολογικού σχεδιασμού (conceptual design)** της βάσης

Η διαδικασία αυτή καταλήγει στο να παραχθεί μια διαγραμματική αναπαράσταση του σχήματος της βάσης δεδομένων που ονομάζεται **διάγραμμα Οντοτήτων - Συσχετίσεων (Entity-Relationship diagram - ERD)**.

ERD - Παράδειγμα



Οντότητες

Οντότητα (entity): μια αυτόνομη και αυθύπαρκτη μονάδα που μπορεί να θεωρηθεί τόσο ως πομπός ως και ως αποδέκτης πληροφορίας

Οι οντότητες είναι αναγνωρίσιμες έννοιες, είτε αφηρημένες είτε συγκεκριμένες, σε άμεση σχέση με τη ΒΔ (π.χ. πρόσωπα, τόποι, γεγονότα κλπ)

Οι οντότητες συνήθως αποδίδονται με ουσιαστικά π.χ. «εταιρεία», «εργαζόμενος», «διδάσκων», «μαθητής», κλπ.

Στις σχεσιακές ΒΔ οι οντότητες αποδίδονται με **πίνακες**

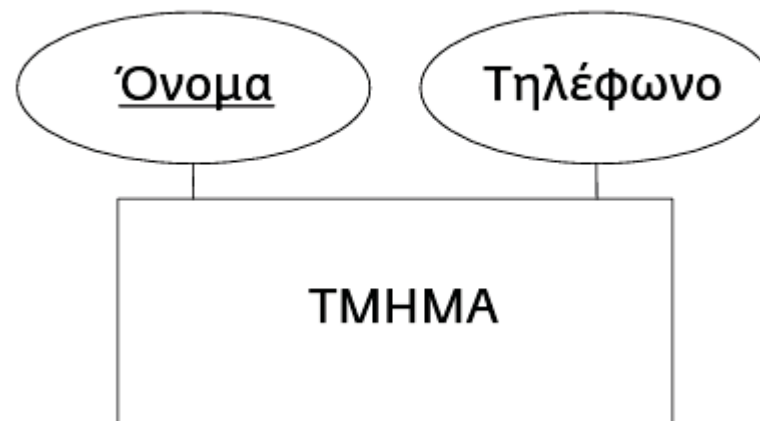
Πεδία

Η κάθε οντότητα περιγράφεται από ένα σύνολο ιδιοτήτων οι οποίες στην ορολογία του μοντέλου, ονομάζονται **χαρακτηριστικά ή γνωρίσματα ή πεδία (attributes)**

Συνήθως τα πεδία παίρνουν απλές τιμές που δεν μπορούν να διαχωριστούν σε μικρότερες μονάδες. (π.χ. η ηλικία) και ονομάζονται **απλά πεδία (simple attributes)** ή **ατομικά πεδία (atomic attributes)**.

Σε κάποιες περιπτώσεις, ένα πεδίο είναι δυνατό να διασπαστεί σε ένα πλήθος απλών πεδίων. Στην περίπτωση αυτή, το πεδίο λέγεται **σύνθετο (complex attribute)**.

Συμβολισμός οντοτήτων και πεδίων



Τύποι οντοτήτων

Ένα βασικό χαρακτηριστικό που εμφανίζεται σε όλες σχεδόν τις ΒΔ, είναι η ομαδοποίηση των οντοτήτων τους με κριτήριο τα κοινά τους πεδία (γνωρίσματα).

Το σύνολο όλων αυτών των οντοτήτων οι οποίες περιγράφονται από τα ίδια πεδία, ορίζει ένα **τύπο οντότητας (entity type)**

Κάθε ένας από αυτούς τους τύπους, περιγράφεται από ένα όνομα και από το σύνολο των πεδίων που περιέχει.

Το σύνολο αυτής της πληροφορίας ορίζει το **σχήμα (schema)** για το συγκεκριμένο τύπο οντότητας.



Κενή ή Άκυρη Τιμή (NULL)

Όταν για κάποια οντότητα δεν ορίζεται ή δεν είναι γνωστή η τιμή κάποιου από τα πεδία που περιέχει, το πεδίο αυτό θα λάβει την επονομαζόμενη **κενή ή άκυρη τιμή (NULL value)**.

Απλά και Σύνθετα Πεδία

Όταν τα χαρακτηριστικά μιας οντότητα παίρνουν απλές τιμές, που δεν μπορούν να διαχωριστούν σε μικρότερες μονάδες, αυτά ονομάζονται **απλά πεδία / χαρακτηριστικά / γνωρίσματα (simple attributes)** ή **ατομικά πεδία / χαρακτηριστικά / γνωρίσματα (atomic attributes)**.

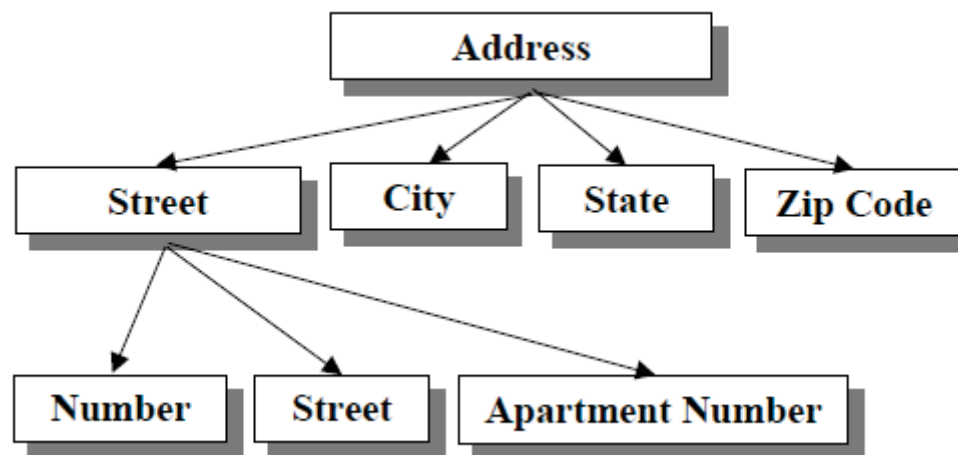
Σε άλλες περιπτώσεις, ένα χαρακτηριστικό είναι δυνατό να διασπαστεί σε ένα πλήθος απλών χαρακτηριστικών. Στην περίπτωση αυτή, το χαρακτηριστικό λέγεται **σύνθετο (complex attribute)**. Τότε η τιμή που αποδίδεται σε αυτό προκύπτει από τη **συνένωση (concatenation)** των τιμών των απλών χαρακτηριστικών που περιέχει.

Ανάμεσα στα σύνθετα χαρακτηριστικά μπορεί να υπάρξει και η έννοια της **ιεραρχίας (hierarchy)**, δηλαδή των πολλαπλών επιπέδων οργάνωσης της πληροφορίας του σύνθετου χαρακτηριστικού.

Σύνθετα πεδία - Ιεραρχία

Ανάμεσα στα σύνθετα πεδία μιας οντότητας μπορεί να υπάρξει και ένα είδος **ιεραρχίας (hierarchy)** :

Π.χ. 341 Taylor St.
Appartment 3
Sacramento
3467 CA



Συσχετιζόμενα πεδία

Δύο ή περισσότερα πεδία λέμε ότι **συσχετίζονται μεταξύ τους (related attributes)** όταν το ένα επιτρέπει τον υπολογισμό της τιμής του άλλου. *Παράδειγμα: ημερομηνία γέννησης και ηλικία*

Το πεδίο που προκύπτει από άλλο λέγεται **συναγόμενο**.

Αν το χαρακτηριστικό **A** είναι εκείνο που καθορίζει την τιμή του χαρακτηριστικού **B**, τότε λέμε ότι το **B** είναι **συναρτησιακά εξαρτημένο από το A** και γράφουμε **B(A)** ή **A→B**

Τιμές πεδίων

Αν ένα πεδίο έχει μόνο μία τιμή (π.χ. η ηλικία) καλείται **πεδίο απλής τιμής (single-valued attribute)**

Αν ένα πεδίο μπορεί να λάβει περισσότερες από μία τιμές ταυτόχρονα (π.χ ο αριθμός τηλεφώνου), τότε είναι ένα **πεδίο πολλαπλών τιμών (multivalued attribute)**.

Στην περίπτωση που ένα πεδίο μπορεί να λάβει περισσότερες από μια τιμές, μπορούμε να θέσουμε περιορισμούς σχετικά με το πλήθος των τιμών

Πρωτεύον πεδίο

Πρωτεύον πεδίο (π.π.): μοναδική τιμή για κάθε στιγμιότυπο. *Παράδειγμα : το ΑΦΜ*

Συνθήκη μοναδικότητας (uniqueness constraint): Δεν είναι δυνατό δύο στιγμιότυπα να έχουν την ίδια τιμή στο π.π. τους.

Το π.π. ενός τύπου οντότητας, μπορεί να μην είναι απλό, αλλά σύνθετο, να αποτελείται δηλαδή από πολλά απλά πεδία.

Ο συνδυασμός των τιμών των απλών πεδίων που αποτελεί και την τιμή του σύνθετου πεδίου, θα πρέπει να είναι μοναδικός για κάθε στιγμιότυπο.

Παραδείγματα συμβολισμού

Παραδείγματα των τεσσάρων ειδών πεδίων (γνωρισμάτων/χαρακτηριστικών) :

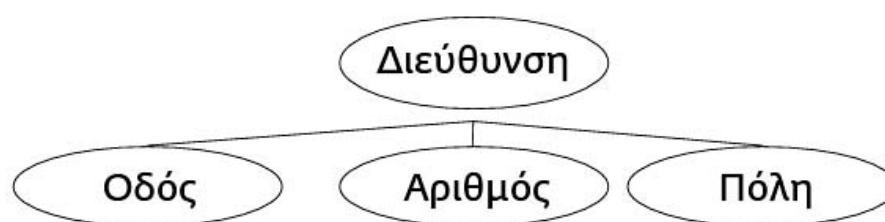
(α) πρωτεύον

(β) σύνθετο

(γ) πολλαπλών τιμών (δ) συναγόμενο



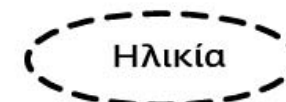
(α)



(β)



(γ)



(δ)

Στιγμιότυπα οντότητας

Οι καταχωρήσεις για τις οποίους τα πεδία του τύπου οντότητας παίρνουν εντελώς συγκεκριμένες τιμές, αποτελούν **στιγμιότυπα (instances)** ή **αντιπροσώπους** αυτού του σχήματος

**EMPLOYEE
SCHEMA**

First Name,	Last Name,	Age,	Salary
-------------	------------	------	--------

EMPLOYEE INSTANCES

EMPLOYEE e₁

John	Smith	55	80000
------	-------	----	-------

EMPLOYEE e₂

Fred	Brown	40	30000
------	-------	----	-------

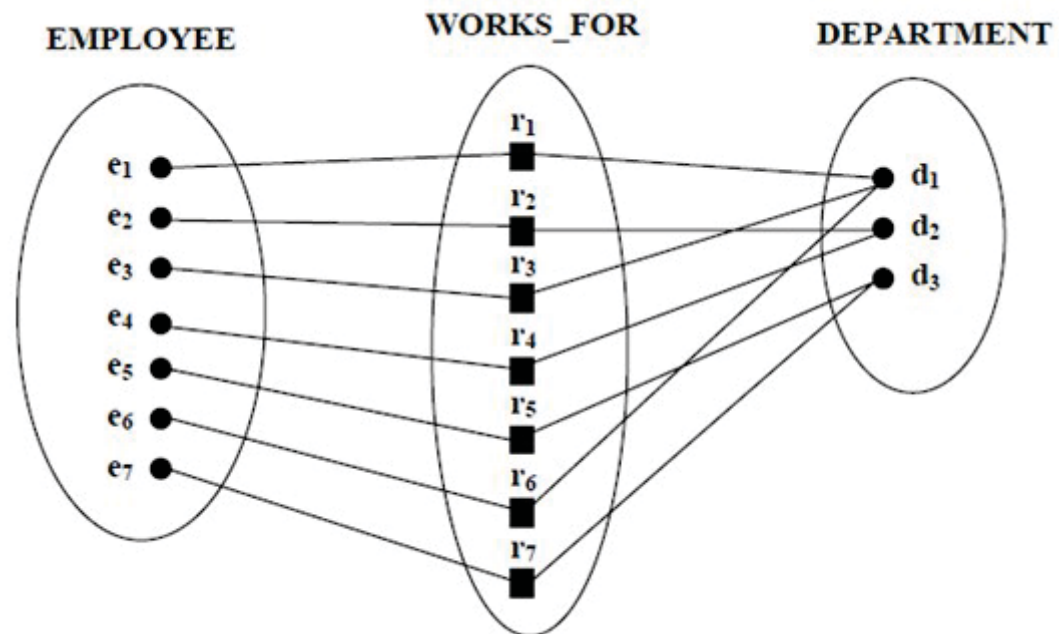
EMPLOYEE e₃

Judy	Clark	25	20000
------	-------	----	-------

Συσχέτιση

Συσχέτιση (relationship): καθορίζει με ποιο τρόπο οι οντότητες μιας ΒΔ, συσχετίζονται μεταξύ τους

Στιγμιότυπο συσχέτισης : Η συσχέτιση μεταξύ στιγμιοτύπων



Τύπος συσχέτισης

Τύπος συσχέτισης R είναι ένα πλήθος από στιγμιότυπα συσχετίσεων r_i , ανάμεσα σε στιγμιότυπα των τύπων οντότητας που συσχετίζονται.

Ο τύπος συσχέτισης, εκτός από τους τύπους οντότητας μεταξύ των οποίων ορίζεται, μπορεί να χαρακτηρίζεται και από την ύπαρξη δικών του πεδίων (*attributes*)

Βαθμός συσχέτισης

Βαθμός (degree) της συσχέτισης R: Το πλήθος n των τύπων οντότητας που συσχετίζονται δια μέσου μιας συσχέτισης.

Συνήθως ο βαθμός συσχέτισης είναι **2 (binary relationship)**, ενώ σε κάποιες περιπτώσεις μπορεί να είναι **3 (ternary relationship)**

Αποδεικνύεται πως στη γενική περίπτωση, μια σχέση με βαθμό 3 ανάμεσα σε τρεις τύπους οντότητας A, B και C, περιέχει περισσότερη πληροφορία από τις τρεις συσχετίσεις με βαθμό 2 που ορίζονται ανάμεσα σε αυτούς τους τρεις τύπους οντότητας.

Ρόλος

Κάθε τύπος οντότητας που συμμετέχει σε ένα τύπο συσχέτισης παίζει και ένα συγκεκριμένο **ρόλο (role)** σε αυτή τη συσχέτιση.

Αν οι τύποι οντότητας που συμμετέχουν σε ένα τύπο συσχέτισης είναι διαφορετικοί μεταξύ τους, ο καθορισμός του ρόλου για τον κάθε τύπο οντότητας δεν είναι μια αναγκαία διαδικασία.

Εάν ένας τύπος οντότητας συμμετέχει περισσότερες από μια φορές σε ένα τύπο συσχέτισης, τότε ο καθορισμός του ρόλου για τον κάθε τύπο οντότητας, μας επιτρέπει να περιγράψουμε καλύτερα τον τύπο αλληλεπίδρασης που υφίσταται ανάμεσα στα δύο αντίγραφα του.

Πολλαπλότητα

Η **πολλαπλότητα (cardinality)** ενός τύπου συσχέτισης καθορίζει το πλήθος των στιγμιότυπων αυτού του τύπου συσχέτισης, στον οποίο μια οντότητα μπορεί να συμμετάσχει.

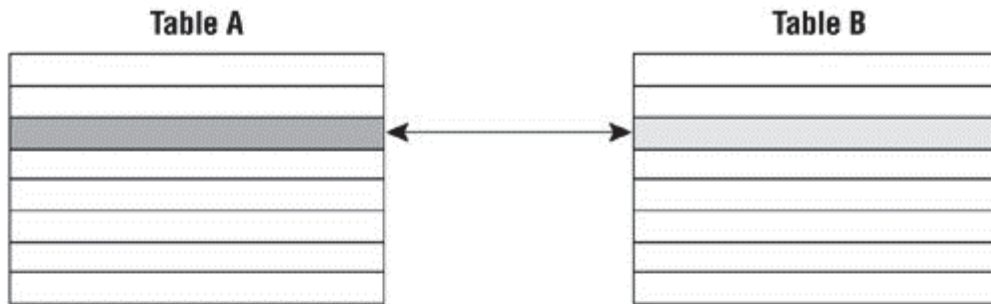
- **1:N (ένα προς πολλά)**
- **1:1 (ένα προς ένα) ή**
- **M:N (πολλά προς πολλά)**

Τύπος συσχέτισης και πολλαπλότητα

- Αν η πολλαπλότητα του τύπου συσχέτισης είναι 1:1 ή 1:N, τότε τα τυχόν πεδία του τύπου συσχέτισης μπορούν να περιληφθούν σε έναν τύπο οντότητας
- Στην περίπτωση 1:1 τα πεδία μπορούν να προστεθούν σε οποιοδήποτε από τους τύπους οντότητας
- Στην περίπτωση 1:N τα πεδία μπορούν να προστεθούν ΜΟΝΟ στον τύπο οντότητας που αντιστοιχεί στην πλευρά του N
- Εάν η σχέση είναι M:N, τότε επειδή τα νέα πεδία δεν αναφέρονται σε ένα από τα δύο στιγμιότυπα των τύπων οντότητας αλλά στο συνδυασμό τους, δεν είναι δυνατή η προσθήκη τους σε κάποιον από τους τύπους οντότητας, και επομένως τα προσθέτουμε υποχρεωτικά στον τύπο συσχέτισης



Σχέσεις 1:1



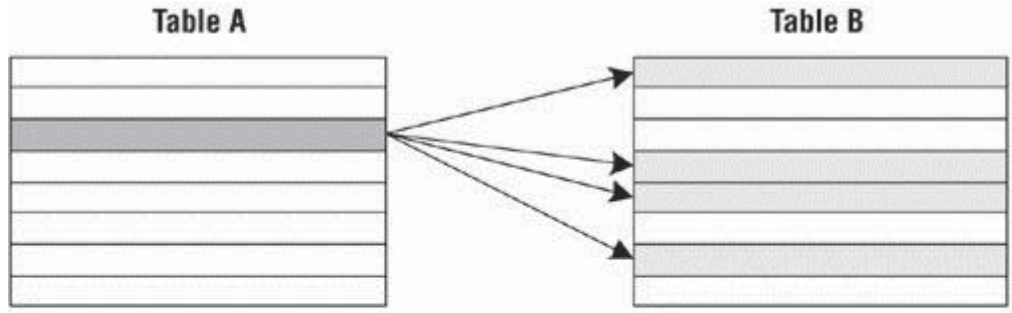
Employees

Employee ID	Employee First Name	Employee Last Name	Home Phone	<< other fields >>
100	Zachary	Erlich	553-3992
101	Susan	McLain	790-3992
102	Joe	Rosales	551-4993

Compensation

Employee ID	Hourly Rate	Commission Rate	<< other fields >>
100	25.00	5.0%
101	19.75	3.5%
102	22.50	5.0%

Σχέσεις 1:N

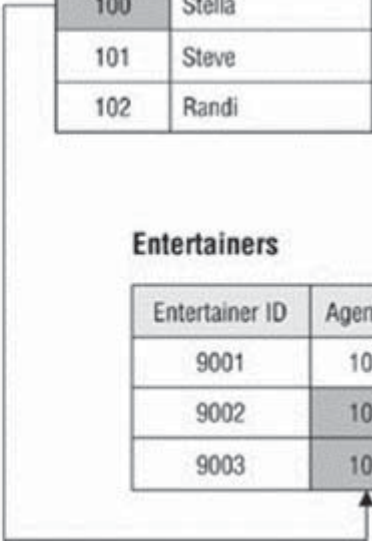


Agents

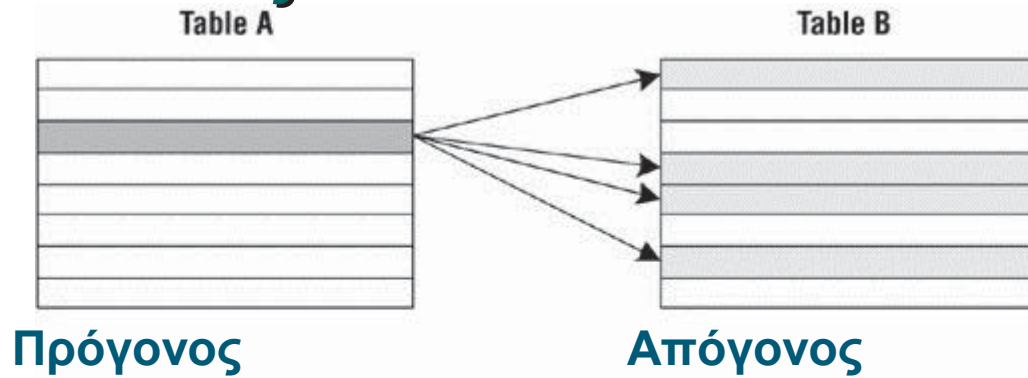
Agent ID	Agent First Name	Agent Last Name	Date of Hire	Agent Home Phone
100	Stella	Rosales	05/16/95	553-3992
101	Steve	Pundt	10/15/95	790-3992
102	Randi	Nathanson	03/01/96	551-4993

Entertainers

Entertainer ID	Agent ID	Entertainer Name	Entertainer Phone	<< other fields >>
9001	101	Jazz Time	555-9928
9002	100	The Mike Hernandez Trio	959-8837
9003	100	The Country Squires	709-3542



Σχέσεις 1:N



Agents

Agent ID	Agent First Name	Agent Last Name	Date of Hire	Agent Home Phone
100	Stella	Rosales	05/16/95	553-3992
101	Steve	Pundt	10/15/95	790-3992
102	Randi	Nathanson	03/01/96	551-4993

Entertainers

Entertainer ID	Agent ID	Entertainer Name	Entertainer Phone	<< other fields >>
9001	101	Jazz Time	555-9928
9002	100	The Mike Hernandez Trio	959-8837
9003	100	The Country Squires	709-3542

Ταυτίζουσες και Μη Σχέσεις

- Μία σχέση καλείται **ταυτίζουσα – identifying** όταν το πρωτεύον πεδίο του πίνακα-προγόνου συμμετέχει στο πρωτεύον πεδίο του πίνακα-απογόνου.
- Μία σχέση καλείται **μη ταυτίζουσα – non-identifying** όταν το πρωτεύον πεδίο του πίνακα-προγόνου εμφανίζεται σαν εξωτερικό πεδίο στον πίνακα-απόγονο και δεν συμμετέχει στο πρωτεύον πεδίο του.
- Η ταυτίζουσα σχέση αποδίδεται με συνεχή γραμμή συσχέτισης, ενώ η μη ταυτίζουσα σχέση σημειώνεται συχνά και με διακεκομμένη γραμμή (crow's foot).
-

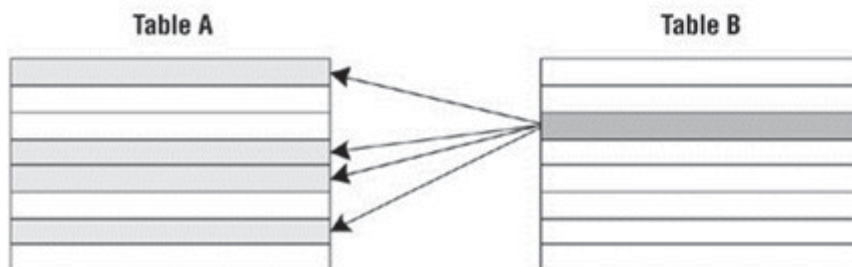
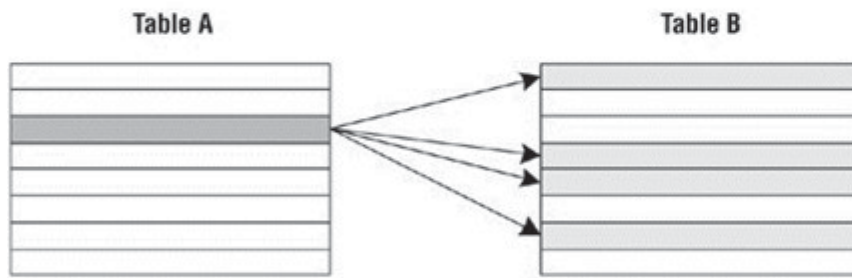
Σχέσεις Μ:Ν

Students

Student ID	Student First Name	Student Last Name	Student Phone	<< other fields >>
60001	Zachary	Erlch	553-3992
60002	Susan	McLain	790-3992
60003	Joe	Rosales	551-4993

Classes

Class ID	Class Name	Instructor ID	<< other fields >>
900001	Intro. to Political Science	220087
900002	Adv. Music Theory	220039
900003	American History	220148



Σχέσεις Μ:Ν – Ανάλυση σε 1:Μ και Ν:1

Students

Student ID	Student First Name	Student Last Name	Student Phone	<< other fields >>
60001	Zachary	Erlich	553-3992
60002	Susan	McLain	790-3992
60003	Joe	Rosales	551-4993

Student Schedule (Linking Table)

Student ID	Class ID
60003	900001
60001	900003
60003	900003
60002	900002
60001	900001

Classes

Class ID	Class Name	Instructor ID	<< other fields >>
900001	Intro. to Political Science	220087
900002	Adv. Music Theory	220039
900003	American History	220148

Συμμετοχή

- Ένας τύπος οντότητας χαρακτηρίζεται από **ολική / υποχρεωτική συμμετοχή (total / mandatory participation)** σε ένα *τύπο συσχέτισης*, εάν όλα τα στιγμιότυπά του οφείλουν να συμμετέχουν στον τύπο συσχέτισης, ή εναλλακτικά, αν η ύπαρξή του εξαρτάται από τη συμμετοχή του ή όχι σε αυτή τη συσχέτιση
- Σε διαφορετική περίπτωση ο τύπος οντότητας χαρακτηρίζεται από **μερική / προαιρετική συμμετοχή (partial/optional participation)**

Βαθμός συμμετοχής

- **Βαθμός συμμετοχής:** καθορίζει τον ελάχιστο αριθμό καταχωρήσεων που ένας δεδομένος πίνακας πρέπει να έχει συσχετισμένες με μία καταχώρηση στον δεύτερο πίνακα, καθώς και το μέγιστο αριθμό καταχωρήσεων που ο δεδομένος πίνακας επιτρέπεται να έχει συσχετισμένες με μία καταχώρηση στον δεύτερο πίνακα.

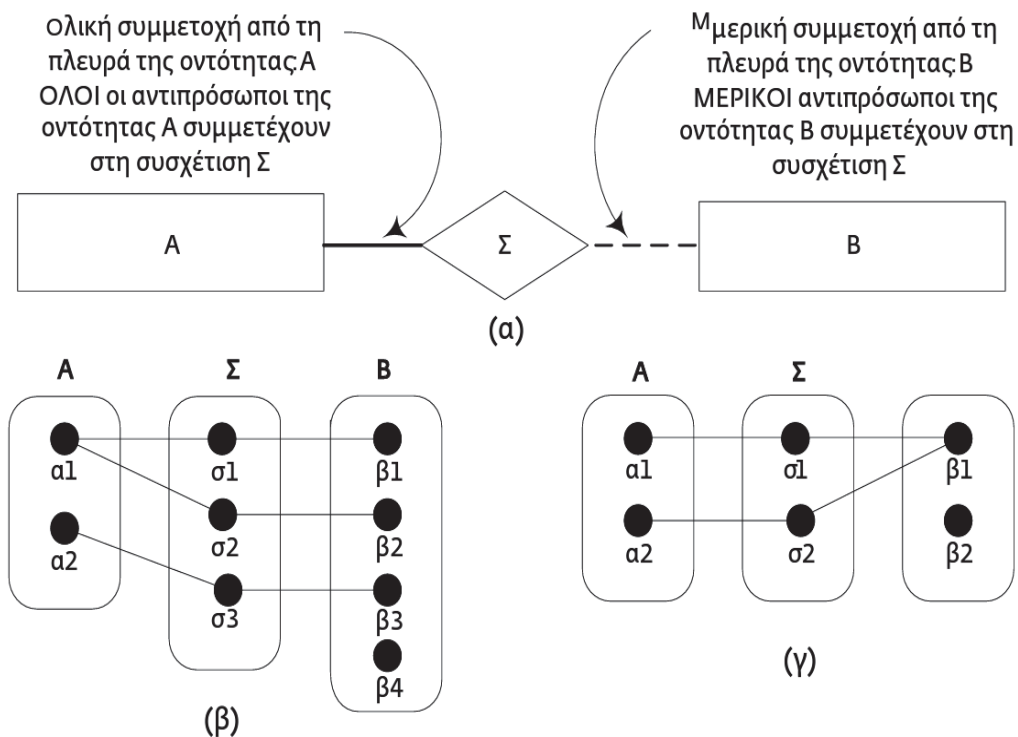
Agents

Agent ID	Agent First Name	Agent Last Name	Date of Hire	Agent Home Phone
100	Stella	Rosales	05/16/95	553-3992
101	Steve	Pundt	10/15/95	790-3992
102	Randi	Nathanson	03/01/96	551-4993

Clients

Client ID	Agent ID	Client First Name	Client Last Name	Client Home Phone
9001	100	Stewart	Jameson	553-3992
9002	101	Shannon	McLain	790-3992
9003	102	Scott	Barker	551-4993

Συμβολισμός συμμετοχής



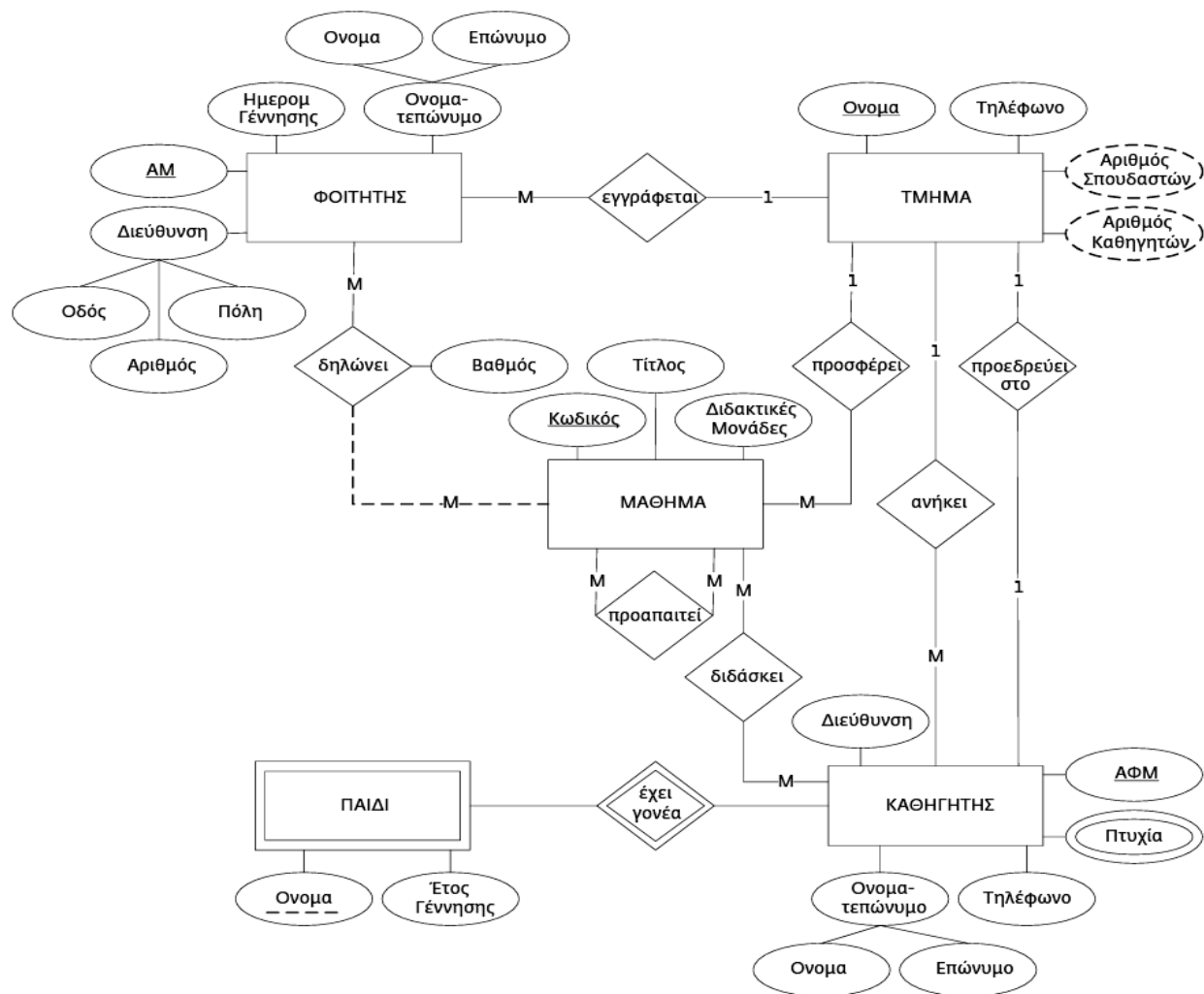
Εικόνα 2.4. Παράδειγμα συμμετοχή συσχέτισης. (α) Η συσχέτιση Σ έχει ολική συμμετοχή από την πλευρά της οντότητας A και μερική από την πλευρά της οντότητας B. (β), (γ) Παραδείγματα συσχέτισης αντιπροσώπων με ολική συμμετοχή των αντιπροσώπων της οντότητας A και μερική συμμετοχή των αντιπροσώπων της οντότητας B.

Αδύναμος τύπος οντότητας

Αδύναμος τύπος οντότητας (weak entity type): είναι ο τύπο οντότητας για τον οποίο δεν μπορούμε να ορίσουμε πρωτεύον πεδίο, ώστε να ξεχωρίζουν τα στιγμιότυπά του. Οι αδύναμοι τύποι οντότητας ορίζονται μόνο από ένα συνδυασμό κάποιων από τα πεδία που περιέχουν καθώς και από τη **προσδιορίζουσα συσχέτισή** τους (identifying relationship) με κάποια **προσδιορίζουσα οντότητα** (identifying owner entity).

Σε αυτή τη συσχέτιση, οι αδύναμοι οντότητας έχουν **ολική συμμετοχή** (total participation) και συμβολίζονται με **διπλό παραλληλόγραμμο**

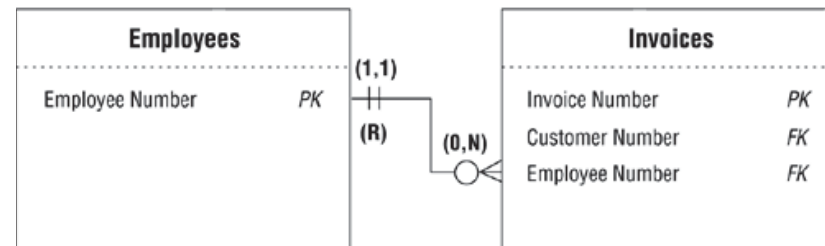
Παράδειγμα ERD κατά Chen



ERD : Crow's Foot

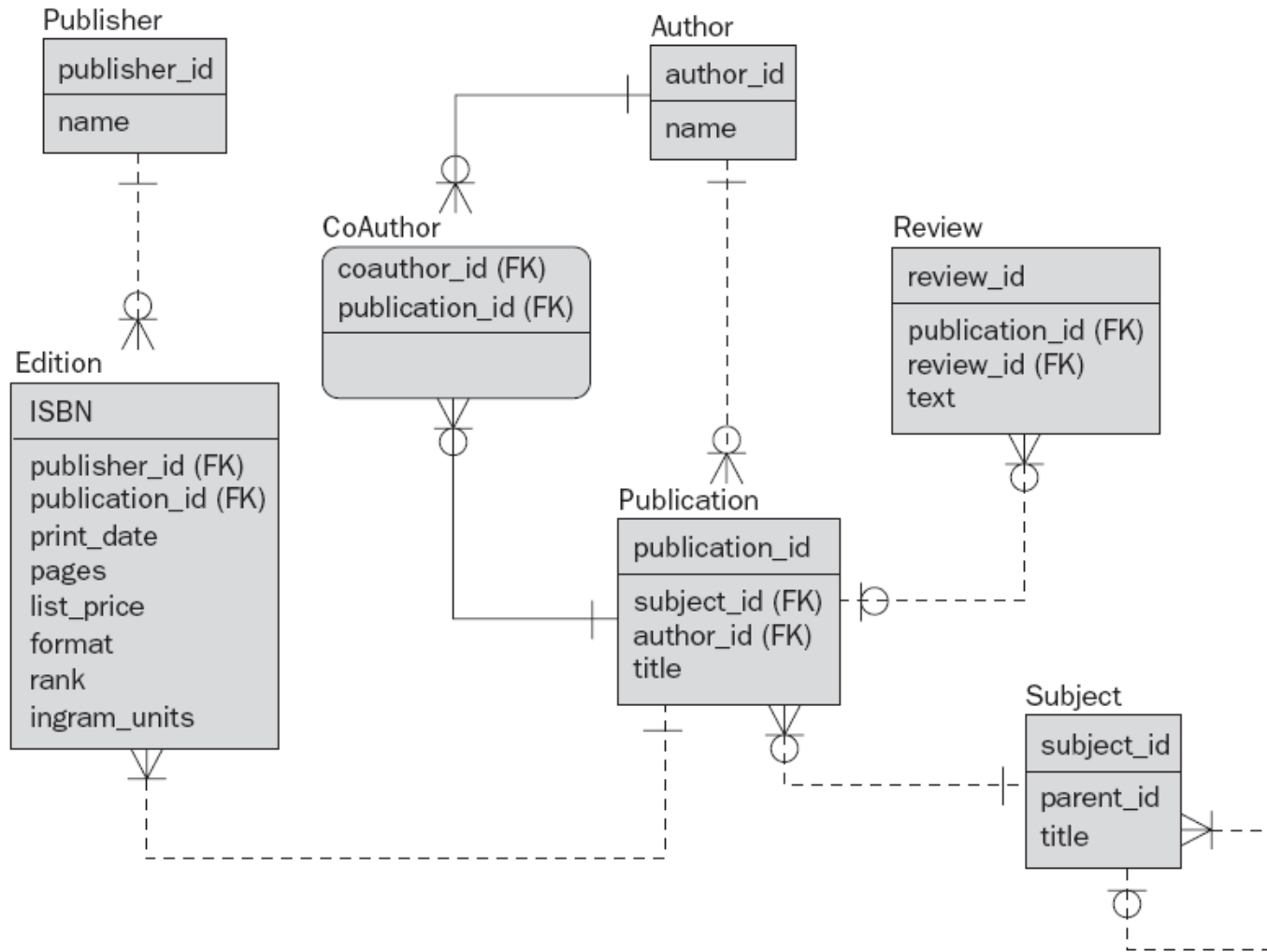
- **Entity Relationship Diagram (ERD)** : Σύνολο πινάκων, γραμμών και συμβόλων που προσδιορίζουν το σχήμα μιας βάσης
- Για το σχεδιασμό του ERD πρέπει να προσδιοριστούν τα εξής :
 - Οι οντότητες και τα πεδία τους
 - Οι συσχετίσεις μεταξύ των οντοτήτων
 - Η κατεύθυνση των σχέσεων (πρόγονος-απόγονος)
 - Το είδος των σχέσεων

Είδη Σχέσεων και Απόδοση Λόγου



Χαρακτηρισμός	Σύμβολο	Σημασία
Δακτύλιος	⊖	Κανένα
Κάθετος	⊥	Ένα
Πόδι κορακιού	⊂	Πολλά
Δακτύλιος και Κάθετος	⊖⊥	Ένα ή κανένα (Μπορεί να έχει ένα...)
Διπλή Κάθετος	⊥⊥	Ακριβώς ένα
Δακτύλιος με Πόδι	⊖⊂	Κανένα ή περισσότερα (Μπορεί να έχει...)
Κάθετος με Πόδι	⊥⊂	Ένα ή περισσότερα (Έχει τουλάχιστον ένα...)

Παράδειγμα ERD με σήμανση γραμμών



Παράδοση 4 : Το Σχεσιακό Μοντέλο

- Πότε καθιερώθηκε το σχεσιακό μοντέλο αναπαράστασης;
- Πώς αναπαρίστανται τα δεδομένα στο μοντέλο αναπαράστασης;
- Τι είναι οι πλειάδες (tuples);
- Τι αναπαριστά κάθε γραμμή ενός πίνακα στο μοντέλο αναπαράστασης;
- Ποιος περιορισμός ισχύει για τις τιμές κάθε στήλης ενός πίνακα;
- Πως ορίζεται το σχήμα μιας σχέσης;
- Ποιος είναι ο βαθμός μιας σχέσης;
- Πως ορίζεται το στιγμιότυπο μιας σχέσης;
- Τι σημασία έχει η διάταξη των πλειάδων σε μια σχέση;
- Τι σημασία έχει η διάταξη των πεδίων της πλειάδας;
- Τι είναι οι ατομικές τιμές;
- Πότε χρησιμοποιούμε το NULL σε ένα πεδίο;
- Γιατί πρέπει να είναι οι πλειάδες διαφορετικές μεταξύ τους και πώς εξασφαλίζεται αυτό;
- Τι είναι το υπέρ κλειδί;
- Τι σχέση έχει το πεδίο κλειδί με το υπέρ κλειδί;
- Τι είναι το πρωτεύον κλειδί και πώς σημειώνεται;
- Τι είναι το σχεσιακό σχήμα μιας βάσης δεδομένων;
- Τι είναι και τι λένε οι κανόνες ακεραιότητας;
- Τι συνεπάγεται η μη τήρηση του 3^{ου} κανόνα ακεραιότητας;
- Σε περίπτωση μη τήρησης του 3^{ου} κανόνα ακεραιότητας κατά τη διαδικασία διαγραφής εγγραφής, ποιες κινήσεις μπορούν να γίνουν;
- Τι είναι το ξένο κλειδί;
- Ποιο είναι το χαρακτηριστικό της αναδρομικής συσχέτισης;
- Τι είναι η απεικόνιση και σε πόσα βήματα συντελείται ;
- Ποιο είναι το 1^ο βήμα κατά την απεικόνιση;
- Πώς αντιμετωπίζονται οι αδύναμοι τύποι οντότητας κατά την απεικόνιση;
- Πώς χειριζόμαστε τους τύπους συσχέτισης με βαθμό 2 και πολλαπλότητα 1:1 κατά την απεικόνιση;
- Πώς χειριζόμαστε τους τύπους συσχέτισης με βαθμό 2 και πολλαπλότητα 1:N κατά την απεικόνιση;
- Πώς χειριζόμαστε τους τύπους συσχέτισης με βαθμό 2 και πολλαπλότητα M:N κατά την απεικόνιση;
- Πώς χειριζόμαστε τα πεδία πολλαπλής τιμής κατά την απεικόνιση;

Το σχεσιακό μοντέλο

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Στόχος και αποτελέσματα

Στόχος του κεφαλαίου αυτού είναι να παρουσιάσει αναλυτικά το σχεσιακό μοντέλο που είναι το πιο δημοφιλές.

Προσδοκώμενα αποτελέσματα: Όταν ολοκληρώσετε τη μελέτη αυτού του κεφαλαίου, θα έχετε κατανοήσει:

- Τις έννοιες πίνακας, πρωτεύον κλειδί, ξένο κλειδί, που είναι οι βασικές έννοιες του σχεσιακού μοντέλου.
- Τον τρόπο οργάνωσης των δεδομένων σε πίνακες, που αποτελεί τη μόνη οργανωτική δομή του σχεσιακού μοντέλου για την αποθήκευση των δεδομένων.
- Τους περιορισμούς που υπάρχουν στο σχεσιακό μοντέλο και σχετίζονται με το πρωτεύον και το ξένο κλειδί.
- Τις πράξεις της εισαγωγής, διαγραφής και ενημέρωσης που είναι διαθέσιμες στο σχεσιακό μοντέλο.

Βιβλιογραφική παραπομπή :

- Κεχρής, Ε., “Σχεσιακές βάσεις δεδομένων – Θεωρία & εργαστηριακές ασκήσεις”, Κεφ. 4,5
- Μάργαρης, Α.Ι., “Εισαγωγή στις Βάσεις Δεδομένων” – Κεφ. 3

Σχεσιακό Μοντέλο (1)

- Το **σχεσιακό μοντέλο αναπαράστασης (relational data model)** καθιερώθηκε από τον Codd το 1970
- Αποτελεί ένα από τα πιο απλά και ευέλικτα μοντέλα αυτού του είδους.
- Τα δεδομένα αναπαρίστανται ως ένα σύνολο από πίνακες.
- Συνήθως υιοθετείται η χρήση **πινάκων (tables)** που περιέχουν ένα πλήθος **γραμμών (rows)** και **στηλών (columns)**.

Σχεσιακό Μοντέλο (2)

- Οι γραμμές ονομάζονται και **πλειάδες (tuples)** ή **εγγραφές (records)**
- Η πλειάδα περιέχει ένα σύνολο **απλών πεδίων (attributes)** τα οποία συσχετίζονται μεταξύ τους.
- Οι πίνακες χρησιμοποιούνται για την αναπαράσταση των τύπων οντοτήτων καθώς και των τύπων συσχετίσεων που υφίστανται ανάμεσά τους
- Κάθε μια από τις γραμμές ενός πίνακα θεωρείται ένα **στιγμιότυπο οντότητας ή συσχέτισης** ανάλογα με το αντικείμενο στο οποίο αναφέρεται.

Παράδειγμα πίνακα

- Η κάθε πλειάδα ή εγγραφή (record) αναφέρεται και σε ένα ξεχωριστό στιγμιότυπο του τ. οντότητας STUDENT
- Η κάθε μια από αυτές τις εγγραφές αποτελείται από ένα σύνολο χαρακτηριστικών ή πεδίων τα οποία λαμβάνουν συγκεκριμένες τιμές, διαφορετικές εν γένει για κάθε εγγραφή.
- Επειδή οι τιμές της κάθε στήλης αναφέρονται ουσιαστικά στο ίδιο πεδίο του πίνακα, είναι προφανές πως θα πρέπει να ανήκουν στον ίδιο τύπο δεδομένων.

STUDENT						
Name	SSN	Home Phone	Address	Office Phone	Age	GPA
Bayer	305612435	3731616	2918 Bluebonnet Lane	NULL	19	3.21
Ashly	381621245	3754409	125 Kirby Road	NULL	18	2.89
David	422112320	NULL	3452 Elgin Road	7491253	25	3.53
Cooper	489221100	3769821	265 Lark Lane	7496492	28	3.93
Benson	533691238	8398461	7384 Fontana Lane	NULL	19	3.25

Σύνολο τιμών (domain)

- Κάθε πεδίο ή στήλη ενός πίνακα, δέχεται τιμές από ένα συγκεκριμένο και καθορισμένο εκ των προτέρων **σύνολο τιμών (domain)**
- Το είδος των τιμών αυτού του συνόλου καθορίζεται από τον **τύπο δεδομένων** του πεδίου του πίνακα
- Ο τύπος δεδομένων ορίζεται κατά το στάδιο της λογικής σχεδίασης
- Όσον αφορά το σύνολο των επιτρεπτών τιμών για αυτό το πεδίο, αυτό εξαρτάται και πάλι από τη φύση του προβλήματος.

Σχήμα σχέσης

- **Σχήμα σχέσης (relation schema) R** : ένα σύνολο πεδίων $\{A_1, A_2, \dots, A_n\}$ κάθε ένα εκ των οποίων παίρνει τιμές μέσα από ένα σύνολο τιμών (**domain**) D
- Για κάθε πεδίο A_i , το πεδίο τιμών του, D , θα συμβολίζεται με **$\text{dom}(A_i)$**
- Το πλήθος των πεδίων του συνόλου R ονομάζεται **βαθμός (degree)** της σχέσης
- Η συμβολοσειρά R είναι το **όνομα της σχέσης**

R A_1 A_2 A_3 A_4 A_5 A_6 A_7
STUDENT {Name, SSN, HomePhone, Address, OfficePhone, Age, GPA}

Για $R = \langle \text{STUDENT} \rangle$, η σχέση R έχει βαθμό 7

Στιγμιότυπο σχέσης

- **Σχέση ή στιγμιότυπο σχέσης (relation instance) r** , του σχήματος σχέσης $\mathbf{R}\{A_1, A_2, \dots, A_n\}$ ορίζεται ένα σύνολο από πλειάδες $r = \{t_1, t_2, t_3, \dots, t_m\}$ (**Πίνακας $m \times n$**)
- Κάθε πλειάδα είναι μια διατεταγμένη λίστα (ordered list) από n τιμές $t_i = (a_{i1}, a_{i2}, \dots, a_{in})$ με την κάθε τιμή a_{ij} ($1 \leq i \leq m, 1 \leq j \leq n$) να παίρνει τιμές από το πεδίο τιμών $\text{dom}(A_i)$ ενώ εναλλακτικά μπορεί να πάρει και την τιμή **NULL**
- Ένα στιγμιότυπο r ενός σχήματος σχέσης \mathbf{R} , μπορεί να γραφεί και ως **$r(\mathbf{R})$**
- Το σύνολο των πλειάδων που συνιστά το στιγμιότυπο σχέσης **$r(\mathbf{R})$** είναι υποσύνολο του καρτεσιανού γινομένου των συνόλων $\text{dom}(A_i)$:
 $r(\mathbf{R}) \subseteq (\text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n))$

Διάταξη

- Η **διάταξη (ordering)** των πλειάδων σε μια σχέση, δεν έχει καμιά σημασία. Αυτό γίνεται διότι μια σχέση ορίζεται ως ένα σύνολο από πλειάδες, και σύμφωνα με τον ορισμό του συνόλου, τα στοιχεία του δε χρειάζεται να είναι διατεταγμένα.
- Αντίθετα, η κάθε πλειάδα αποτελείται από ένα πλήθος διατεταγμένων πεδίων, και επομένως η σειρά με την οποία αναγράφονται οι τιμές μιας πλειάδας είναι σημαντική.

Ατομικές τιμές

- Οι τιμές που καταχωρούνται σε μια πλειάδα, θα πρέπει να είναι **ατομικές (atomic)** να μην υποδιαιρούνται δηλαδή σε μικρότερες μονάδες πληροφορίας.
- Σύνθετες τιμές δεδομένων **δεν** μπορούν να χρησιμοποιηθούν αλλά θα πρέπει να γραφούν με τρόπο που να μην παραβιάζει την παραπάνω αρχή
- Αυτός ο τρόπος φορμαλισμού ενός πίνακα έτσι ώστε όλα του τα πεδία να έχουν ατομικές τιμές, αναφέρεται και ως **πρώτη κανονική μορφή (1st Normal Form, 1NF)**
- Για όσα πεδία δεν γνωρίζουμε ή δεν διαθέτουμε την τιμή που τα χαρακτηρίζει, χρησιμοποιούμε γι' αυτά την τιμή **NULL**

Πεδία-κλειδιά

- Ένα από τα βασικά χαρακτηριστικά μιας σχέσης είναι η **μοναδικότητα** κάθε μιας από τις πλειάδες που περιέχει
- Οι πλειάδες πρέπει να είναι διαφορετικές μεταξύ τους ώστε να εξασφαλίζεται ότι σε καμιά περίπτωση δεν πρόκειται να καταχωρήσουμε δυο φορές την ίδια πλειάδα
- Σε καθεμιά από τις πλειάδες ενός πίνακα, υπάρχει ένας συνδυασμός πεδίων που ο συνδυασμός των τιμών αυτών των πεδίων είναι μοναδικός για κάθε εγγραφή

Υπέρ κλειδί

- **Υπερ-κλειδί (superkey):** Το σύνολο των πεδίων που ο συνδυασμός των τιμών τους είναι μοναδικός για κάθε εγγραφή του σχεσιακού σχήματος
- Η κάθε πλειάδα διαθέτει τουλάχιστον ένα superkey, που είναι το σύνολο των πεδίων που περιέχει
- Το **πεδίο κλειδί** μιας σχέσης είναι ένα superkey από το οποίο δεν μπορούμε να αφαιρέσουμε κανένα πεδίο (minimal superkey).

Υποψήφιο και πρωτεύον κλειδί

- **Υποψήφια κλειδιά (candidate keys):** Τα πεδία που να μπορούν να χρησιμοποιηθούν ως πεδία κλειδιά μιας σχέσης.
- **Πρωτεύον κλειδί (primary key):** ένα από υποψήφια κλειδιά που επιλέγεται για να γίνει το κλειδί του σχήματος της σχέσης.
- Η επιλογή του πρωτεύοντος κλειδιού εξαρτάται από τη μορφή της πληροφορίας που διατηρούμε για κάθε σχέση.
- Το πρωτεύον κλειδί ενός πίνακα σημειώνεται υπογραμμισμένο

Σχεσιακό σχήμα

- Σχεσιακό σχήμα μιας βάσης δεδομένων (**relational database schema**), S : ένα σύνολο σχημάτων σχέσεων $S = \{R_1, R_2, \dots, R_k\}$ και ένα σύνολο κανόνων ακεραιότητας (integrity constraint) για αυτά τα σχήματα
- Ένα **στιγμιότυπο του σχεσιακού σχήματος** της βάσης δεδομένων (**relational database instance**) ορίζεται ως ένα σύνολο στιγμιότυπων σχέσεων $\{r_1, r_2, \dots, r_m\}$ με το κάθε στιγμιότυπο r_i της σχέσης R_i να υπακούει στους κανόνες ακεραιότητας του σχεσιακού σχήματος της βάσης

Παράδειγμα σχήματος ΒΔ

EMPLOYEE

FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
-------	-------	-------	------------	-------	---------	-----	--------	----------	-----

DEPARTMENT

DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE
-------	----------------	--------	--------------

DEPT_LOCATIONS

<u>DNUMBER</u>	<u>DLOCATION</u>
----------------	------------------

DEPENDENT

<u>ESSN</u>	<u>DEPENDENT-NAME</u>	SEX	BDATE	RELATIONSHIP
-------------	-----------------------	-----	-------	--------------

PROJECT

PNAME	<u>PNUMBER</u>	PLOCATION	DNUM
-------	----------------	-----------	------

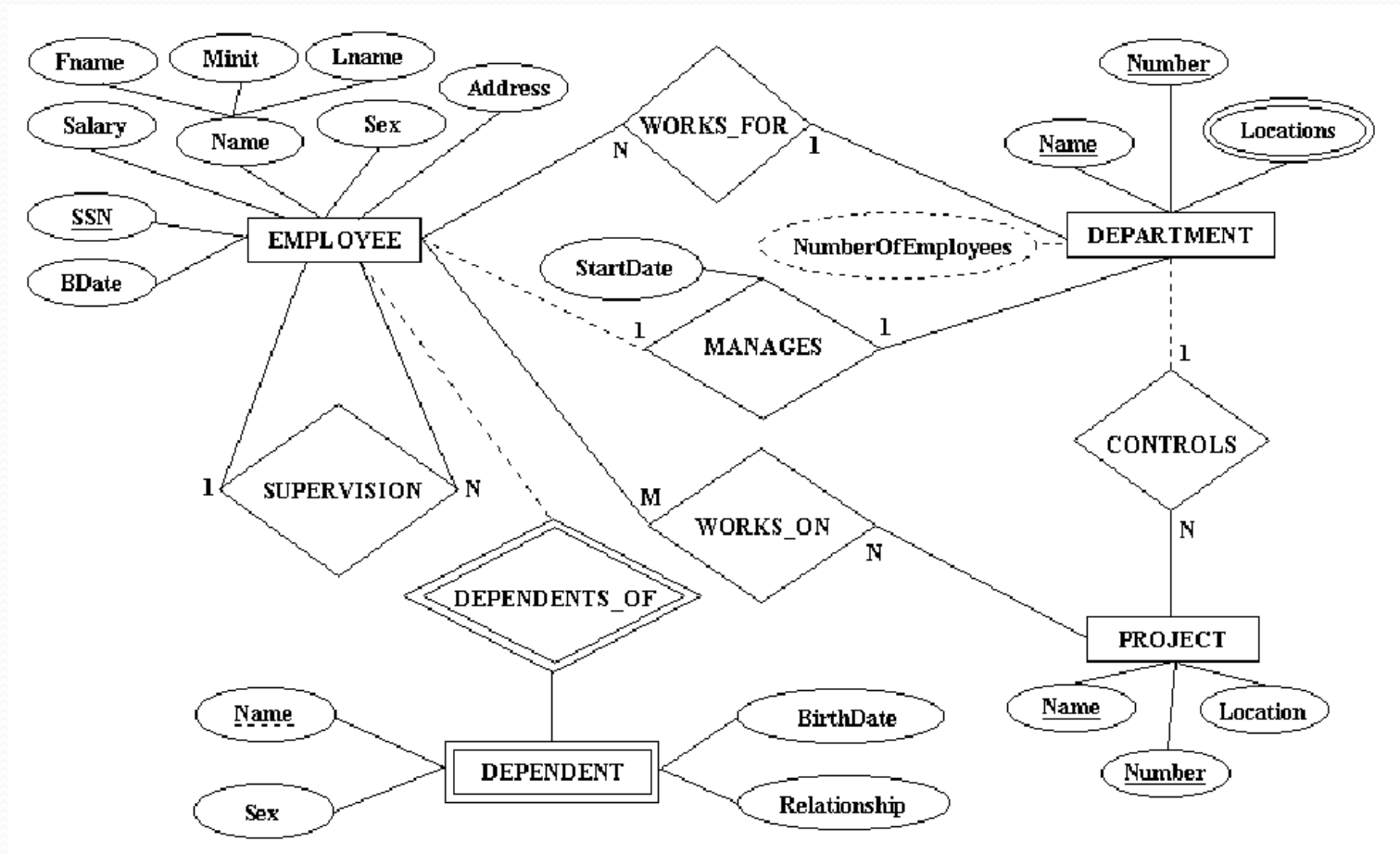
WORKS_ON

<u>ESSN</u>	<u>PNO</u>	HOURS
-------------	------------	-------

Ξένο κλειδί

- **Ξένο κλειδί (foreign key):** Ένα πεδίο ενός πίνακα **A** ονομάζεται ξένο κλειδί, όταν αποτελεί πρωτεύον κλειδί κάποιου άλλου πίνακα **B** και έχει τοποθετηθεί στον πίνακα **A** ως αποτέλεσμα της συσχέτισης που υφίσταται ανάμεσα στους πίνακες **A** και **B** σύμφωνα με το μοντέλο οντοτήτων – συσχετίσεων.

Παράδειγμα ERD



Παράδειγμα ξένων κλειδιών

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fine Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

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

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
333445555	Alice	F	05-Apr-76	Daughter
333445555	Theodore	M	25-Oct-73	Son
333445555	Joy	F	03-May-48	Spouse
987654321	Abner	M	29-Feb-32	Spouse
123456789	Michael	M	01-Jan-78	Son
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse

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	10	15.0
888665555	20	NULL

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
2	Stafford
5	Bellaire
5	Sugarland
5	Houston

EMPLOYEE.DNO = DEPARTMENT.DNUMBER

Παράδειγμα ξένων κλειδιών

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE	LOC
Research	5	333445555	22-May-78	
Administration	4	987654321	01-Jan-85	
Headquarters	1	888665555	19-Jun-71	

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	10	15.0
888665555	20	NULL

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
2	Stafford
5	Bellaire
5	Sugarland
5	Houston

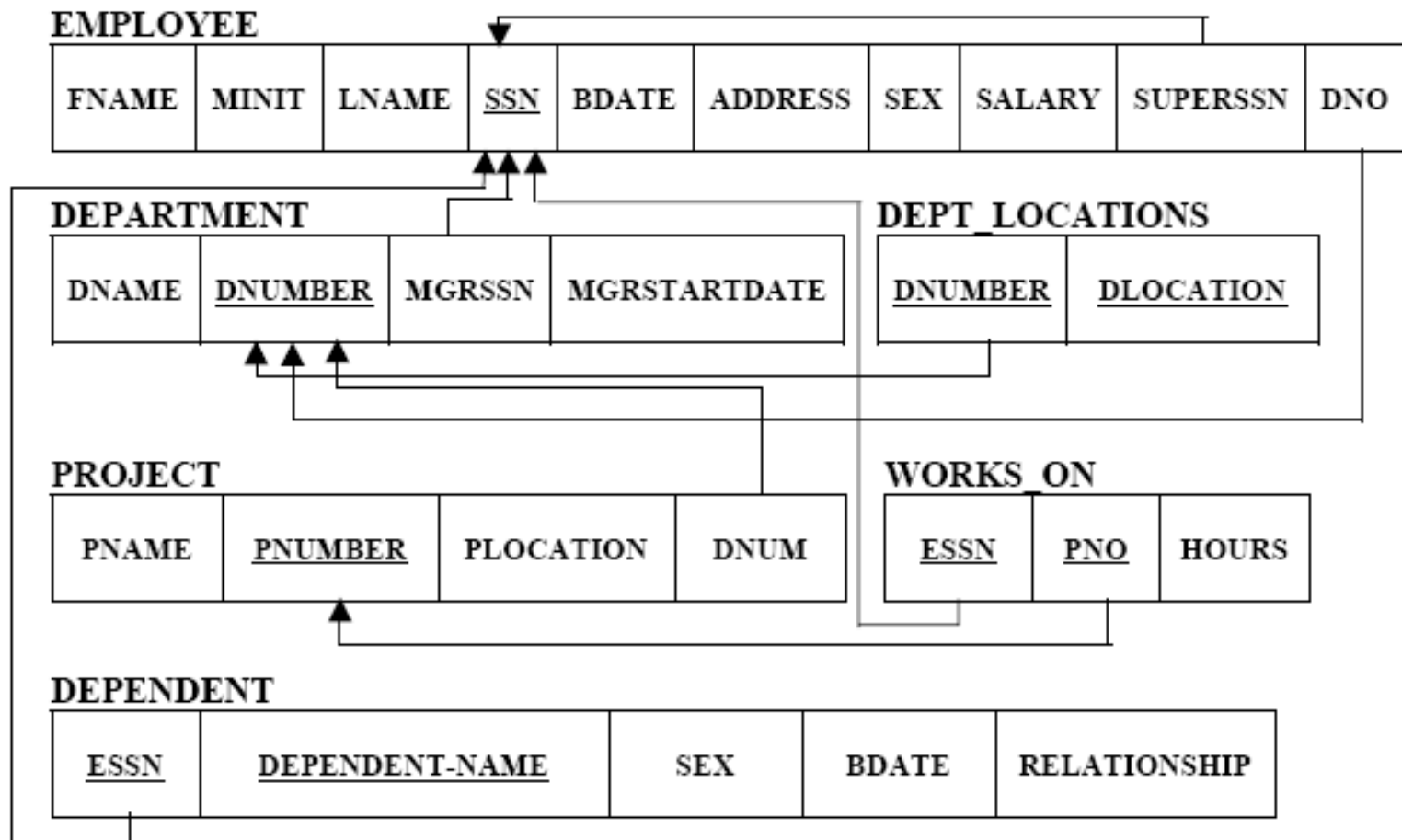
PROJECT

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

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
333445555	Alice	F	05-Apr-76	Daughter
333445555	Theodore	M	25-Oct-73	Son
333445555	Joy	F	03-May-48	Spouse
987654321	Abner	M	29-Feb-32	Spouse
123456789	Michael	M	01-Jan-78	Son
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse

Παράδειγμα κλειδιών



Αναδρομική συσχέτιση

- **Αναδρομική (recursive) συσχέτιση**: ορίζεται ανάμεσα σε δύο αντίγραφα του ίδιου τύπου οντότητας
- Το ξένο κλειδί του πίνακα, αναφέρεται σε κάποια εγγραφή του ίδιου πίνακα.
- **Παράδειγμα** : το πεδίο **SUPERSSN** του υπαλλήλου **John Smith** έχει τιμή **333445555**. Αυτό σημαίνει πως θα πρέπει στον ίδιο πίνακα να υπάρχει κάποιος εργαζόμενος με **SSN = 333445555** – αυτός ο εργαζόμενος πράγματι υπάρχει και είναι ο **Franklin Wong**, και επομένως στην περίπτωση αυτή, ο τρίτος κανόνας ακεραιότητας ικανοποιείται.

Κανόνες ακεραιότητας

- Οι **κανόνες ακεραιότητας (integrity constraints)** ενός σχεσιακού σχήματος μιας βάσης δεδομένων, διασφαλίζουν τη συνέπεια των δεδομένων της βάσης, και ισχύουν για κάθε στιγμιότυπό της.
- **1^{ος} κανόνας (key constraint)**: απόδοση μοναδικών τιμών σε όλα τα πεδία του πίνακα που έχουν χαρακτηριστεί ως υποψήφια κλειδιά (candidate keys).
- **2^{ος} κανόνας (entity integrity constraint)**: απαγορεύει την απόδοση της τιμής **NULL**, στο πρωτεύον κλειδί της κάθε εγγραφής του πίνακα.
- **3^{ος} κανόνας (referential integrity constraint)**: εάν η τιμή κάποιου πεδίου μιας από τις εγγραφές ενός πίνακα αναφέρεται σε εγγραφή κάποιου άλλου πίνακα, αυτή θα πρέπει να είναι μια υπάρχουσα εγγραφή και να μην οδηγούμαστε σε **ασυνεπή δεδομένα (inconsistent data)**.

Παραδείγματα INSERT

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

Insert <'Cecilia','F','Kolonsky','677678989','05-Apr-50','6357 Windy Lane, Katy, TX','F',28000,NULL,4>
into EMPLOYEE

Insert <'Alicia','J','Zelaya','999887777','05-Apr-50','6357 Windy Lane, Katy, TX','F',28000,'987654321',4>
into EMPLOYEE

Insert <'Cecilia','F','Kolonsky',NULL,'05-Apr-50','6357 Windy Lane, Katy, TX','F',28000,NULL,4> into
EMPLOYEE

Insert <'Cecilia','F','Kolonsky','677678989','05-Apr-50','6357 Windy Lane, Katy,
TX','F',28000,'987654321',7> into EMPLOYEE

Παραδείγματα DELETE

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

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	10	15.0
888665555	20	NULL

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
2	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

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
333445555	Alice	F	05-Apr-76	Daughter
333445555	Theodore	M	25-Oct-73	Son
333445555	Joy	F	03-May-48	Spouse
987654321	Abner	M	29-Feb-32	Spouse
123456789	Michael	M	01-Jan-78	Son
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse

- DELETE the WORKS_ON tuple with SSN='999887777' AND PNO=10
- DELETE the EMPLOYEE tuple with SSN='999887777'
- DELETE the EMPLOYEE tuple with SSN='333445555'

Παραδείγματα DELETE

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	658 Voss, Houston, TX	M	48888	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Joseph	C	Wallace	987654321	29-Jun-24	281 Berry, Bellaire, TX	F	42000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

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	10	15.0
888665555	20	NULL

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
2	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

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
333445555	Alice	F	05-Apr-76	Daughter
333445555	Theodore	M	25-Oct-73	Son
333445555	Joy	F	03-May-48	Spouse
987654321	Abner	M	29-Feb-32	Spouse
123456789	Michael	M	01-Jan-78	Son
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse

- DELETE the WORKS_ON tuple with SSN='999887777' AND PNO=10
- **DELETE the EMPLOYEE tuple with SSN='999887777'**
- DELETE the EMPLOYEE tuple with SSN='333445555'

Παραδείγματα DELETE

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Redman, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelayer	999887777	19-Jul-58	3521 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

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	10	15.0
888665555	20	NULL

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
2	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

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
333445555	Alice	F	05-Apr-76	Daughter
333445555	Theodore	M	25-Oct-73	Son
333445555	Joy	F	03-May-48	Spouse
987654321	Abner	M	29-Feb-32	Spouse
123456789	Michael	M	01-Jan-78	Son
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse

- DELETE the WORKS_ON tuple with SSN='999887777' AND PNO=10
- DELETE the EMPLOYEE tuple with SSN='999887777'
- DELETE the EMPLOYEE tuple with SSN='333445555'

Παραδείγματα DELETE

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-78	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-46	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-38	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-58	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-68	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-88
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-81

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	10	15.0
888665555	20	NULL

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
2	Stafford
5	Bellaire
5	Sugarland
5	Houston

PROJECT

PNAME	PNO	DNO	M
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
333445555	Alice	F	05-Apr-76	Daughter
333445555	Theodore	M	25-Oct-73	Son
333445555	Joy	F	03-May-48	Spouse
987654321	Abner	M	29-Feb-32	Spouse
123456789	Michael	M	01-Jan-78	Son
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse

- DELETE the WORKS_ON tuple with SSN='999887777' AND PNO=10
- DELETE the EMPLOYEE tuple with SSN='999887777'
- DELETE the EMPLOYEE tuple with SSN='333445555'

Παράβαση 3^{ου} κανόνα από DELETE

- Στις περιπτώσεις αυτές έχουμε γενικά τρεις επιλογές:
 - να απαγορεύσουμε τη διαγραφή αυτής της πλειάδας
 - να διαγράψουμε εκτός από την εν λόγω πλειάδα και όλες τις πλειάδες των άλλων πινάκων που αναφέρονται σε αυτή (cascade delete operation)
 - να μη διαγράψουμε τις πλειάδες των άλλων πινάκων αλλά να τοποθετήσουμε **NULL** στα πεδία που αναφέρονται στην πλειάδα που επιχειρούμε να διαγράψουμε (referencing attributes).
- Για να είναι δυνατή η πραγματοποίηση της 3^{ης} επιλογής δεν θα πρέπει να πεδία που θα λάβουν την τιμή **NULL** να αποτελούν τμήμα του πρωτεύοντος κλειδιού διότι θα παραβιαστεί ο πρώτος κανόνας ακεραιότητας (key constraint).

Απεικόνιση

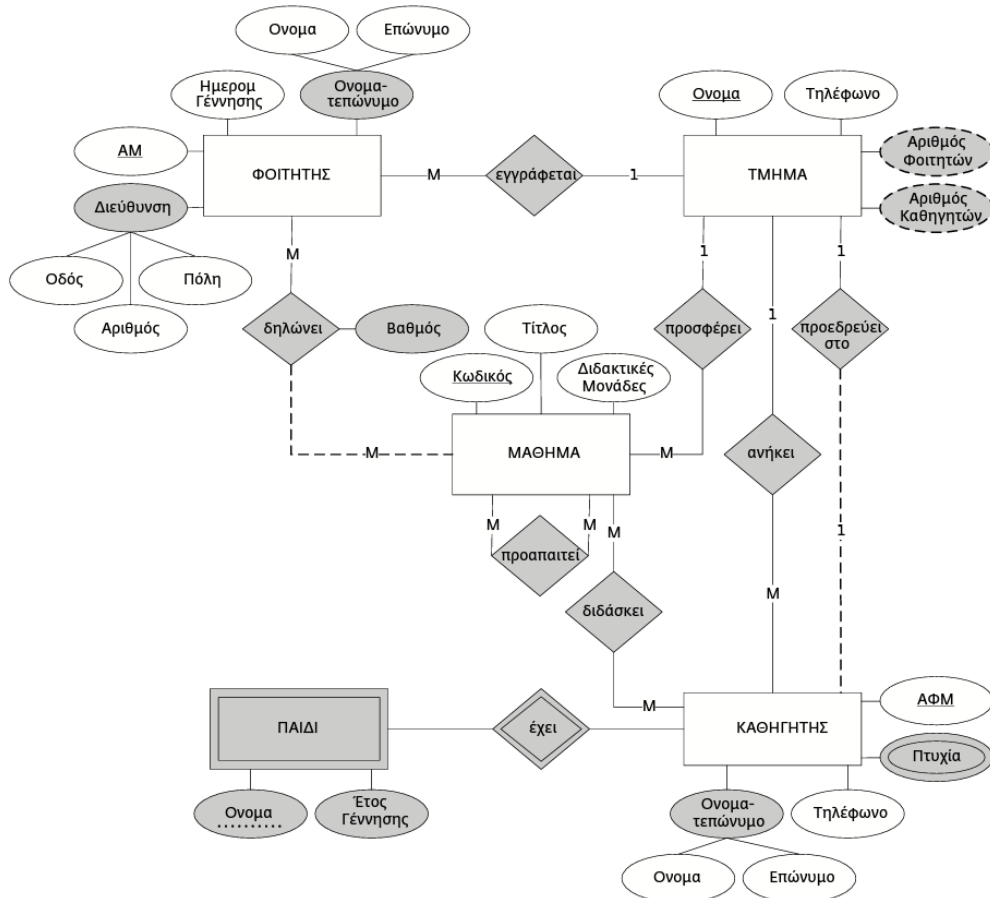
- **Απεικόνιση (mapping)** : Η διαδικασία δημιουργίας του σχεσιακού μοντέλου μιας βάσης δεδομένων από το μοντέλο οντοτήτων συσχετίσεων (ERD)
- Η απεικόνιση του ενός μοντέλου στο άλλο πραγματοποιείται δια της εφαρμογής μιας διαδικασίας **έξι βημάτων**

Βήμα 1: Ισχυρές Οντότητες

- Δημιουργείται μια σχέση του σχεσιακού μοντέλου για κάθε ισχυρή οντότητα του ERD
- Τα πεδία της σχέσης είναι τα απλά χαρακτηριστικά της οντότητας.
- Το πρωτεύον κλειδί της σχέσης είναι το πρωτεύον χαρακτηριστικό της οντότητας. Το πρωτεύον κλειδί μπορεί να είναι σύνθετο κλειδί, και να αποτελείται από περισσότερα από ένα πεδία.
- Τα χαρακτηριστικά πολλαπλών τιμών (multivalued attributes) δεν χρησιμοποιούνται ακόμη.

Παράδειγμα 1ου βήματος:

Δημιουργία σχέσεων (πινάκων) για τις ισχυρές οντότητες:



ΦΟΙΤΗΤΗΣ

ΑΜ	Όνομα	Επώνυμο	Οδός	Αριθμός	Πόλη	ΗμΓέν
----	-------	---------	------	---------	------	-------

ΜΑΘΗΜΑ

Κωδικός	Τίτλος	ΔιδΜονάδες
---------	--------	------------

ΤΜΗΜΑ

Όνομα	Τηλέφωνο
-------	----------

ΚΑΘΗΓΗΤΗΣ

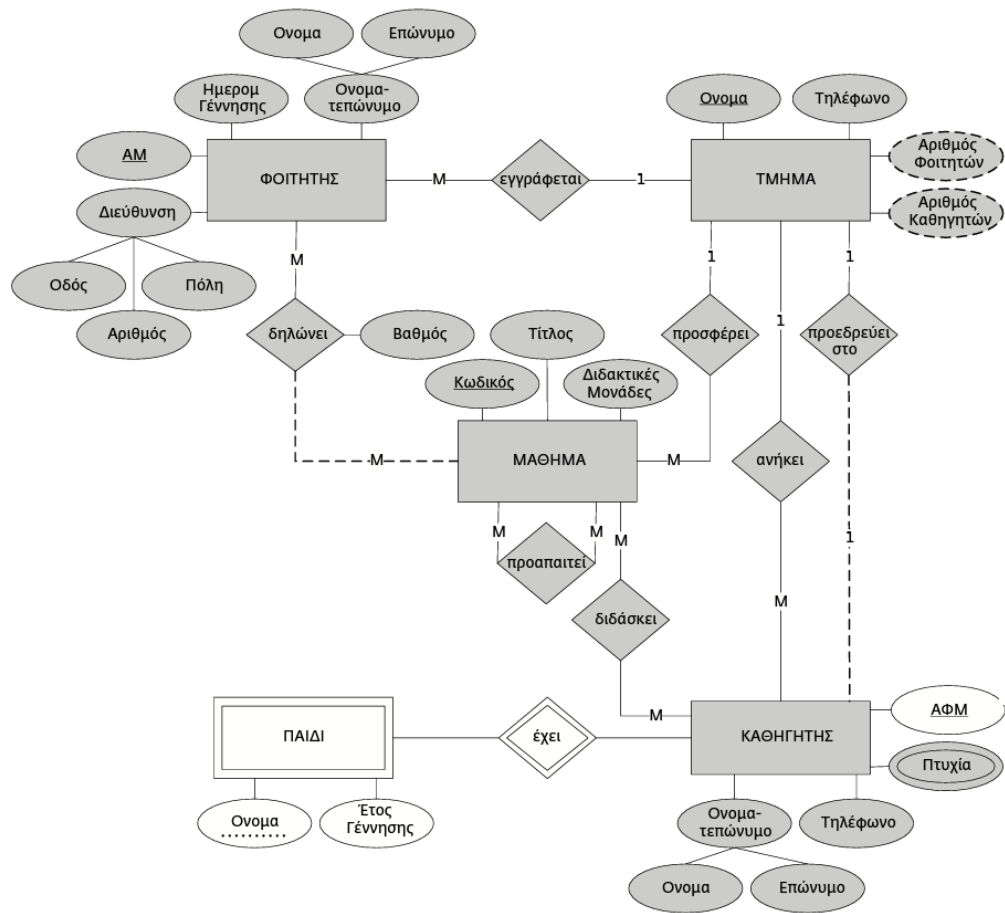
ΑΦΜ	Όνομα	Επώνυμο	Τηλέφωνο
-----	-------	---------	----------

Παρατήρηση:

Οι πίνακες είναι ακόμη ασύνδετοι μεταξύ τους καθώς δεν υφίστανται συσχετίσεις.

Βήμα 2: Ασθενείς οντότητες

- Δημιουργείται μια σχέση του σχεσιακού μοντέλου για κάθε ασθενή οντότητα του ERD
- Τα πεδία της σχέσης είναι :
 - όλα τα απλά χαρακτηριστικά της ασθενούς οντότητας και
 - το πρωτεύον χαρακτηριστικό της προσδιορίζουσας οντότητας.
- Το πρωτεύον κλειδί της σχέσης είναι ο συνδυασμός του μερικού πρωτεύοντος χαρακτηριστικού της ασθενούς οντότητας και του πρωτεύοντος χαρακτηριστικού της προσδιορίζουσας οντότητας.
- Το πρωτεύον χαρακτηριστικό της προσδιορίζουσας οντότητας είναι ταυτόχρονα ξένο κλειδί και αναφέρεται στο πρωτεύον κλειδί της σχέσης που αντιστοιχεί στην προσδιορίζουσα οντότητα.



Παράδειγμα 2ου βήματος:

Δημιουργία σχέσεων (πίνακων) για τις ασθενείς οντότητες:

ΦΟΙΤΗΤΗΣ

ΑΜ	Όνομα	Επώνυμο	Οδός	Αριθμός	Πόλη	ΗμΓέν
----	-------	---------	------	---------	------	-------

ΜΑΘΗΜΑ

Κωδικός	Τίτλος	ΔιδΜονάδες
---------	--------	------------

ΤΜΗΜΑ

Όνομα	Τηλέφωνο
-------	----------

ΚΑΘΗΓΗΤΗΣ

ΑΦΜ	Όνομα	Επώνυμο	Τηλέφωνο
-----	-------	---------	----------

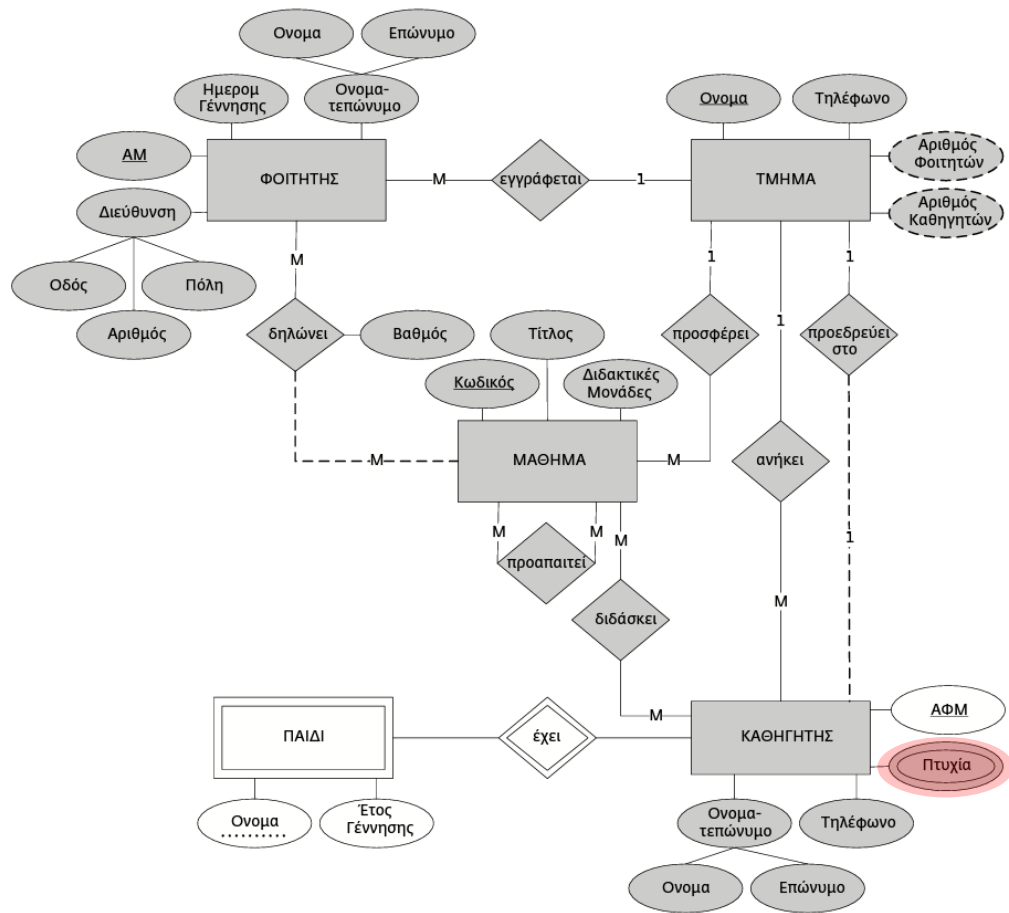
ΠΑΙΔΙ

ΑΦΜ	Όνομα	ΈτοςΓέννησης
-----	-------	--------------

ξ.κ

Βήμα 3: Χαρακτηριστικά πολλαπλών τιμών

- Δημιουργείται μια σχέση του σχεσιακού μοντέλου για κάθε χαρακτηριστικό πολλαπλών τιμών του ERD
- Τα πεδία της σχέσης είναι
 - το χαρακτηριστικό πολλαπλών τιμών και
 - το πρωτεύον χαρακτηριστικό της οντότητας στην οποία ανήκει το χαρακτηριστικό πολλαπλών τιμών.
- Το πρωτεύον κλειδί της σχέσης απαρτίζεται από όλα τα πεδία της σχέσης που δημιουργήθηκε στο βήμα αυτό.
- Το πεδίο της σχέσης που αντιστοιχεί στο πρωτεύον χαρακτηριστικό της οντότητας που έχει το χαρακτηριστικό πολλαπλών τιμών ορίζεται ξένο κλειδί.



Παράδειγμα 3ου βήματος:

Δημιουργία νέας σχέσης (πίνακα)
 Για κάθε χαρακτηριστικό πολλαπλών τιμών :

ΦΟΙΤΗΤΗΣ

ΑΜ	Όνομα	Επώνυμο	Οδός	Αριθμός	Πόλη	ΗμΓέν
----	-------	---------	------	---------	------	-------

ΜΑΘΗΜΑ

Κωδικός	Τίτλος	ΔιδΜονάδες
---------	--------	------------

ΤΜΗΜΑ

Όνομα	Τηλέφωνο
-------	----------

ΚΑΘΗΓΗΤΗΣ

ΑΦΜ	Όνομα	Επώνυμο	Τηλέφωνο	<table border="1"> <tr> <th colspan="2">ΠΤΥΧΙΑ ΚΑΘΗΓΗΤΗ</th> </tr> <tr> <td>ΑΦΜ</td> <td>Πτυχίο</td> </tr> </table>	ΠΤΥΧΙΑ ΚΑΘΗΓΗΤΗ		ΑΦΜ	Πτυχίο
ΠΤΥΧΙΑ ΚΑΘΗΓΗΤΗ								
ΑΦΜ	Πτυχίο							

Ξ.κ

ΠΑΙΔΙ

ΑΦΜ	Όνομα	ΈτοςΓέννησης
-----	-------	--------------

Ξ.κ

Βήμα 4: 1:1 συσχετίσεις

- Προστίθεται ένα ξένο κλειδί για κάθε 1:1 συσχέτιση.
- Ως ξένο κλειδί επιλέγεται το πρωτεύον χαρακτηριστικό μίας από τις οντότητες που συμμετέχει στην 1:1 συσχέτιση. Αυτό προστίθεται στη σχέση που αντιστοιχεί στην άλλη οντότητα που συμμετέχει στην 1:1 συσχέτιση. Στην ίδια σχέση προστίθεται ως πεδίο και κάθε χαρακτηριστικό της συσχέτισης.

Παράδειγμα 4ου βήματος:

Προσθήκη ξένων κλειδιών για τις συσχετίσεις 1:1 :

ΦΟΙΤΗΤΗΣ

ΑΜ	Όνομα	Επώνυμο	Οδός	Αριθμός	Πόλη	ΗμΓέν
----	-------	---------	------	---------	------	-------

ΜΑΘΗΜΑ

Κωδικός	Τίτλος	ΔιδΜονάδες
---------	--------	------------

ΤΜΗΜΑ

Όνομα	Τηλέφωνο	ΑΦΜΠροέδρου
ξ.κ		ξ.κ

ΚΑΘΗΓΗΤΗΣ

ΑΦΜ	Όνομα	Επώνυμο	Τηλέφωνο	ΑΦΜ	Πτυχία
ξ.κ					

ΠΤΥΧΙΑ ΚΑΘΗΓΗΤΗ

ΑΦΜ	Πτυχίο
-----	--------

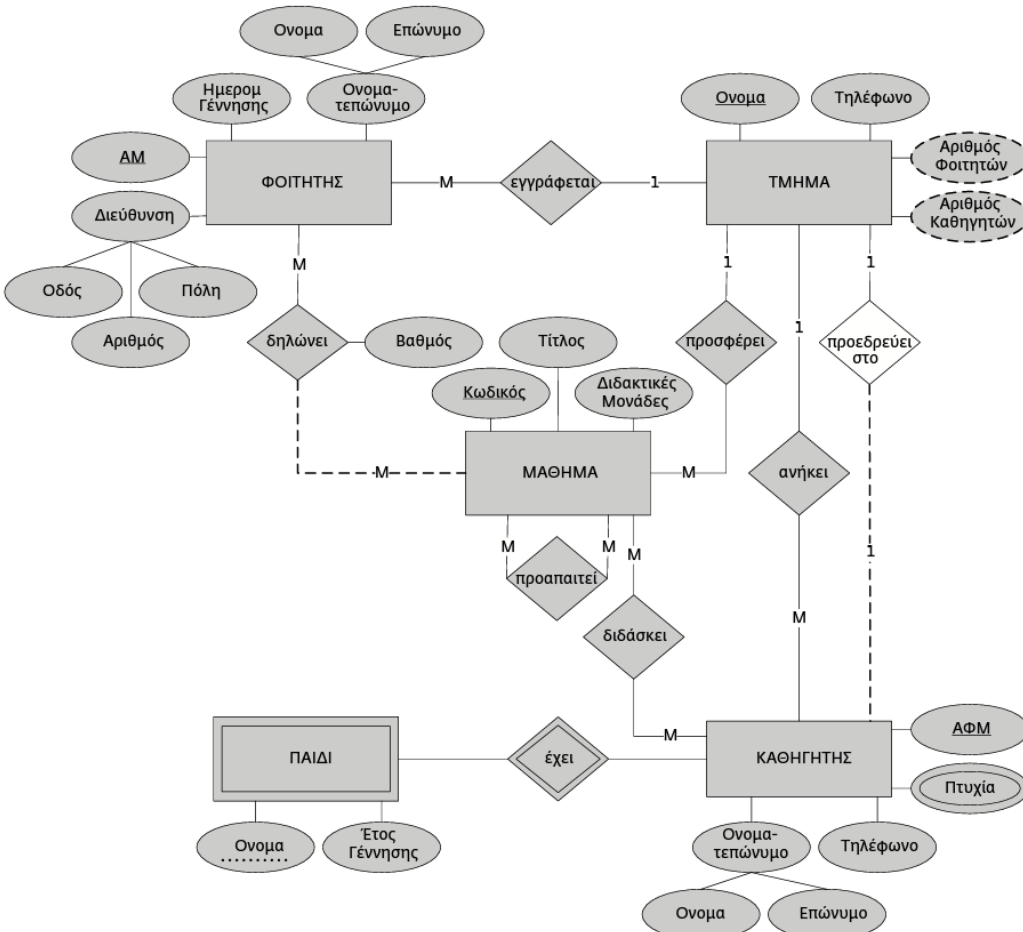
ΠΑΙΔΙ

ΑΦΜ	Όνομα	ΈτοςΓέννησης
ξ.κ		

Εναλλακτικά :

ΚΑΘΗΓΗΤΗΣ

ΑΦΜ	Όνομα	Επώνυμο	Τηλέφ	ΌνομαΤμήματοςΣτοΟποίοΠροεδρεύει
-----	-------	---------	-------	---------------------------------



Παρατηρήσεις

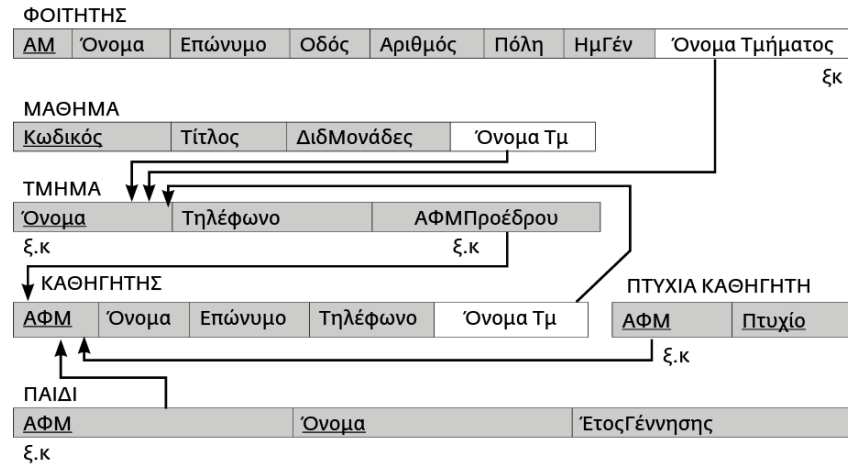
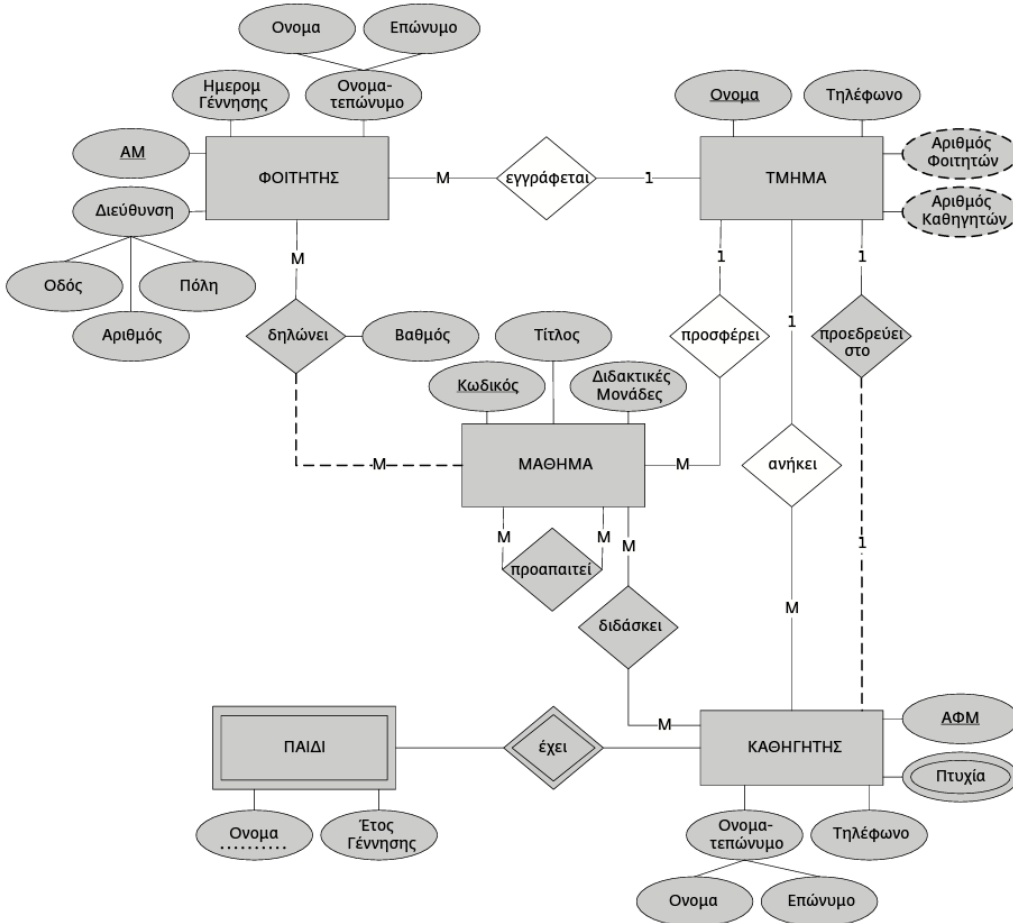
- Στη συσχέτιση «προεδρεύει στο» συμμετέχουν δύο οντότητες: το ΤΜΗΜΑ και ο ΚΑΘΗΓΗΤΗΣ.
- Έχουμε λοιπόν δύο εξίσου αποδεκτές εναλλακτικές λύσεις:
 - να προσθέσουμε ως ξένο κλειδί στον ΚΑΘΗΓΗΤΗ το πρωτεύον κλειδί από το ΤΜΗΜΑ, ή
 - να προσθέσουμε ως ξένο κλειδί στο ΤΜΗΜΑ το πρωτεύον κλειδί του ΚΑΘΗΓΗΤΗ (που είναι το ΑΦΜ).
- Παρατηρούμε ότι η σχέση «προεδρεύει στο» είναι μερικής συμμετοχής από το μέρος του ΚΑΘΗΓΗΤΗ, ενώ είναι ολικής συμμετοχής από τη μεριά της οντότητας ΤΜΗΜΑ. Αυτό σημαίνει ότι ορισμένοι (όχι όλοι) οι καθηγητές είναι πρόεδροι στο Τμήμα. Άρα, όσοι καθηγητές δεν είναι πρόεδροι θα έχουν στο χαρακτηριστικό *ΌνομαΤμήματοςΣτοΟποίοΠροεδρεύει* την τιμή **NULL** και επομένως η σχέση ΚΑΘΗΓΗΤΗ θα έχει πολλές πλειάδες οι οποίες θα έχουν την τιμή **NULL** στο πεδίο *ΌνομαΤμήματοςΣτοΟποίοΠροεδρεύει*. Επομένως η δεύτερη εναλλακτική (δηλαδή η προσθήκη του πρωτεύοντος χαρακτηριστικού του ΚΑΘΗΓΗΤΗ στο ΤΜΗΜΑ) είναι η καλύτερη.

Βήμα 5: 1:N συσχετίσεις

- Προστίθεται ένα ξένο κλειδί για κάθε 1:N συσχέτιση.
- Ως ξένο κλειδί επιλέγεται το πρωτεύον χαρακτηριστικό της οντότητας που συμμετέχει στη συσχέτιση από το μέρος του 1. Αυτό προστίθεται ως ξένο κλειδί στη σχέση που αντιστοιχεί στην οντότητα που συμμετέχει στη συσχέτιση από το μέρος του N. Στην ίδια σχέση προστίθεται και κάθε χαρακτηριστικό της συσχέτισης.

Παράδειγμα 5ου βήματος:

Προσθήκη ξένων κλειδιών για τις συσχετίσεις 1:N :

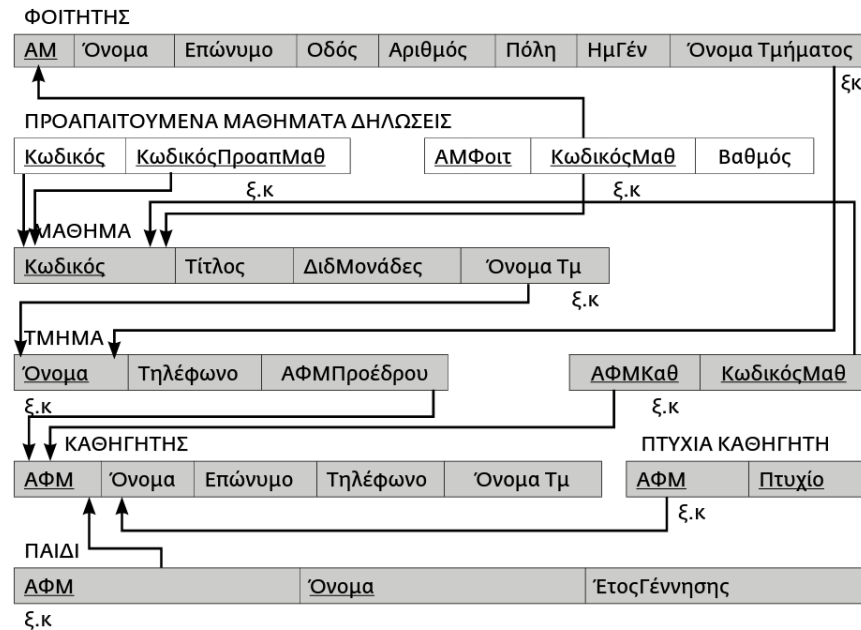
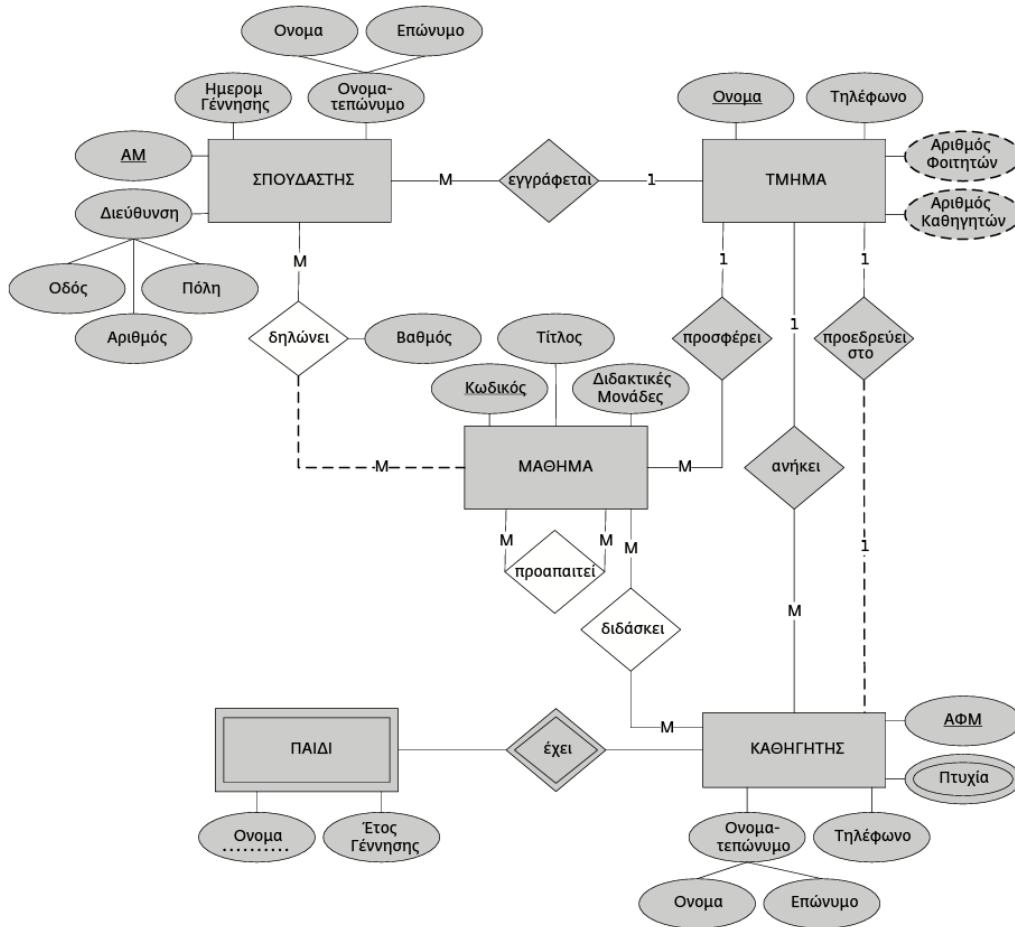


Βήμα 6: M:N συσχετίσεις

- Προστίθεται μία σχέση για κάθε M:N συσχέτιση.
- Η νέα σχέση περιλαμβάνει ως πεδία τα πρωτεύοντα χαρακτηριστικά των οντοτήτων που συμμετέχουν στη M:N συσχέτιση. Αυτά αποτελούν και το πρωτεύον κλειδί της σχέσης. Ταυτόχρονα κάθε ένα από αυτά ξεχωριστά είναι ξένο κλειδί.
- Στη σχέση προστίθεται ως πεδίο και κάθε χαρακτηριστικό της συσχέτισης.

Παράδειγμα βου βήματος:

Προσθήκη νέων σχέσεων για τις συσχετίσεις M:N :



Στιγμιότυπο

ΦΟΙΤΗΤΗΣ

ΑΜ	Όνομα	Επώνυμο	Οδός	Αριθμός	Πόλη	ΗμΓέν	ΌνομαΤμήματος
1234	Ανέστης	Δήμου	Ωκεανίας	53	Ιωάννινα	10/4/80	Πληροφορικής
1235	Βασίλης	Γεωργίου	Ποταμιάς	56	Πάτρα	8/5/81	Διοίκησης Επιχ
1236	Γιώργος	Αγγέλου	Αθηνάς	4	Αθήνα	14/2/81	Πληροφορικής
1237	Δήμητρα	Βάσου	Ιονίου	12	Σέρρες	1/4/80	Διοίκησης Επιχ
1238	Ελένη	Έλις	Ίωνος	45	Καβάλα	29/6/80	Διοίκησης Επιχ

ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ

Κωδικός	ΚωδικόςΠροαπΜαθ
ΠΛ002	ΠΛ001
ΠΛ004	ΠΛ001
ΠΛ004	ΠΛ003
ΔΕ005	ΔΕ001
ΔΕ005	ΔΕ004
ΔΕ004	ΔΕ001
ΔΕ004	ΔΕ003

ΔΗΛΩΣΕΙΣ

ΑΜΣΦοιτ	ΚωδικόςΜαθ	Βαθμός
1234	ΠΛ001	Null
1234	ΠΛ003	8
1234	ΠΛ004	9
1235	ΔΕ001	10
1235	ΔΕ003	8
1235	ΔΕ004	9
1235	ΔΕ005	10
1236	ΔΕ001	8
1236	ΔΕ002	7

ΤΜΗΜΑ

Όνομα	Τηλέφωνο	ΑΦΜΠροέδρου
Πληροφορικής	2321054321	012345678
Διοίκησης Επιχ	2321065432	056781234

ΚΑΘΗΓΗΤΗΣ

ΑΦΜ	Όνομα	Επώνυμο	Τηλέφωνο	ΌνομαΤμήματος
012345678	Ανδρέας	Ζάχαρης	2321012345	Πληροφορικής
023456789	Γεωργία	Γλύκα	2321023456	Πληροφορικής
034567891	Δήμητρα	Πίκρα	2321034567	Διοίκησης Επιχ
045678912	Ελένη	Ξύδη	2321045678	Διοίκησης Επιχ
056789123	Ζαχαρίας	Μούστος	2321056789	Διοίκησης Επιχ

ΔΙΔΑΣΚΑΛΙΑ

ΑΦΜΚαθ	ΚωδικόςΜαθ
012345678	ΠΛ001
012345678	ΠΛ002
023456789	ΠΛ003
023456789	ΠΛ004
034567891	ΔΕ001
034567891	ΔΕ002
045678912	ΔΕ003
045678912	ΔΕ004
056789123	ΔΕ005

ΠΤΥΧΙΑ ΚΑΘΗΓΗΤΗ

ΑΦΜ	Πτυχίο
012345678	Μαθηματικού
012345678	Μεταπτυχιακό στην Πληροφορική
012345678	Διδακτορικό στην Παράλληλη Επεξεργασία
023456789	Φυσικού
023456789	Μεταπτυχιακό στην Πληροφορική
023456789	Διδακτορικό στις Βάσεις Δεδομένων
034567891	Οικονομικών Σπουδών
034567891	Μεταπτυχιακό στη Διοίκηση Επιχειρήσεων
034567891	Διδακτορικό στη Συμπεριφορά Καταναλωτή
045678912	Μαθηματικού
045678912	Μεταπτυχιακό στην Επιχειρησιακή Έρευνα
045678912	Διδακτορικό στον Έλεγχο Αποθεμάτων
056789123	Μηχανολόγου Μηχανικού
056789123	Διδακτορικό στην Οργάνωση Παραγωγής

ΜΑΘΗΜΑ

Κωδικός	Τίτλος	ΔιδΜονάδες	ΌνομαΤμ
ΠΛ001	Εισαγωγή στην Πληροφορική	5	Πληροφορικής
ΠΛ002	Βάσεις Δεδομένων	3	Πληροφορικής
ΠΛ003	Προγραμματισμός με Java	3	Πληροφορικής
ΠΛ004	Επεξεργασία Εικόνας	5	Πληροφορικής
ΔΕ001	Εισαγωγή στη Διοίκηση	5	Διοίκησης Επιχ
ΔΕ002	Διαφήμιση και Πωλήσεις	4	Διοίκησης Επιχ
ΔΕ003	Πιθανότητες και Στατιστική	3	Διοίκησης Επιχ
ΔΕ004	Επιχειρησιακή Έρευνα	5	Διοίκησης Επιχ
ΔΕ005	Οργάνωση Παραγωγής	3	Διοίκησης Επιχ

ΠΑΙΔΙ

ΑΦΜ	Όνομα	ΈτοςΓέννησης
012345678	Νικόλαος	1996
034567891	Μαρία	1998
034567891	Ιωάννης	2000

Κανονικοποίηση

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Στόχος και αποτελέσματα

Στόχος του κεφαλαίου αυτού είναι να παρουσιάσει αναλυτικά τη διαδικασία κανονικοποίησης.

Προσδοκώμενα αποτελέσματα: Όταν ολοκληρώσετε τη μελέτη αυτού του κεφαλαίου, θα έχετε κατανοήσει:

- Τους στόχους και τα αποτελέσματα της διαδικασίας κανονικοποίησης
- Τον τρόπο υλοποίησης της διαδικασίας κανονικοποίησης των Σχισιακών Βάσεων Δεδομένων
- Τα χαρακτηριστικά κάθε κανονικής μορφής

Βιβλιογραφική παραπομπή :

- Κεχρής, Ε., “Σχισιακές βάσεις δεδομένων – Θεωρία & εργαστηριακές ασκήσεις”, Κεφ. 8
- Μάργαρης, Α.Ι., “Εισαγωγή στις Βάσεις Δεδομένων” – Κεφ. 4

Από το ΣΜ στους πίνακες

- Η βασική αρχή σχεδίασης του σχεσιακού μοντέλου είναι:
 - η **ομαδοποίηση των πεδίων** των τύπων σχέσεων*, και
 - η **δημιουργία πινάκων** για τους τύπους σχέσεων* από το μοντέλο οντοτήτων συσχετίσεων (ERD)
- Ερωτήματα:
 - Με ποιο κριτήριο γίνεται η επιλογή των πεδίων που θα ομαδοποιηθούν και θα σχηματίσουν ένα πίνακα ;
 - Η επιλογή γίνεται με τον καλύτερο δυνατό τρόπο ;
 - Υπάρχει κάποιος καλύτερος συνδυασμός των πεδίων που θα μπορούσε να οδηγήσει σε μια πιο ευέλικτη και αξιόπιστη βάση δεδομένων ;

**σχέση : οντότητα ή συσχέτιση*

Συναρτησιακή Εξάρτηση

Συναρτησιακή εξάρτηση (functional dependency): η συσχέτιση ανάμεσα στα πεδία ενός πίνακα.

- Ένα **πεδίο B** κάποιου πίνακα θεωρείται συναρτησιακώς εξαρτημένο από ένα **πεδίο A** του ίδιου πίνακα, όταν σε **κάθε χρονική στιγμή**, η τιμή του **A** καθορίζει μονοσήμαντα την τιμή του **B**.
- Για αυτή τη συναρτησιακή εξάρτηση, χρησιμοποιούμε το συμβολισμό **A→B**, ενώ ισοδύναμα μπορούμε να γράψουμε και **B=B (A)**

Επισημάνσεις

- Τα πεδία **A** και **B** μπορεί να είναι είτε **απλά (atomic)** είτε **σύνθετα (composite)** πεδία
- **ΟΛΑ** τα πεδία του πίνακα που δεν ανήκουν στο πρωτεύον κλειδί του, εξαρτώνται συναρτησιακά από τα πεδία του πρωτεύοντος κλειδιού
- Η ισχύς της εξάρτησης **A** \rightarrow **B** δεν προϋποθέτει και την ισχύ της εξαρτησης **B** \rightarrow **A**

Παραδείγματα συν. εξάρτησης

EMPLOYEE

FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
-------	-------	-------	------------	-------	---------	-----	--------	----------	-----

Γνωρίζοντας τον κωδικό αριθμό ενός υπαλλήλου **SSN** πρέπει να μπορούμε να προσδιορίζουμε μονοσήμαντα το επώνυμό του **LNAME** : **SSN**→**LNAME**

Αντίθετα, η συσχέτιση **LNAME→**SSN** δεν είναι έγκυρη, διότι είναι πιθανόν να υπάρχουν υπάλληλοι με το ίδιο επώνυμο, και οι οποίοι ασφαλώς, θα χαρακτηρίζονται από διαφορετικό κωδικό **SSN**.**

WORKS_ON

<u>ESSN</u>	<u>PNO</u>	HOURS
-------------	------------	-------

Ο συνδυασμός των πεδίων **{ESSN, PNO}** – σύνθετο πρωτεύον κλειδί – πρέπει να καθορίζει μονοσήμαντα την τιμή του πεδίου **HOURS**.
{ESSN, PNO}→HOURS.

Ολική & μερική εξάρτηση

- Η συναρτησιακή εξάρτηση μιας ομάδας πεδίων από κάποια άλλη, μπορεί να είναι τόσο **ολική (full dependency)** όσο και **μερική (partial dependency)**.
- Δύο ομάδες πεδίων X και Y λέμε ότι χαρακτηρίζονται από **ολική εξάρτηση**, όταν η αφαίρεση έστω και ενός από τα πεδία του X , καταργεί την εξάρτηση $X \rightarrow Y$
- Όταν η εν λόγω εξάρτηση διατηρείται ακόμη και μετά την αφαίρεση ενός πεδίου από το σύνολο πεδίων X , τότε οι δύο ομάδες πεδίων χαρακτηρίζονται από **μερική εξάρτηση**

Χαρακτηρισμοί εξαρτήσεων

- Δύο πεδία **A** και **B** χαρακτηρίζονται από **έμμεση** συναρτησιακή εξάρτηση όταν υπάρχει ένα πεδίο **C** τέτοιο ώστε $A \rightarrow C$ και $C \rightarrow B$
ή ισοδύναμα $C = C(A)$ και $B = B(C)$
- Εάν επιπλέον το πεδίο **C** δεν ανήκει στα πεδία-κλειδιά του πίνακα, τότε η εν λόγω εξάρτηση λέγεται **μεταβατική εξάρτηση (transitive dependency)**
- Μια συναρτησιακή εξάρτηση μπορεί να είναι **μόνιμη** ή **προσωρινή**, ανάλογα με το χρονικό διάστημα για το οποίο ισχύει.

Παραδείγματα χαρακτηρισμών

EMPLOYEE

FNAME	MINIT	LNAME	<u>SSN</u>	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
-------	-------	-------	------------	-------	---------	-----	--------	----------	-----

Η συσχέτιση **SSN** → **LNAME** του πίνακα **EMPLOYEE** είναι **μόνιμη** διότι ο κωδικός **SSN** ενός υπαλλήλου δεν πρόκειται να αλλάξει ποτέ.

Η συσχέτιση **DNUMBER** → **MGRSSN** του πίνακα **DEPARTMENT** είναι **προσωρινή**, διότι είναι πιθανόν σε κάποια χρονική στιγμή να λάβει χώρα αντικατάσταση του **MANAGER** κάποιου τμήματος με κάποιο άλλο υπάλληλο της εταιρείας.

DEPARTMENT

DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE
-------	----------------	--------	--------------

Πλειότιμη εξάρτηση

- **Πλειότιμη εξάρτηση (multivalued dependency):** Η συναρτησιακή εξάρτηση δύο πεδίων **A** και **B** εκ των οποίων το ένα πεδίο είναι πεδίο πολλαπλής τιμής
- Στην περίπτωση αυτή η γνώση του **A** δεν καθορίζει μονοσήμαντα την τιμή του **B**, αλλά ένα σύνολο τιμών για το **B** που είναι γνωστό και καθορισμένο εκ των προτέρων.

Παράδειγμα: Στον πίνακα *DEPARTMENT* υπάρχει το πεδίο για τις τοποθεσίες του τμήματος *DLOCATIONS*. Επειδή ένα τμήμα μπορεί να έχει γραφεία σε περισσότερες από μια τοποθεσίες, η γνώση του κωδικού *DNUMBER* για ένα τμήμα δεν προσδιορίζει μονοσήμαντα την τοποθεσία του τμήματος, αλλά συσχετίζεται με ένα μια ολόκληρη ομάδα τιμών.

DEPARTMENT

DNAME	<u>DNUMBER</u>	MGRSSN	MGRSTARTDATE	DLOCATIONS
-------	----------------	--------	--------------	------------

Κανονικοποίηση

- Η **κανονικοποίηση (normalization)** εφαρμόζεται πάνω σε ένα σχεσιακό σχήμα, και προσπαθεί να το μετασχηματίσει σε μια νέα μορφή, που να είναι απαλλαγμένη από τις διάφορες ανωμαλίες εισαγωγής, διαγραφής και τροποποίησης εγγραφών
- Είναι η διαδικασία αλλαγής των σχέσεων του σχήματος μιας βάσης δεδομένων, ώστε το τελικό σχήμα να διασφαλίζει την ακεραιότητα των δεδομένων και να απαλείφει τα περιττά δεδομένα.
- Η διαδικασία της κανονικοποίησης (normalization process) προτάθηκε από τον Codd το 1970.
- Η βασική της λειτουργία είναι να δέχεται ως είσοδο το σχεσιακό σχήμα μιας βάσης δεδομένων και να εφαρμόζει πάνω του μια σειρά από ελέγχους προκειμένου να διαπιστώσει εάν ανήκει ή όχι σε κάποια **κανονική μορφή (normal form)**.

Κανονικές μορφές

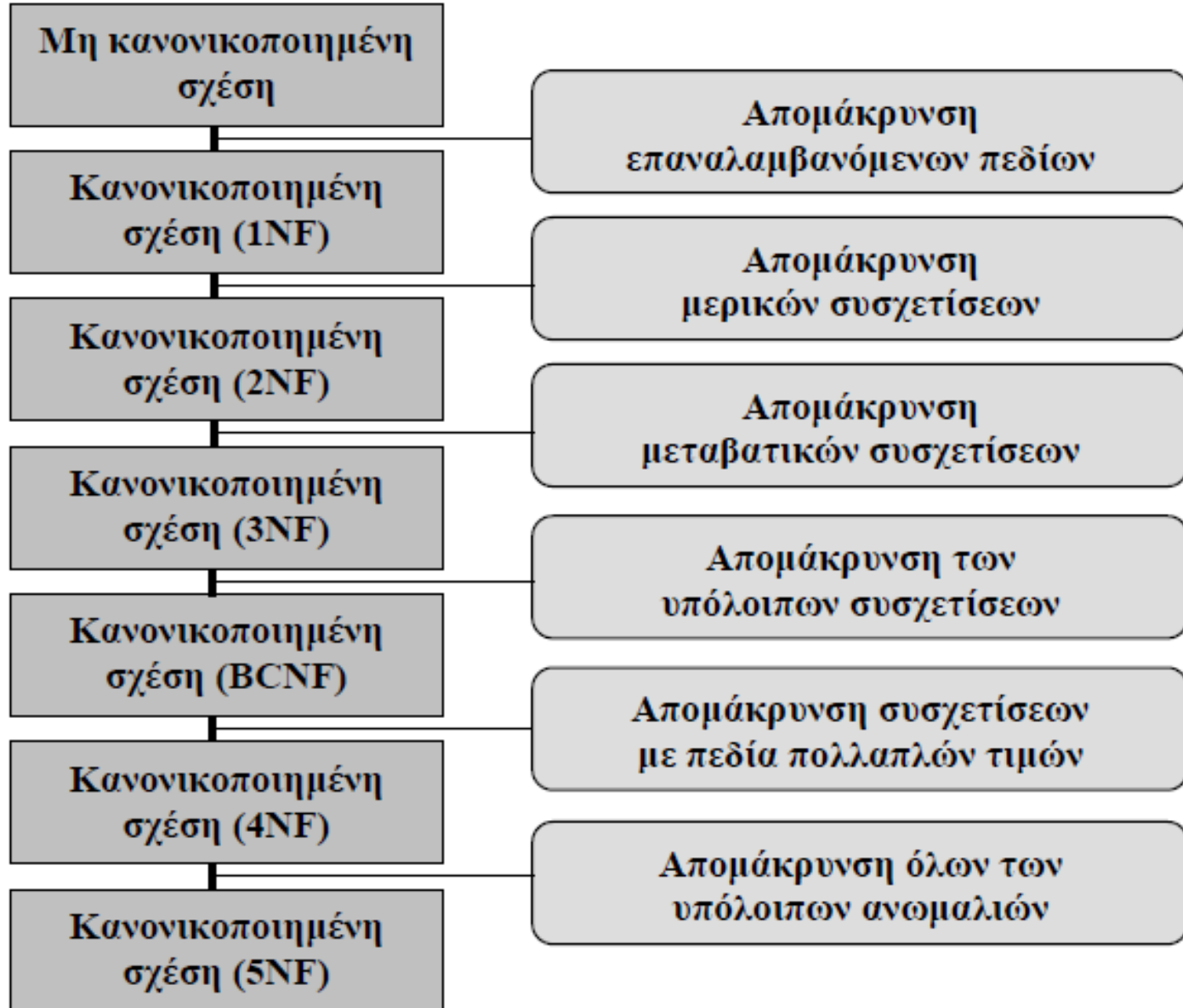
- **Κανονική μορφή (normal form):** ορίζεται ως η «κατάσταση» που βρίσκεται μια σχέση όσον αφορά τους τύπους των εξαρτήσεων που υφίστανται ανάμεσα στα πεδία της.
- Ο Codd εισήγαγε τρεις κανονικές μορφές, την *πρώτη (1NF)*, *δεύτερη (2NF)* και *τρίτη (3NF)* κανονική μορφή οι οποίες συσχετίζονται με τις συναρτησιακές εξαρτήσεις που υφίστανται ανάμεσα στα πεδία των πινάκων της βάσης.
- Αργότερα προτάθηκαν άλλες δύο κανονικές μορφές – η *τέταρτη (4NF)* και η *πέμπτη (5NF)* – οι οποίες στηρίζονται σε άλλου είδους εξαρτήσεις (**multivalued dependencies** και **join dependencies**).

Διαδικασίες μορφών I

- **Πρώτη Κανονική Μορφή (1st Normal Form, 1NF):** Να απομακρύνουμε τις επαναλαμβανόμενες ομάδες πεδίων, έτσι ώστε η τομή μιας γραμμής και μιας στήλης του πίνακα, να αντιστοιχεί πάντα σε μια απλή τιμή.
- **Δεύτερη Κανονική Μορφή (2nd Normal Form, 2NF):** για ένα πίνακα σε 1NF, να απομακρύνουμε όλες τις μερικές συναρτησιακές εξαρτήσεις (partial dependencies) που υφίστανται ανάμεσα στα πεδία του.
- **Τρίτη Κανονική Μορφή (3rd Normal Form, 3NF):** για ένα πίνακα σε 2NF, να απομακρύνουμε όλες τις μεταβατικές συναρτησιακές εξαρτήσεις (transitive dependencies) που υφίστανται ανάμεσα στα πεδία του.

Διαδικασίες μορφών II

- **Τέταρτη Κανονική Μορφή (4th Normal Form, 4NF):** για ένα πίνακα σε 3NF, να απομακρύνουμε τις πλειότιμες συναρτησιακές εξαρτήσεις (multivalued dependencies)
- **Πέμπτη Κανονική Μορφή (5th Normal Form, 5NF):** για ένα πίνακα σε 4NF, να απομακρύνουμε τις υπόλοιπες εξαρτήσεις που ενδεχομένως παραμένουν στη δομή του
- **Boyce-Codd (Boyce Codd Normal Form, BCNF):** μπορεί να θεωρηθεί μια πιο αυστηρή 3NF. Ένας πίνακας που βρίσκεται σε BCNF βρίσκεται αυτόματα και σε 3NF – **το αντίστροφο όμως δεν ισχύει.**



1NF

- Ένας πίνακας λέμε ότι βρίσκεται σε 1NF, όταν η τιμή του κάθε πεδίου σε κάθε πλειάδα, είναι **ατομική** δηλαδή δεν μπορεί να διασπαστεί σε μικρότερες μονάδες πληροφορίας
- Αποτρέπει την εμφάνιση **επαναλαμβανόμενων πεδίων** και την εμφάνιση **σύνθετων ή πολλαπλών** τιμών καθώς και **συνδυασμούς αυτών** των δύο.

1NF



- Απαιτήσεις 1NF :
 - Όλες οι πλειάδες πρέπει να προσδιορίζονται μοναδικά με ένα πρωτεύον κλειδί. Το πρωτεύον κλειδί είναι μοναδικό για την κάθε πλειάδα (εγγραφή/καταχώρηση).
 - Όλα τα πεδία, πέραν του πρωτεύοντος κλειδιού, πρέπει να εξαρτώνται από αυτό.
 - Όλα τα πεδία πρέπει να περιέχουν μια απλή τιμή.
 - Όλες οι τιμές στο ίδιο πεδίο πρέπει να είναι του ίδιου τύπου δεδομένων – *datatype*.
- Η 1NF δημιουργεί συσχετίσεις πολλαπλότητας 1:N

1^ο Παράδειγμα 1NF

AUTHOR	TITLE_01	PAGES_01	TITLE_02	PAGES_02	TITLE_03	PAGES_03	TITLE_04	PAGES_04	TITLE_05	PAGES_05
James Blish	A Case of Conscience	256	Cities of Flight	590						
Larry Niven	Footfall	608	Lucifer's Hammer	640	Ringworld	352				



Primary keys uniquely identify rows on each table

Foreign key points to master table

AUTHOR
Isaac Azimov
James Blish
Larry Niven

Author

AUTHOR	TITLE	ISBN	PAGES
Isaac Azimov	Foundation	893402095	435
Isaac Azimov	Foundation	345308999	
Isaac Azimov	Foundation	345336275	285
Isaac Azimov	Foundation	5557076654	
Isaac Azimov	Foundation	246118318	234
Isaac Azimov	Foundation	345334787	
Isaac Azimov	Foundation	5553673224	
Isaac Azimov	Foundation and Empire	553293370	320
Isaac Azimov	Foundation's Edge	553293389	480
Isaac Azimov	Prelude to Foundation	553298398	480
Isaac Azimov	Second Foundation	553293362	304
James Blish	A Case of Conscience	345438353	256
James Blish	Cities in Flight	1585670081	590
Larry Niven	Footfall	345323440	608
Larry Niven	Lucifer's Hammer	449208133	640
Larry Niven	Ringworld	345333926	352

Book

1NF

N

No restriction on the number of books per author

1

2^ο Παράδειγμα 1NF

1

DEPARTMENT

DNAME	<u>DNUMBER</u>	DMGRSSN	DLOCATIONS
Research	5	333445555	Bellaire Sugarland Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

N



DEPARTMENT

DNAME	<u>DNUMBER</u>	DMGRSSN	<u>DLOCATION</u>
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

1NF

3^ο Παράδειγμα 1NF

- Το πεδίο **Telephone Numbers** έχει μετατραπεί σε πεδίο συμβολοσειράς για να καταστεί δυνατή η αποθήκευση πολλαπλών αριθμών, ενώ ένας αριθμός τηλεφώνου είναι της μορφής : **xxx-xxx-xxxx** όπου **x** ψηφίο 0-9

Customer

Customer ID	First Name	Surname	Telephone Numbers
123	Robert	Ingram	555-861-2025
456	Jane	Wright	555-403-1659, 555-776-4100
789	Maria	Fernandez	555-808-9633

- Από τον πίνακα **Customer** προκύπτουν 2 νέοι πίνακες με πολλαπλότητα 1:N :

Customer Name

<u>Customer ID</u>	First Name	Surname
123	Robert	Ingram
456	Jane	Wright
789	Maria	Fernandez

Customer Telephone Number

<u>Customer ID</u>	<u>Telephone Number</u>
123	555-861-2025
456	555-403-1659
456	555-776-4100
789	555-808-9633

2NF



- Ένας πίνακας λέμε ότι βρίσκεται σε 2NF, όταν όλα τα πεδία που δεν ανήκουν στο πρωτεύον κλειδί του πίνακα, εξαρτώνται συναρτησιακά **μόνο** από τα πεδία του πρωτεύοντος κλειδιού, και μάλιστα, μέσω **πλήρους** συναρτησιακής εξάρτησης (**full dependency**)
- Η 2NF προκύπτει από την 1NF, εάν μετασχηματίσουμε τη δομή του πίνακα, ώστε να απομακρυνθούν όλες οι μερικές συναρτησιακές εξαρτήσεις (partial dependencies) που υφίστανται ανάμεσα στα πεδία του
- **Η αναγωγή ενός πίνακα στη 2NF έχει νόημα μόνο όταν το πρωτεύον κλειδί του είναι σύνθετο**

2NF



- Απαιτήσεις 2NF:
 - Ο πίνακας πρέπει να είναι στην 1NF.
 - Πρέπει να εξαλειφθούν τα πεδία που είναι μερικώς εξαρτημένα ή ανεξάρτητα από το πρωτεύον κλειδί.
 - Πρέπει να δημιουργηθούν νέοι πίνακες με τα με τα πεδία που μερικώς εξαρτημένα ή ανεξάρτητα από το πρωτεύον κλειδί.
- Η 2NF κάνει το αντίστοιχο της 1NF, όχι όμως στα πεδία, αλλά στις τιμές
- Η 2NF δημιουργεί συσχετίσεις πολλαπλότητας N:1

1^ο Παράδειγμα 2NF

DEPARTMENT

DNAME	<u>DNUMBER</u>	DMGRSSN	DLOCATIONS
Research	1 5	333445555	Bellaire Sugarland Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston



DEPARTMENT {DNAME, DNUMBER, DMGRSSN, DLOCATION})
DNUMBER → {DNAME, DMGRSSN}

DEPARTMENT

DNAME	<u>DNUMBER</u>	DMGRSSN	<u>DLOCATION</u>
Research	5	333445555	Bellaire
Research	N 5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

2NF

DEPARTMENT

DNAME	<u>DNUMBER</u>	DMGRSSN
Research	5	333445555
Administration	1 4	987654321
Headquarters	1	888665555

DEPT LOCATIONS

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

2^ο Παράδειγμα 2NF

EMPL_PROJ {SSN, PNUMBER, HOURS, ENAME, PNAME, PLOCATION} 1NF

$\{SSN, PNUMBER\} \rightarrow HOURS$: Ο κωδικός του εργαζόμενου και του **project** πρέπει να προσδιορίζει τις ώρες απασχόλησης

Αλλά και :

$SSN \rightarrow ENAME$: Ο κωδικός του εργαζόμενου πρέπει να προσδιορίζει μονοσήμαντα το όνομά του

$PNUMBER \rightarrow \{PNAME, PLOCATION\}$: Ο κωδικός του project προσδιορίζει μονοσήμαντα την ονομασία του και τον τόπο που εκτελείται

- Άρα υπάρχουν πεδία **μερικώς εξαρτημένα** από το πρωτεύον κλειδί, γιατί προσδιορίζονται ακόμη και χωρίς να προσδιορίζονται και τα δύο πεδία του πρωτεύοντος κλειδιού

Λύση για 2NF : *EMPL_PROJ* {*SSN, PNUMBER, HOURS*},
EMPLOYEE {*SSN, ENAME*} και
PROJECT {*PNUMBER, PNAME, PLOCATION*}

3NF

- Ένας πίνακας λέμε ότι βρίσκεται σε 3NF, όταν δεν υφίστανται **μεταβατικές εξαρτήσεις (transitive dependencies)** μεταξύ των πεδίων του.
- Αν υπάρχουν μεταβατικές εξαρτήσεις, αυτές πρέπει να εξαλειφθούν και να δημιουργηθούν νέοι πίνακες με τα με τα πεδία που δημιουργούν μεταβατικές εξαρτήσεις.
- Η 3NF επιβάλλει ο πίνακας να είναι 2NF και όλα τα πεδία που δεν είναι ξένα κλειδιά να είναι πλήρως εξαρτημένα **μόνο** από το πρωτεύον κλειδί

1^ο Παράδειγμα 3NF

Tournament_Winners {Tournament, Year, Winner, Winner_Date_of_Birth}
{Tournament, Year} → Winner
Winner → Winner_Date_of_Birth

Λύση για 3NF :

Tournament_Winners {Tournament, Year, Winner}
Player_Dates_of_Birth {Player, Date_of_Birth}

Tournament Winners

<u>Tournament</u>	<u>Year</u>	Winner	Winner Date of Birth
Indiana Invitational	1998	Al Fredrickson	21 July 1975
Cleveland Open	1999	Bob Albertson	28 September 1968
Des Moines Masters	1999	Al Fredrickson	21 July 1975
Indiana Invitational	1999	Chip Masterson	14 March 1977

Tournament Winners

<u>Tournament</u>	<u>Year</u>	Winner
Indiana Invitational	1998	Al Fredrickson
Cleveland Open	1999	Bob Albertson
Des Moines Masters	1999	Al Fredrickson
Indiana Invitational	1999	Chip Masterson

Player Dates of Birth

<u>Player</u>	<u>Date of Birth</u>
Chip Masterson	14 March 1977
Al Fredrickson	21 July 1975
Bob Albertson	28 September 1968

2^ο Παράδειγμα 3NF

EMPL_DEPT { *ENAME*, *SSN*, *BDATE*, *ADDRESS*, *DNUMBER*, *DCITY* } **2NF**

Εξαρτήσεις :

SSN → *DNUMBER* : Ο κωδικός προσδιορίζει το τμήμα που δουλεύει ο υπάλληλος

DNUMBER → *DCITY* : Το τμήμα προσδιορίζει την πόλη

SSN → *DCITY* : Η πόλη κάθε υπαλλήλου προσδιορίζεται από το τμήμα που αυτός εργάζεται (μεταβατική εξάρτηση)

Λύση για 3NF :

EMPLOYEE { *ENAME*, *SSN*, *BDATE*, *ADDRESS*, *DNUMBER* } ^{f.k.} και
DEPARTMENT { *DNUMBER*, *DNAME*, *DCITY* }

Boyce-Codd Normal Form

- **Ορίζον πεδίο (determinant):** οποιοδήποτε πεδίο (απλό ή σύνθετο) μιας σχέσης από το οποίο κάποιο άλλο είναι πλήρως εξαρτημένο
- **Μια σχέση είναι στη BCNF αν και μόνο αν κάθε ορίζον πεδίο είναι υποψήφιο κλειδί**
- Η κανονική μορφή των **Boyce και Codd** αποτελεί μια πιο αυστηρή διατύπωση της 3NF
- Χρησιμοποιείται για να απομακρύνει ανωμαλίες που ενδέχεται να προκύψουν σε περιπτώσεις που ένας πίνακας έχει περισσότερα από ένα υποψήφια κλειδιά, τα οποία μάλιστα είναι σύνθετα.
- Ένας μη-BCNF πίνακας **δεν μπορεί πάντα** να αναλυθεί σε πίνακες BCNF π.χ. όταν $\{A, B \rightarrow C, C \rightarrow B\}$, αντίθετα από τις τρεις πρώτες κανονικές μορφές

1^ο Παράδειγμα BCNF

I

- Κάθε μαθητής μπορεί να δηλώσει πολλά μαθήματα
- Κάθε μάθημα μπορεί να έχει πολλούς δασκάλους
- Κάθε μαθητής σε ένα μάθημα έχει ένα μόνο δάσκαλο
- Κάθε δάσκαλος διδάσκει ένα μάθημα μόνο
- Κάθε δάσκαλος έχει πολλούς μαθητές

CLASS{*StudentID*, *Class*, *Teacher*}

- *StudentID* → *Lesson* (multivalued dep.)
- {*StudentID*, *Lesson*} → *Teacher*
- *Lesson* → *Teacher* (multivalued dep.)
- *Teacher* → *Lesson*
- *Teacher* → *StudentID* (multivalued dep.)

StudentID	Lesson	Teacher
123	physics	Brown
123	music	Jones
456	biology	Stewart
789	physics	Lenon
999	physics	Brown

Υποψήφια κλειδιά : {*StudentID*, *Lesson*} και {*StudentID*, *Teacher*}
και υπερ-κλειδί το {*StudentID*, *Class*, *Teacher*}

1^ο Παράδειγμα BCNF

I

- Πρωτεύον κλειδί {**StudentID**, **Lesson**}: Το πεδίο **Teacher** ταυτοποιεί μονοσήμαντα το πεδίο **Lesson**, αλλά δεν αποτελεί μέρος του πρωτεύοντος κλειδιού
- Προβλήματα :
 - Αν ο μαθητής 456 αλλάξει επιλογή, θα χαθεί η πληροφορία του μαθήματος **biology** με δάσκαλο τον **Stewart**
 - Αν θέλαμε να κρατήσουμε την πληροφορία για το μάθημα και το δάσκαλο, θα έπρεπε να υπάρχει μια εγγραφή σαν {**NULL**, **biology**, **Stewart**}, πράγμα που δεν επιτρέπεται όμως, καθώς το πεδίο **StudentID** είναι μέρος του πρωτεύοντος κλειδιού.
- Προτεινόμενες λύσεις :
 - **CLASS_1** {**StudentID**, **Lesson**}, **CLASS_2** {**Teacher**, **Lesson**}
 - **CLASS_1** {**StudentID**, **Teacher**}, **CLASS_2** {**Teacher**, **Lesson**}

2^ο Παράδειγμα BCNF

I

StudentNo	StudentName	Appointment Code	Time	Advisor
1	John	1	09:00	Zorro
2	Kerr	1	09:00	Killer
3	Adam	2	10:00	Zorro
4	Robert	1	13:00	Killer
5	Zane	2	14:00	Zorro
6	Bob	3	11:00	Killer
6	Bob	4	12:00	Zorro

ΠΡΟΓΡΑΜΜΑ ΚΑΘΗΜΕΡΙΝΩΝ ΣΥΝΑΝΤΗΣΕΩΝ

- Κάθε φοιτητής μπορεί να έχει 4 είδη συναντήσεων : 1.Ενημερωση, 2.Υποβολή πρότασης, 3.Αξιολόγηση και 4.Ολοκλήρωση.
- Δεν χρειάζονται όλοι οι φοιτητές και τις 4 συναντήσεις.
- Όσων το όνομα αρχίζει από Α-Ο έρχονται πρωί, ενώ οι υπόλοιποι (Ρ-Ζ) απόγευμα.
- Η συνάντηση 1 γίνεται στις 09:00 ή 13:00, η συνάντηση 2 στις 10:00 ή 14:00, κοκ.
- Κάθε ώρα μπορεί να απασχολείται ένας από τους δύο ή και οι δύο Σύμβουλοι

2^ο Παράδειγμα BCNF

II

DB{StudentNo, StudentName, AppointmentCode, Time, Advisor}

- StudentNo → StudentName, StudentName → StudentNo
- {StudentNo, AppointmentCode} → {Time, Advisor}
- Time → AppointmentCode, AppointmentCode → Time (mv.d.)
- {StudentNo, Time} → {AppointmentCode, Advisor}
- Time → Advisor (mv. d.)

Επιλέγουμε το πρωτεύον κλειδί :

- DB{StudentNo, StudentName, AppointmentCode, Time, Advisor} 1NF

2NF :

- DB{StudentNo, AppointmentCode, Time, Advisor} 3NF
- R1{StudentNo, StudentName} 3NF, BCNF

Ελέγχουμε BCNF την DB:

- Time → AppointmentCode αλλά το Time δεν είναι υποψήφιο κλειδί

BCNF :

- DB{StudentNo, Time, Advisor} 3NF, BCNF
- R1{StudentNo, StudentName} 3NF, BCNF
- R2{Time, AppointmentCode} 3NF, BCNF

2° Παράδειγμα BCNF

III

DB{StudentNo, StudentName, AppointmentCode, Time, Advisor}

- StudentNo \rightarrow StudentName, StudentName \rightarrow StudentNo
- {StudentNo, AppointmentCode} \rightarrow {Time, Advisor}
- Time \rightarrow AppointmentCode, AppointmentCode \rightarrow Time (mv.d.)
- {StudentNo, Time} \rightarrow {AppointmentCode, Advisor}
- Time \rightarrow Advisor (mv. d.)

Επιλέγουμε το πρωτεύον κλειδί :

- DB{StudentNo, StudentName, AppointmentCode, Time, Advisor} 1NF

2NF :

- DB{StudentNo, AppointmentCode, Time, Advisor} 2NF (μερική εξάρτηση)
- R1{StudentNo, StudentName} 3NF, BCNF

2NF διόρθωση:

- DB{StudentNo, Time, Advisor} 3NF, BCNF
- R1{StudentNo, StudentName} 3NF, BCNF
- R2{Time, AppointmentCode} 3NF, BCNF

2^ο Παράδειγμα BCNF

VI

<u>StudentNo</u>	StudentName
1	John
2	Kerr
3	Adam
4	Robert
5	Zane
6	Bob

<u>StudentNo</u>	<u>Time</u>	Advisor
1	09:00	Zorro
2	09:00	Killer
3	10:00	Zorro
4	13:00	Killer
5	14:00	Zorro
6	11:00	Killer
6	12:00	Zorro

<u>Time</u>	AppointmentCode
9:00	1
10:00	2
11:00	3
12:00	4
13:00	1
14:00	2
15:00	3
16:00	4

$DB\{\underline{StudentNo}, \underline{Time}, Advisor\}$
 $R1\{\underline{StudentNo}, StudentName\}$
 $R2\{\underline{Time}, AppointmentCode\}$

4NF

Για ένα πίνακα σε 3NF/BCNF, να απομακρύνουμε τις **μη-τετριμμένες*** πλειότιμες συναρτησιακές εξαρτήσεις (multivalued dependencies)

Παράδειγμα :

- Κάθε μάθημα μπορεί να διδαχθεί από πολλούς καθηγητές
- Σε κάθε μάθημα μπορούν να χρησιμοποιηθούν περισσότερα από ένα βιβλία
- Για κάθε μάθημα χρησιμοποιούνται πάντα τα ίδια βιβλία, ανεξάρτητα από τον καθηγητή που το διδάσκει.

OFFERNG {COURSE, INSTRUCTOR, TEXTBOOK}

- Οι εξαρτήσεις είναι οι **Course** → **Instructor** και **Course** → **Textbook**, με τα πεδία **Instructor** και **Textbook**, να μπορούν να λάβουν περισσότερες από μια τιμές, σε κάθε εξάρτηση.

Παράδειγμα 4NF

- *OFFERNG* {COURSE, INSTRUCTOR, TEXTBOOK}
- *Course* → *Instructor* (*multivalued*)
- *Course* → *Textbook* (*multivalued*)

ΛΥΣΗ 4NF

- COURSE_INSTRUCTOR {COURSE, INSTRUCTOR}
- COURSE_TEXTBOOK {COURSE, TEXTBOOK}

COURSE	INSTRUCTOR	TEXTBOOK
Management	White Green Black	Drucker Peters
Finance	Gray	Weston Gilford

Μη κανονικοποιημένη μορφή του πίνακα OFFERING

Λογική & Θεωρία Συνόλων

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων
(ΕΒΔΟ100)

Τι είναι λογική;

- 1. επιστήμη που ασχολείται με τη δομή, με τις μορφές και με τους νόμους της νόησης
- 2. ακολουθία ιδεών, σκέψεων, γεγονότων με εσωτερική συνέπεια, συνάφεια

Λεξικό της Κοινής Νεοελληνικής

Ινστιτούτο Νεοελληνικών Σπουδών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης

http://www.greek-language.gr/greekLang/modern_greek/tools/lexica/triantafyllides/search.html?lq=λογική&dq=

- Η λογική ασχολείται με δύο θεμελιώδεις δεξιότητες, τις οποίες θα πρέπει να έχει κάθε επιστήμονας :
 - Αφαίρεση
 - Τυπικότητα

Μαθηματική Λογική

- **Μαθηματική Λογική** : η ανάλυση των μεθόδων συλλογισμού.
- http://el.wikipedia.org/wiki/Μαθηματική_λογική
- Η Λογική ενδιαφέρεται για τη **μορφή** και όχι για το περιεχόμενο του συλλογισμού :
 - Όλα τα κουνέλια αγαπούν τα καρότα.
 - Ο Πλάτων είναι κουνέλι.
 - Άρα ο Πλάτων αγαπάει τα καρότα.
- Η **μορφή** του συλλογισμού :
 - Όλα τα A είναι B. Το S είναι A. Άρα το S είναι B.
- Η αλήθεια ή το ψεύδος των επιμέρους συλλογισμών δεν ενδιαφέρουν τη Λογική.

Ιστορική αναδρομή

- Πρώτα βήματα από τους φυσικούς φιλόσοφους της Ιωνίας, τους Ελεάτες και τους Σοφιστές
- Θεμελίωση από τον Αριστοτέλη
- Στα μοντέρνα χρόνια αναζωπύρωση με την ανακάλυψη μη-Ευκλείδιων γεωμετριών και τη θεωρητική θεμελίωση της Ανάλυσης
- Σύγχρονοι Σταθμοί :
 - Frege (1879) – τυπική γλώσσα για τα μαθηματικά και τη λογική
 - Cantor (1895-7) – Θεωρία Συνόλων
 - Russel (1930) – Αξιωματική θεμελίωση της Λογικής
 - Robinson (1965) - μέθοδος της Επίλυσης (resolution) για το χειρισμό συμβόλων και εκτέλεση μηχανικών αποδείξεων
 - Αρχές δεκαετίας 1970: οι Kowalski και Colmerauer προτείνουν τη Λογική σαν γλώσσα προγραμματισμού (Prolog)

Λογική και Βάσεις Δεδομένων και Γνώσεων

- Η Λογική και η Θεωρία Συνόλων αποτέλεσαν τη βάση για το Σχεσιακό Μοντέλο Δεδομένων
- Οι Βάσεις Γνώσεων αναπαριστούν γνώσεις με τη μορφή λογικών προτάσεων
- Η παραγωγή αποτελεσμάτων από βάσεις δεδομένων / γνώσεων γίνεται με *γλώσσες ερωτημάτων (query languages)* και *συμπερασματικούς κανόνες (inference rules)*

Προτασιακός λογισμός :

- Ασχολείται με τη δομή των προτάσεων και τη χρήση τους στην εξαγωγή συμπερασμάτων.
- Οδηγεί στην κατασκευή μιας **τυπικής** γλώσσας, με αυστηρούς κανόνες σχηματισμού και προτάσεων.
- Βασικά σύμβολα :
 - \neg άρνηση (δεν) - negation
 - \wedge σύζευξη (και) - conjunction
 - \vee διάζευξη (ή) - disjunction
 - \rightarrow συνεπαγωγή (αν ... τότε) - implication
 - \forall για καθένα
 - \exists υπάρχει

Σύνθετες προτάσεις

- Σε σύνθετες προτάσεις, χρησιμοποιούμε παρενθέσεις για να μην έχουμε σύγχυση εννοιών :

$$(p \wedge q) \vee r \quad \text{ή} \quad p \wedge (q \vee r)$$

- Κατά σύμβαση η σειρά ισχύος είναι :
 - Παρενθέσεις
 - Άρνηση
 - Σύζευξη
 - Διάζευξη
 - Συνεπαγωγή

Από «προτάσεις» σε προτάσεις

	Προτάσεις	Τυπική Γλώσσα
K	παρατηρούνται σημεία οικονομικής ύφεσης	
C	υπάρχουν περιθώρια κοινωνικής πολιτικής	
	Αν παρατηρούνται σημεία οικονομικής ύφεσης, τότε δεν υπάρχουν περιθώρια κοινωνικής πολιτικής	$(K \rightarrow (\neg C))$
	Υπάρχουν περιθώρια κοινωνικής πολιτικής και παρατηρούνται σημεία οικονομικής ύφεσης	$(C \wedge K)$
	Είτε δεν παρατηρούνται σημεία οικονομικής ύφεσης ή δεν υπάρχουν περιθώρια κοινωνικής πολιτικής	$((\neg K) \vee (\neg C))$
	Ούτε παρατηρούνται σημεία οικονομικής ύφεσης ούτε υπάρχουν περιθώρια κοινωνικής πολιτικής	$((\neg K) \wedge (\neg C)) \text{ ή } (\neg (C \vee K))$

Προτασιακοί τύποι

- Εκφράσεις : Πεπερασμένες ακολουθίες συμβόλων (π.χ. $\rightarrow \wedge \neg () \neg$)

Γενικευμένος επαγωγικός ορισμός :

- Προτασιακοί τύποι είναι οι εκφράσεις που ορίζονται, επαγωγικά, ως εξής :
 1. Τα σύμβολα προτάσεων είναι προτασιακοί τύποι
 2. Αν \mathbf{p}, \mathbf{q} προτασιακοί τύποι, τότε οι εκφράσεις $(\mathbf{p} \wedge \mathbf{q}), (\mathbf{p} \vee \mathbf{q}), (\mathbf{p} \rightarrow \mathbf{q}), (\neg \mathbf{p})$ είναι επίσης προτασιακοί τύποι
 3. Μόνο οι εκφράσεις που σχηματίζονται με διαδοχική εφαρμογή των 1. και 2. είναι προτασιακοί τύποι.

Σημασιολογική (semantics)

- Ως **ατομική, απλή πρόταση** θεωρείται μια απλή δηλωτική πρόταση, η οποία είναι είτε **αληθής** είτε **ψευδής**
- Η αλήθεια ή το ψεύδος μιας πρότασης σημειώνεται με την απόδοση μιας εκ των **τιμών αληθείας - truth values** **T** (αληθές - true) ή **F** (ψευδές - false)

Πίνακες αληθείας

P	¬P
T	F
F	T

Άρνηση

«Καθηγητές ή φοιτητές έχουν ειδική έκπτωση» :
Μη αποκλειστική διάζευξη

«Απόψε θα πάμε θέατρο ή κινηματογράφο» :
Αποκλειστική διάζευξη

P	Q	P ∧ Q
T	T	T
T	F	F
F	T	F
F	F	F

Σύζευξη

P	Q	P ∨ Q
T	T	T
T	F	T
F	T	T
F	F	F

Διάζευξη (μη αποκλειστική)

P	Q	P → Q
T	T	T
T	F	F
F	T	T
F	F	T

Συνεπαγωγή

Παράδειγμα 1

A_1	A_2	$(A_1 \wedge A_2) \vee A_1$	$(A_1 \vee A_2) \wedge A_1$
T	T	?	?
T	F	?	?
F	T	?	?
F	F	?	?

Παράδειγμα 1

A_1	A_2	$(A_1 \wedge A_2) \vee A_1$	$(A_1 \vee A_2) \wedge A_1$
T	T	?	?
T	F	?	?
F	T	?	?
F	F	?	?

$(A_1 \wedge A_2)$	$(A_1 \wedge A_2) \vee A_1$	$(A_1 \vee A_2)$	$(A_1 \vee A_2) \wedge A_1$
T	T	T	T
F	T	T	T
F	F	T	F
F	F	F	F

Παράδειγμα 2

A_1	A_2	$(A_1 \wedge A_2) \vee A_2$	$(A_1 \vee A_2) \wedge A_2$
T	T	?	?
T	F	?	?
F	T	?	?
F	F	?	?

Παράδειγμα 2

A_1	A_2	$(A_1 \wedge A_2) \vee A_2$	$(A_1 \vee A_2) \wedge A_2$
T	T	?	?
T	F	?	?
F	T	?	?
F	F	?	?

$(A_1 \wedge A_2)$	$(A_1 \wedge A_2) \vee A_2$	$(A_1 \vee A_2)$	$(A_1 \vee A_2) \wedge A_2$
T	T	T	T
F	F	T	F
F	T	T	T
F	F	F	F

Τι είναι σύνολο;

Ορισμός Cantor :

Σύνολο είναι μια οποιαδήποτε συνάθροιση σε ολότητα οριστικών και διακεκριμένων στοιχείων της διαίσθησης ή του στοχασμού μας

ΣΥΝΟΛΟ : Μια καλώς ορισμένη συλλογή αντικειμένων

- Τα σύνολα είναι μη διατεταγμένες συλλογές στοιχείων
- Τα στοιχεία συνήθως ονομάζονται με μικρά γράμματα
- Τα σύνολα συνήθως ονομάζονται με κεφαλαία γράμματα

Σύνολα αριθμών

- \mathbb{P} , το σύνολο όλων των πρώτων αριθμών.
- \mathbb{N} , το σύνολο όλων των φυσικών αριθμών.
- \mathbb{Z} , το σύνολο όλων των ακεραίων αριθμών.
- \mathbb{Q} , το σύνολο όλων των ρητών αριθμών.
- \mathbb{R} , το σύνολο όλων των πραγματικών αριθμών.
- \mathbb{C} , το σύνολο όλων των μιγαδικών αριθμών. Αυτό γράφεται και ως $\{z | z = x + \psi i, i^2 = -1\}$.

Ορισμοί και υποσύνολα

Δύο τρόποι ορισμού ενός συνόλου:

- Απαρίθμηση $A = \{1, 3, 5, 7, 9\}$
- Περιγραφή $B = \{x \mid x \text{ είναι περιττός}\}$
- Αν x ανήκει / δεν ανήκει σε ένα σύνολο C : $x \in C$ / $x \notin C$

Υποσύνολο : Το μέρος ενός συνόλου ,
π.χ. το B είναι υποσύνολο του A

- Αν το σύνολο A έχει έστω και ένα στοιχείο παραπάνω από το B , τότε το B είναι γνήσιο υποσύνολο του A :
 $B \subset A$. Εάν δεν μπορεί να διασφαλιστεί αυτό, $B \subseteq A$.

Μερικοί ακόμη ορισμοί

Για τα $A, B \subseteq U$ ισχύουν :

a) $A \cup B = \{x \mid x \in A \vee x \in B\}$

ένωση

b) $A \cap B = \{x \mid x \in A \wedge x \in B\}$

τομή

c) $A \Delta B = \{x \mid x \in A \cup B \wedge x \notin A \cap B\}$

συμμετρική διαφορά

Αμοιβαίως ασύνδετα

$$A \cap B = \phi$$

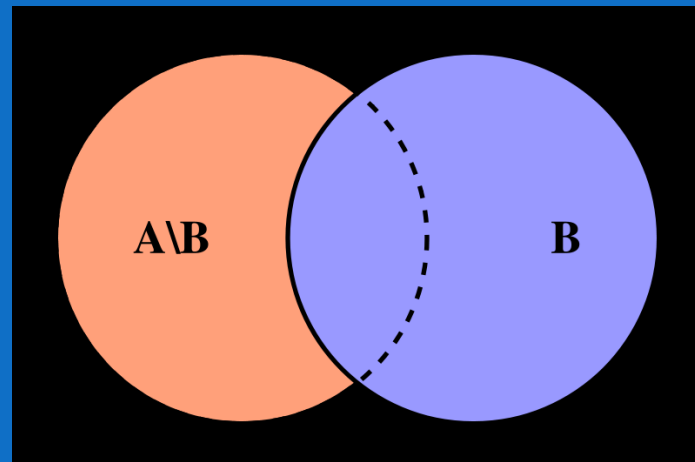
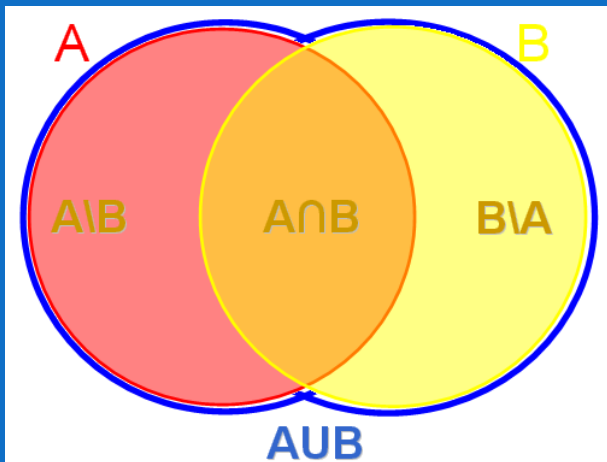
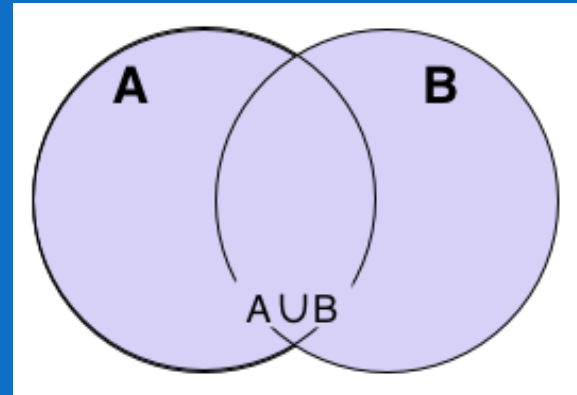
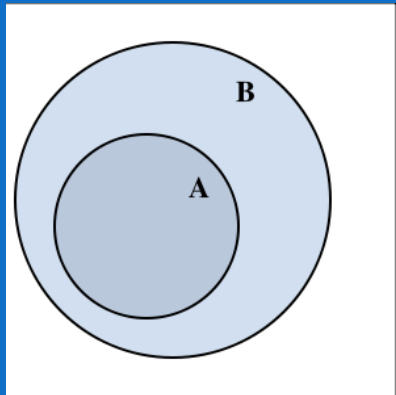
Συμπλήρωμα

$$\bar{A} = U - A = \{x \mid x \in U \wedge x \notin A\}$$

Σχετικό συμπλήρωμα του A στο B

$$B - A = \{x \mid x \in B \wedge x \notin A\}$$

Διαγράμματα Venn



Τυπικοί Ορισμοί

$$A \subseteq B \Leftrightarrow \forall x(x \in A \rightarrow x \in B)$$

$$A \subset B \Leftrightarrow (A \subseteq B) \wedge (A \neq B)$$

$$A = B \Leftrightarrow (A \subseteq B) \wedge (A \supseteq B)$$

$$A \neq B \Leftrightarrow A \subseteq B \vee A \supseteq B$$

Όταν τα A και B έχουν σχέση υπερσυνόλου - υποσυνόλου

$$A \cup B = \{x \mid x \in A \vee x \in B\}$$

$$A \cap B = \{x \mid x \in A \wedge x \in B\}$$

$$A \setminus B = \{x \mid x \in A \wedge x \notin B\}$$

$$\mathcal{P}(A) = \{x \mid x \subseteq A\} \quad \text{Δυναμοσύνολο του A}$$

$$\emptyset = \{x \mid x \neq x\}$$

Νόμοι συνόλων (1)

$$(1) \overline{\overline{A}} = A$$

Νόμος Διπλού Συμπληρώματος

$$(2) \overline{A \cup B} = \overline{A} \cap \overline{B}$$

Νόμος De Morgan

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

$$(3) A \cup B = B \cup A$$

Αντιμεταθετικοί Νόμοι

$$A \cap B = B \cap A$$

$$(4) A \cup (B \cup C) = (A \cup B) \cup C \quad \text{Προσεταιριστικοί Νόμοι}$$

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

$$(5) A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \text{Επιμεριστικοί Νόμοι}$$

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

Νόμοι συνόλων (2)

$$(6) A \cup A = A, A \cap A = A$$

$$(7) A \cup \emptyset = A, A \cap U = A$$

$$(8) A \cup \bar{A} = U, A \cap \bar{A} = \emptyset$$

$$(9) A \cup U = U, A \cap \emptyset = \emptyset$$

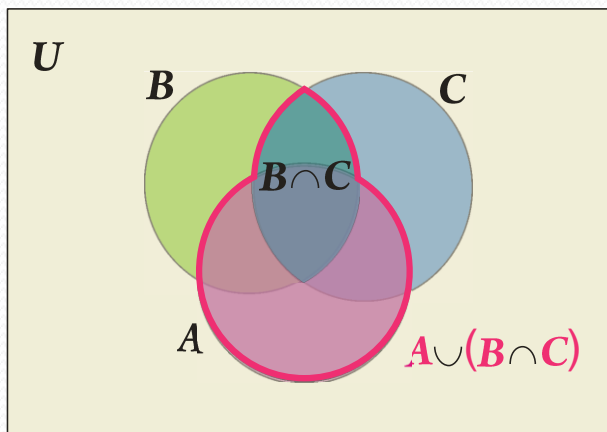
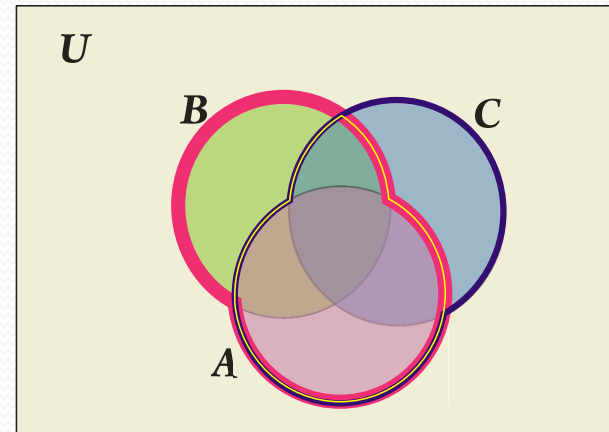
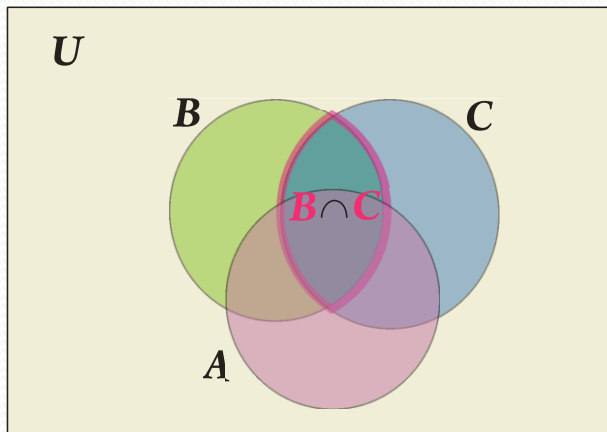
$$(10) A \cup (A \cap B) = A$$

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

$$(11) A \cup B \cup C =$$

$$A + B + C - (A \cap B + B \cap C + C \cap A) + A \cap B \cap C$$

Παράδειγμα με διαγράμματα



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

ΠΡΟΒΛΗΜΑ

Σε μια έρευνα για τους τρόπους επικοινωνίας των νέων σε καθημερινή βάση με 451 έγκυρες απαντήσεις καταγράφηκαν τα εξής :

- 406 άτομα χρησιμοποιούν το κινητό τους
- 405 άτομα χρησιμοποιούν το email τους
- 371 άτομα παρακολουθούν το προφίλ τους στο facebook

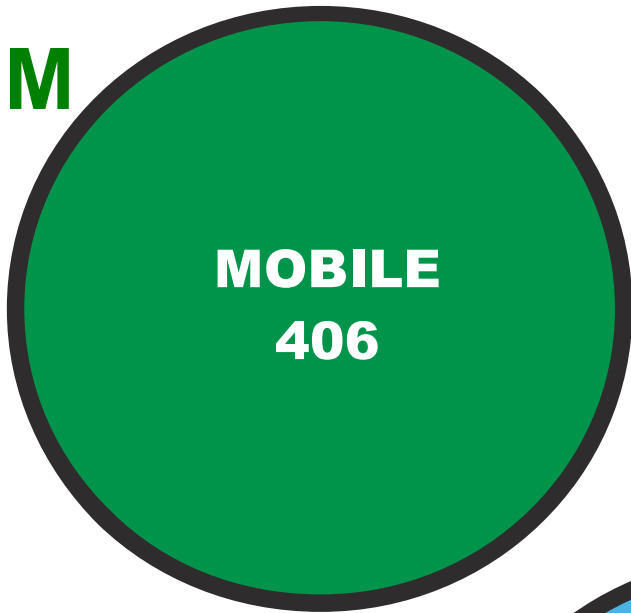
Επίσης διαπιστώθηκε ότι :

- 29 άτομα χρησιμοποιούν ΜΟΝΟ κινητό
- 377 άτομα χρησιμοποιούν κινητό και email
- 369 άτομα χρησιμοποιούν email και facebook
- 346 άτομα χρησιμοποιούν κινητό, email και facebook.

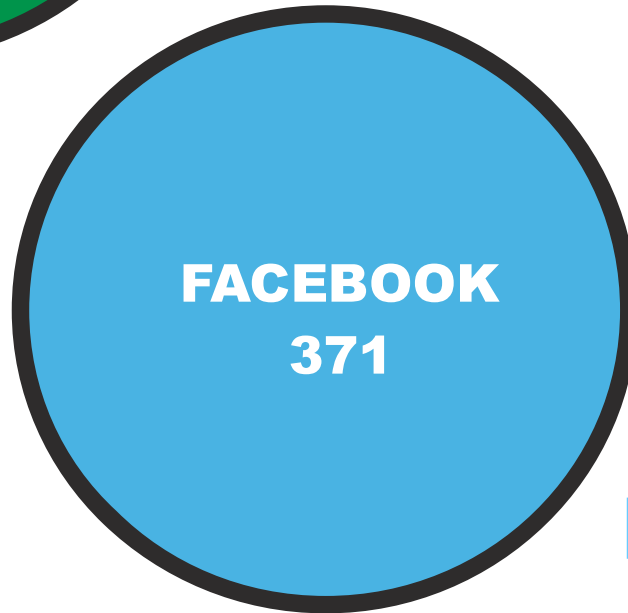
⇒Πόσα άτομα χρησιμοποιούν ΜΟΝΟ facebook ;

⇒Πόσα άτομα δεν χρησιμοποιούν τίποτε από τα παραπάνω ;

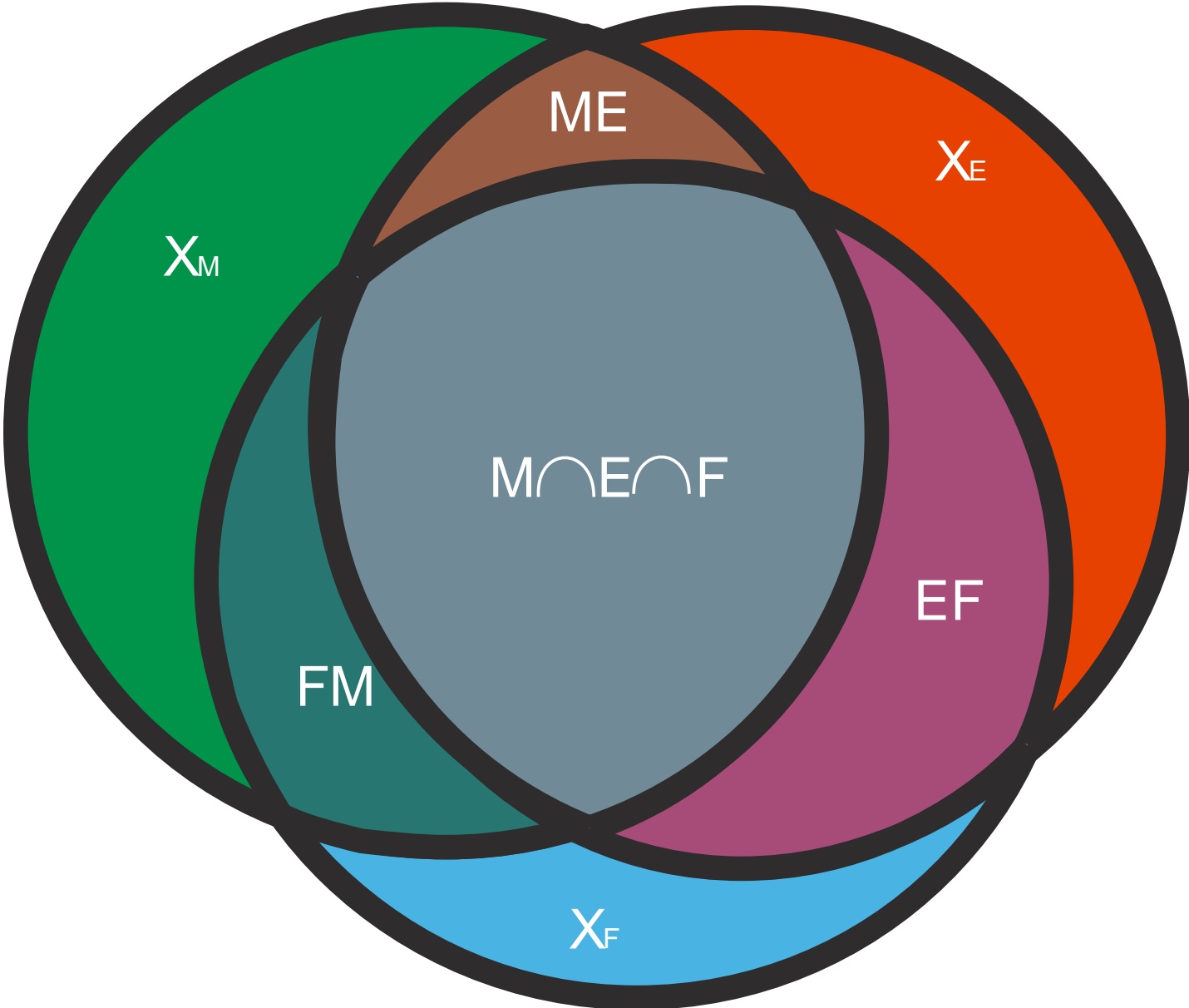
M

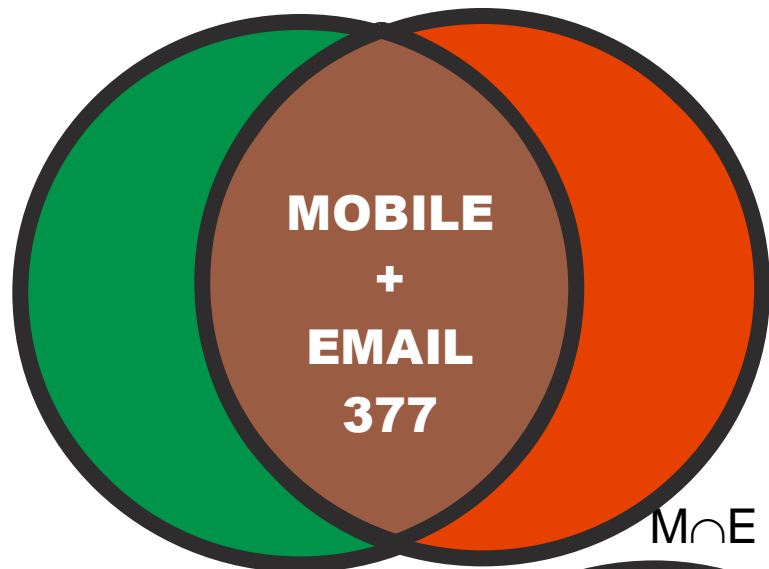


E

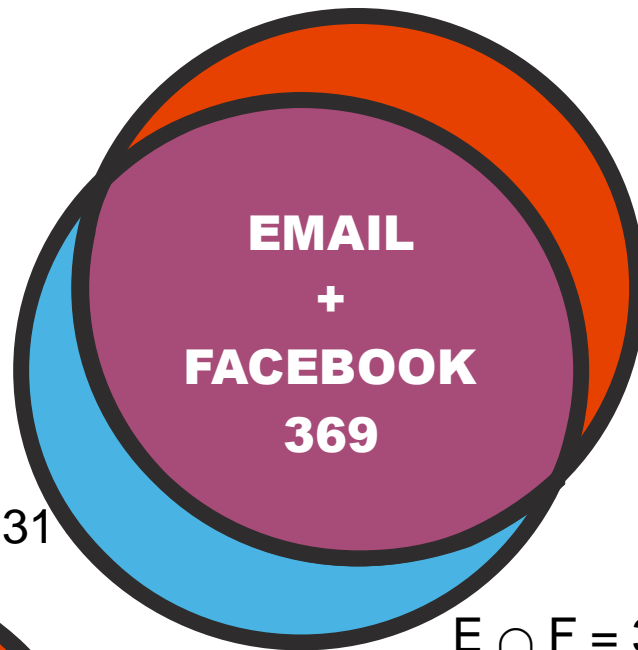


F



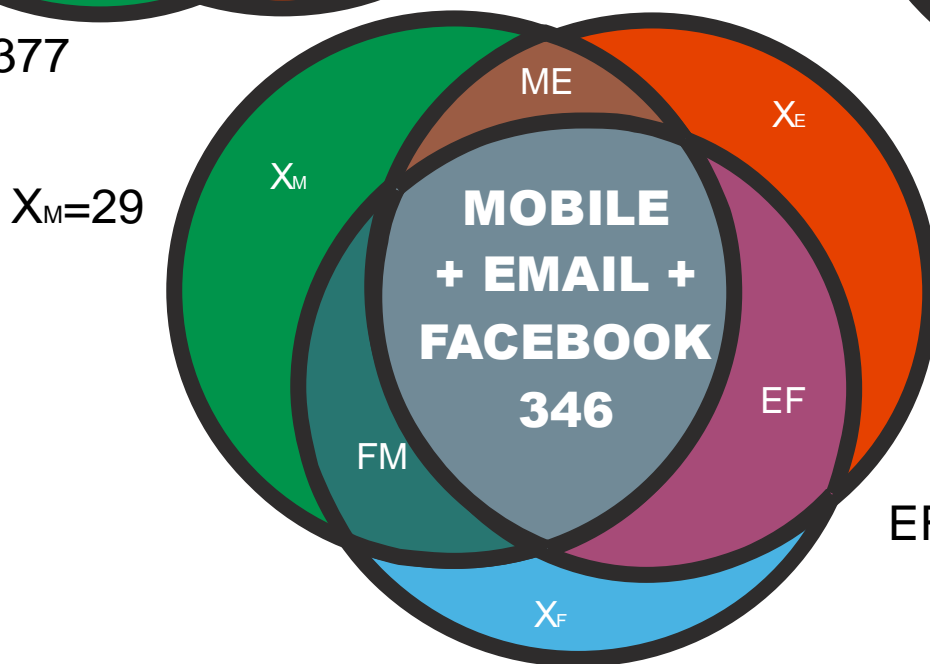


$$M \cap E = 377$$



$$E \cap F = 369$$

$$ME = M \cap E - M \cap E \cap F = 31$$



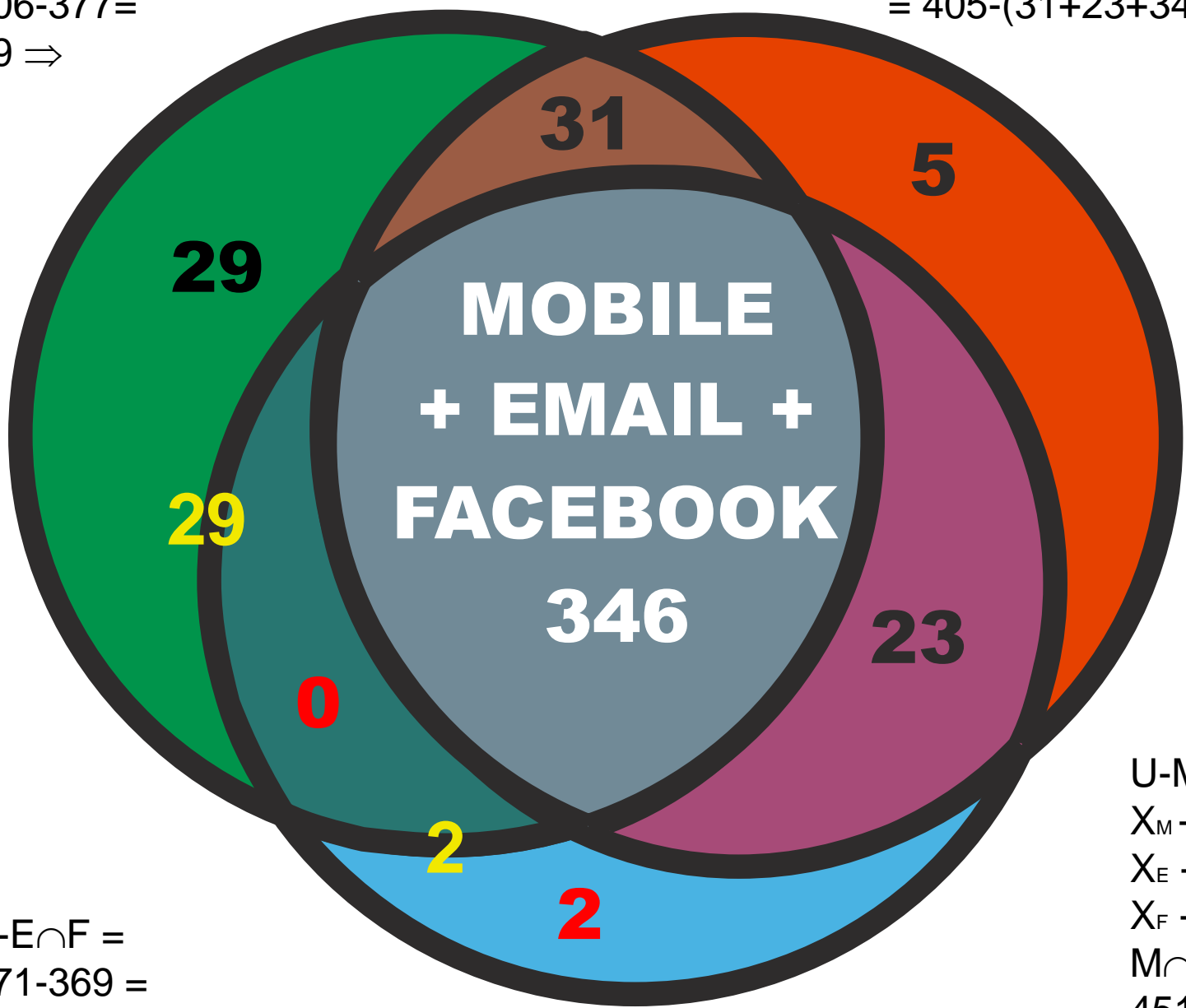
$$M \cap E \cap F = 346$$

$$EF = E \cap F - M \cap E \cap F = 23$$

$$\begin{aligned}
 X_M + FM &= M - M \cap E = \\
 &= 406 - 377 = \\
 &= 29 \Rightarrow
 \end{aligned}$$

$$FM = 0$$

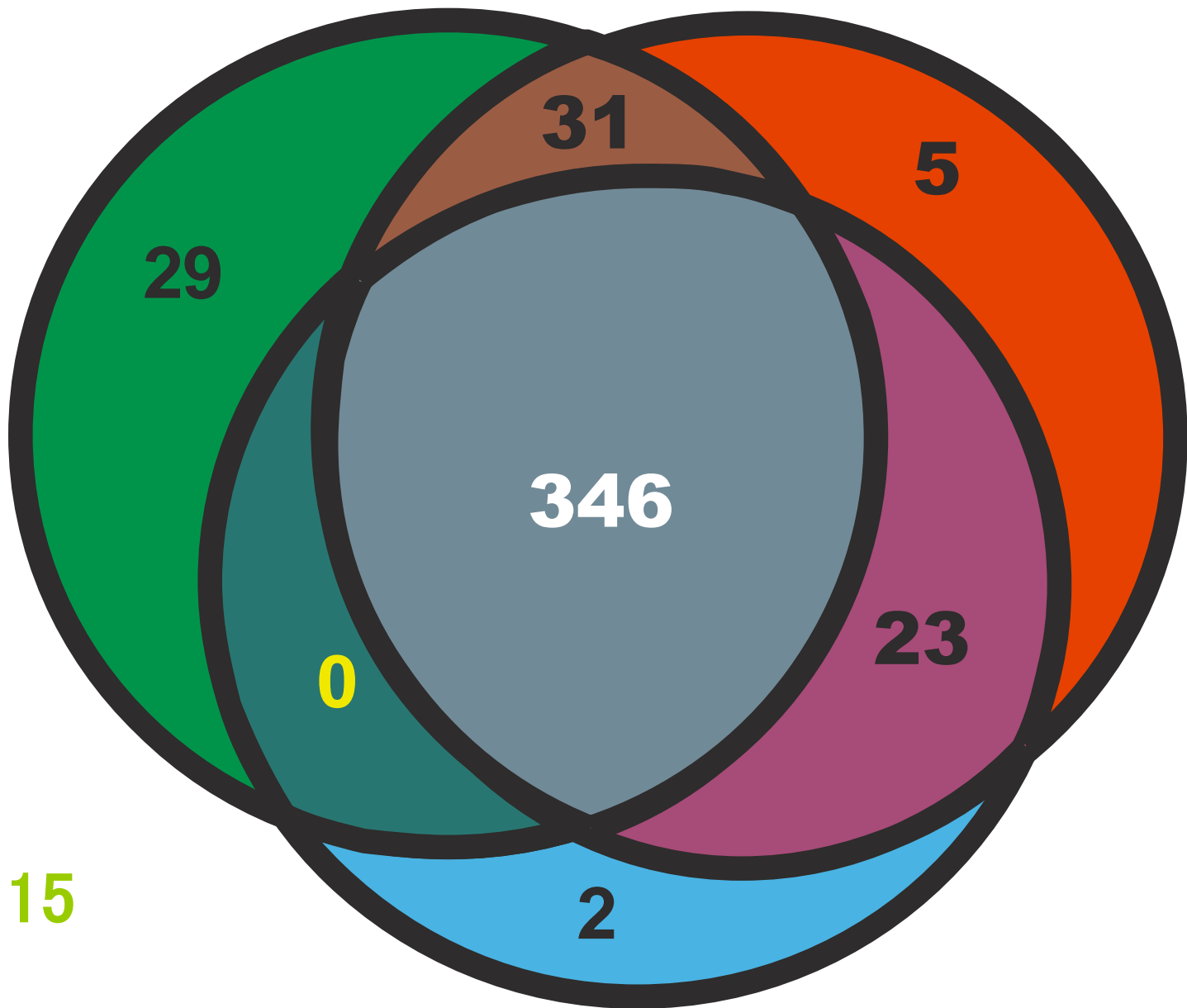
$$\begin{aligned}
 X_E &= E - (ME + EF + M \cap E \cap F + X_E) \\
 &= 405 - (31 + 23 + 346) = 5
 \end{aligned}$$



$$\begin{aligned}
 X_F + FM &= F - E \cap F = \\
 &= 371 - 369 = \\
 &= 2 \Rightarrow
 \end{aligned}$$

$$X_F = 2$$

$$\begin{aligned}
 U - M \cup E \cup F &= \\
 &= X_M + ME + \\
 &= X_E + EF + \\
 &= X_F + FM + \\
 &= M \cap E \cap F = \\
 &= 451 - 436 = 15
 \end{aligned}$$



Σχεσιακή Άλγεβρα

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Σχεσιακή Άλγεβρα

- Είναι το μαθηματικό υπόβαθρο του σχεσιακού μοντέλου δεδομένων.
- Αποτελεί μια θεωρητική γλώσσα που βασίζεται στη θεωρία των συνόλων.
- Προσφέρει ένα σύνολο πράξεων για τη διαχείριση των σχέσεων. Ορισμένες από τις πράξεις της προέρχονται από τη θεωρία συνόλων, ενώ άλλες αναπτύχθηκαν ειδικά για τη διαχείριση σχεσιακών β.δ.
- Η διαχείριση περιλαμβάνει την επιλογή γραμμών και στηλών οι οποίες πληρούν κάποια συγκεκριμένα κριτήρια, τόσο από ένα απλό πίνακα όσο και από το συνδυασμό δύο ή περισσότερων πινάκων.
- Όλες οι πράξεις της σχεσιακής άλγεβρας εφαρμόζονται πάνω σε μια σχέση και ως αποτέλεσμα παράγουν μια άλλη σχέση. Με τον τρόπο αυτό είναι τελικά δυνατή η πραγματοποίηση οποιασδήποτε ενέργειας, όσο πολύπλοκη και αν είναι.

Πίνακες και Πράξεις ΣΑ

- Ένας πίνακας (δηλ. μια σχέση) ορίζεται ως ένα σύνολο πλειάδων (tuples)
- Ο πίνακας διαθέτει όλες τις ιδιότητες ενός **συνόλου**
- Οι **πράξεις συνόλου (set operators)** αναφέρονται σε σύνολα:
 - Ένωση (Union), Τομή (Intersection), Διαφορά (Difference), Καρτεσιανό Γινόμενο (Cartesian Product)
- Οι **εγγενείς πράξεις (native operators)** αναφέρονται σε πλειάδες ή στήλες:
 - Επιλογή (Select), Προβολή (Project), Σύζευξη (Join) και η Διαίρεση (Division)

Επιλογή (select)

Επιλογή ενός υποσυνόλου των πλειάδων μιας σχέσης R που ικανοποιεί μια **συνθήκη επιλογής (selection condition)**

Συμβολισμός - Σύνταξη: $\sigma_{\text{συνθήκη-επιλογής}}(R)$

Ιδιότητες:

- Μοναδιαίος τελεστής (unary) : Εφαρμόζεται σε ένα μόνο πίνακα
- Αντιμεταθετική πράξη

$$\sigma_{\text{συνθ-1}}(\sigma_{\text{συνθ-2}}(R)) = \sigma_{\text{συνθ-2}}(\sigma_{\text{συνθ-1}}(R)) = \sigma_{\text{συνθ-1} \wedge \text{συνθ-2}}(R)$$

- βαθμός εξόδου = βαθμός εισόδου
- επιλεκτικότητα (selectivity): ποσοστό των πλειάδων που επιλέγονται

Παρατήρηση: η νέα σχέση που προκύπτει με την πράξη της επιλογής περιέχει όλα τα χαρακτηριστικά της αρχικής σχέσης.

Τελεστές Λογικοί και Σύγκρισης

- Στη συνθήκη επιλογής, μπορούν να χρησιμοποιηθούν οι **τελεστές σύγκρισης** $\{=, \neq, <, \leq, >, \geq\}$
- Οι τελεστές σύγκρισης $\{<, \leq, >, \geq\}$ προϋποθέτουν οι τιμές που παίρνουν τα πεδία πάνω στα οποία εφαρμόζονται να είναι **διατεταγμένες**.
- Εκτός από τους τελεστές σύγκρισης, χρησιμοποιούνται και οι **λογικοί τελεστές** AND (\wedge), OR (\vee) και NOT (\neg)
- **Μετονομασία (Rename)** : Το αποτέλεσμα της επιλογής μπορούμε να το αποθηκεύσουμε σε ενδιάμεσο πίνακα με τον τελεστή « \leftarrow »:

$$\text{RESULT} \leftarrow \sigma_{\text{DNO}=5 \wedge \text{SALARY}>25000}(\text{EMPLOYEE})$$

Παράδειγμα – επιλογή I

TAINIA

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

Ταινίες με διάρκεια μεγαλύτερη των 100 λεπτών)

$\sigma_{\text{διάρκεια} > 100}$ (TAINIA)

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

Παράδειγμα – επιλογή II

TAINIA

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

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

σ διάρκεια > 100 \wedge χρόνος > 1995 (TAINIA)

τίτλος	χρόνος	διάρκεια	είδος
Star Wars	1997	124	έγχρωμη

Προβολή (project)

Επιλογή ορισμένων στηλών από τον πίνακα μιας σχέσης και απορρίπτει τις υπόλοιπες

Συμβολισμός - Σύνταξη: $\pi_{\text{λίστα-πεδίων}}(R)$

Ιδιότητες:

- Μοναδιαίος τελεστής
- Απομακρύνονται οι διπλές πλειάδες (duplicate elimination)
- Μη αντιμεταθετική πράξη

$$\pi_{\text{λίστα-1}}(\pi_{\text{λίστα-2}}(R)) \neq \pi_{\text{λίστα-2}}(\pi_{\text{λίστα-1}}(R)) \text{ [εν γένει]}$$

$$\pi_{\text{λίστα-1}}(\pi_{\text{λίστα-2}}(R)) = \pi_{\text{λίστα-1}}(R) \text{ εφόσον η λίστα-1 περιέχει τα πεδία της λίστα-2}$$

- βαθμός εξόδου \leq βαθμός εισόδου

Παράδειγμα – προβολή I

ΤΑΙΝΙΑ

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

Τίτλος, χρόνος, διάρκεια των ταινιών

π τίτλος, χρόνος, διάρκεια (ΤΑΙΝΙΑ)

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

Παράδειγμα – προβολή II

TAINIA

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

Είδος ταινιών

$\pi_{\text{είδος}}$ (TAINIA)

είδος
έγχρωμη

Προσοχή: απαλοιφή διπλότυπων

Παράδειγμα project-select

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

$\pi_{\text{FNAME, LNAME, SALARY}} (\sigma_{\text{DNO}=5 \wedge \text{SALARY}>25000} (\text{EMPLOYEE}))$

Παράδειγμα project-select

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

$\pi_{\text{FNAME, LNAME, SALARY}} (\sigma_{\text{DNO}=5 \wedge \text{SALARY}>25000} (\text{EMPLOYEE}))$

Παράδειγμα project-select

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

$\pi_{\text{FNAME, LNAME, SALARY}} (\sigma_{\text{DNO}=5 \wedge \text{SALARY}>25000} (\text{EMPLOYEE}))$

FNAME	LNAME	SALARY
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000

Συμβατοί ως προς την ένωση

- Συμβατοί ως προς την ένωση πίνακες (**union compatible**):

Έστω $R(r_1, r_2, \dots, r_n)$, $S=(s_1, s_2, \dots, s_m)$. Αυτοί είναι συμβατοί, εάν $m=n$ και $dom(r_i) = dom(s_i)$ ($1 \leq i \leq n$)

- Η **τομή**, η **ένωση** και η **διαφορά** απαιτούν οι πίνακες επί των οποίων εφαρμόζονται, να είναι συμβατοί ως προς ένωση.

Ένωση, Τομή και Διαφορά

- Η **ένωση** δύο πινάκων, $R \cup S$, είναι ένας νέος πίνακας που περιλαμβάνει **όλες τις πλειάδες** των δύο πινάκων, οι οποίες μπορεί να ανήκουν μόνο στον R , μόνο στον S ή και στους δύο πίνακες ταυτόχρονα (**όχι διπλές πλειάδες**)
- Η **τομή** δύο πινάκων, $R \cap S$, είναι ένας νέος πίνακας, που περιλαμβάνει **όλες τις κοινές πλειάδες** από τους δύο πίνακες
- Η **διαφορά** των πινάκων, $R - S$, είναι ένας νέος πίνακας, που περιλαμβάνει όλες τις πλειάδες που ανήκουν στον R αλλά όχι στον S
- Ισχύουν $R \cup S = S \cup R$ και $R \cap S = S \cap R$ αλλά $R - S \neq S - R$
- Οι πράξεις μεταξύ πινάκων είναι **δυαδικές (binary)** -- δύο πίνακες

Παραδείγματα Πράξεων

Example

R

A	B	C
a_1	b_1	c_1
a_1	b_2	c_3
a_2	b_1	c_2

S

A	B	C
a_1	b_1	c_1
a_1	b_1	c_2
a_1	b_2	c_3
a_3	b_2	c_3

Διαφορά: Σχετικό Συμπλήρωμα

$R - S$

A	B	C
a_2	b_1	c_2

$S - R$

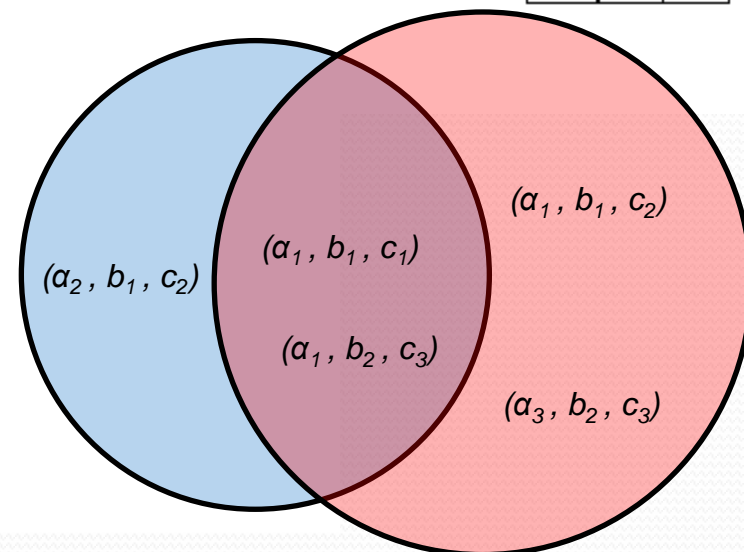
A	B	C
a_1	b_1	c_2
a_3	b_2	c_3

$R \cup S$

A	B	C
a_1	b_1	c_1
a_1	b_1	c_2
a_1	b_2	c_3
a_2	b_1	c_2
a_3	b_2	c_3

$R \cap S$

A	B	C
a_1	b_1	c_1
a_1	b_2	c_3



STUDENT	
FIRSTNAME	LASTNAME
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

INSTRUCTOR	
FNAME	LNAME
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

FIRSTNAME	LASTNAME
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert
John	Smith
Ricardo	Browne
Francis	Johnson

STUDENT \cup INSTRUCTOR

FIRSTNAME	LASTNAME
Susan	Yao
Ramesh	Shah

STUDENT \cap INSTRUCTOR

FIRSTNAME	LASTNAME
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

STUDENT - INSTRUCTOR

FIRSTNAME	LASTNAME
John	Smith
Ricardo	Browne
Francis	Johnson

INSTRUCTOR - STUDENT

Καρτεσιανό Γινόμενο

Το καρτεσιανό γινόμενο $R \times S$ όπου $R (r_1, r_2, \dots, r_n)$ και $S (s_1, s_2, \dots, s_m)$ ορίζεται ως ένας πίνακας με $m+n$ πεδία

$$R \times S = \{(r_1, r_2, \dots, r_n, s_1, s_2, \dots, s_m) \mid (r_1, r_2, \dots, r_n) \in R, (s_1, s_2, \dots, s_m) \in S\}$$

ή

$$R \times S = \{(r, s) \mid r \in R \wedge s \in S\}$$

Εάν ο πίνακας R περιέχει n πλειάδες και ο πίνακας S περιέχει m πλειάδες, ο πίνακας $R \times S$ θα περιέχει συνολικά $n \times m$ πλειάδες.

Παράδειγμα Καρτεσιανού I

R

A	B	C
a_1	b_1	c_1
a_1	b_2	c_3
a_2	b_1	c_2

S

B	C	D
b_1	c_1	d_1
b_1	c_1	d_3
b_2	c_2	d_2
b_1	c_2	d_4

$R \times S$

R.A	R.B	R.C	S.B	S.C	S.D
a_1	b_1	c_1	b_1	c_1	d_1
a_1	b_1	c_1	b_1	c_1	d_3
a_1	b_1	c_1	b_2	c_2	d_2
a_1	b_1	c_1	b_1	c_2	d_4
a_1	b_2	c_3	b_1	c_1	d_1
a_1	b_2	c_3	b_1	c_1	d_3
a_1	b_2	c_3	b_2	c_2	d_2
a_1	b_2	c_3	b_1	c_2	d_4
a_2	b_1	c_2	b_1	c_1	d_1
a_2	b_1	c_2	b_1	c_1	d_3
a_2	b_1	c_2	b_2	c_2	d_2
a_2	b_1	c_2	b_1	c_2	d_4

Παράδειγμα Καρτεσιανού II

TEACHER

<u>Teacher ID</u>	TFNAME	TLNAME
1	ΒΑΣΙΛΕΙΟΣ	ΒΑΣΙΛΕΙΟΥ
2	ΓΕΩΡΓΙΟΣ	ΓΕΩΡΓΙΟΥ
3	ΑΝΔΡΕΑΣ	ΑΝΔΡΕΟΥ

COURSE

<u>Course ID</u>	CNAME	CHOURS
211	ΒΔ	4
212	MATH 1	3
210	MATH 2	3

TEACHER x COURSE

Teacher_ID	TFNAME	TLNAME	Course_ID	CNAME	CHOURS
1	ΒΑΣΙΛΕΙΟΣ	ΒΑΣΙΛΕΙΟΥ	211	ΒΔ	4
1	ΒΑΣΙΛΕΙΟΣ	ΒΑΣΙΛΕΙΟΥ	212	MATH 1	3
1	ΒΑΣΙΛΕΙΟΣ	ΒΑΣΙΛΕΙΟΥ	210	MATH 2	3
2	ΓΕΩΡΓΙΟΣ	ΓΕΩΡΓΙΟΥ	211	ΒΔ	4
2	ΓΕΩΡΓΙΟΣ	ΓΕΩΡΓΙΟΥ	212	MATH 1	3
2	ΓΕΩΡΓΙΟΣ	ΓΕΩΡΓΙΟΥ	210	MATH 2	3
3	ΑΝΔΡΕΑΣ	ΑΝΔΡΕΟΥ	211	ΒΔ	4
3	ΑΝΔΡΕΑΣ	ΑΝΔΡΕΟΥ	212	MATH 1	3
3	ΑΝΔΡΕΑΣ	ΑΝΔΡΕΟΥ	210	MATH 2	3

Παράδειγμα Καρτεσιανού III

ΒΙΒΛΙΑ

Κωδικός	Τίτλος
7071	Νοημοσύνη
403	Ψυχολογία
7023	C++

(α)

ΣΥΓΓΡΑΦΕΙΣ

Κωδικός Βιβλίου	Όνομα
7071	Γεωργίου
7071	Δημητρίου
403	Ασκητής
7023	Ζαφειριού
7023	Σταματίου

(β)

ΒΙΒΛΙΑ × ΣΥΓΓΡΑΦΕΙΣ

Κωδικός	Τίτλος	Κωδικός Βιβλίου	Όνομα
7071	Νοημοσύνη	7071	Γεωργίου
7071	Νοημοσύνη	7071	Δημητρίου
7071	Νοημοσύνη	403	Ασκητής
7071	Νοημοσύνη	7023	Ζαφειριού
7071	Νοημοσύνη	7023	Σταματίου
403	Ψυχολογία	7071	Γεωργίου
403	Ψυχολογία	7071	Δημητρίου
403	Ψυχολογία	403	Ασκητής
403	Ψυχολογία	7023	Ζαφειριού
403	Ψυχολογία	7023	Σταματίου
7023	C++	7071	Γεωργίου
7023	C++	7071	Δημητρίου
7023	C++	403	Ασκητής
7023	C++	7023	Ζαφειριού
7023	C++	7023	Σταματίου

Πράξεις σύζευξης (join)

- Το καρτεσιανό γινόμενο ($R \times S$) συνήθως είναι μια πολύ μεγαλύτερη σχέση από αυτή που χρειαζόμαστε
- Ορίζουμε ένα σύνολο πράξεων **σύζευξης (JOIN)** οι οποίες μπορούν να παραχθούν (εκφραστούν) ως ακολουθίες πράξεων καρτεσιανού γινομένου και επιλογής, με διάφορες παραλλαγές :
 - **θ-σύζευξη (Theta-join)**
 - **Σύζευξη ισότητας (Equijoin)**
 - **Φυσική σύζευξη (Natural join)**
 - **Εξωτερική σύζευξη (Outer join)**

θ-Σύζευξη (θ-join)

θ-Σύζευξη (join - \bowtie) συνδυασμός του καρτεσιανού γινομένου και μιας συνθήκης επιλογής :

$$R \times S = \{ (r, s) \mid r \in R \wedge s \in S \}$$

$$R \bowtie_F S = \{ (r, s) \mid r \in R \wedge s \in S \wedge F((r, s)) \} = \sigma_F(R \times S)$$

- Η **θήτα-σύζευξη** δύο σχέσεων R και S υπό τη συνθήκη (κατηγόρημα) F ορίζει μια σχέση που περιέχει τις πλειάδες του καρτεσιανού γινομένου των R και S οι οποίες ικανοποιούν το κατηγόρημα F .
- Το κατηγόρημα F είναι κριτήριο επιλογής στη μορφή $R.r_{ij}$ op $S.s_{ik}$ όπου op είναι ένας τελεστής σύγκρισης ($<$, $<=$, $>$, $>=$, $=$, $!=$), i : γραμμή, j, k : στήλες

Παράδειγμα θ-σύζευξης

Car

CarModel	CarPrice
CarA	20'000
CarB	30'000
CarC	50'000

Boat

BoatModel	BoatPrice
Boat1	10'000
Boat2	40'000
Boat3	60'000

Car ⋈ *Boat*

CarPrice ≥ *BoatPrice*

CarModel	CarPrice	BoatModel	BoatPrice
CarA	20'000	Boat1	10'000
CarB	30'000	Boat1	10'000
CarC	50'000	Boat1	10'000
CarC	50'000	Boat2	40'000

CarA	20.000	Boat1	10.000
CarA	20.000	Boat2	40.000
CarA	20.000	Boat3	60.000
CarB	30.000	Boat1	10.000
CarB	30.000	Boat2	40.000
CarB	30.000	Boat3	60.000
CarC	50.000	Boat1	10.000
CarC	50.000	Boat2	40.000
CarC	50.000	Boat3	60.000

Σύζευξη ισότητας και φυσική

- Εάν η συνθήκη επιλογής είναι ισότητα, τότε η σύζευξη ονομάζεται **σύζευξη ισότητας (equijoin)** και το χαρακτηριστικό της είναι ότι έχει δύο πεδία που εμφανίζουν **ίδια τιμή**.
- Εάν οι δύο προς σύζευξη σχέσεις έχουν ένα κοινό πεδίο και η σύζευξη γίνεται επ' αυτού, τότε ονομάζεται **φυσική σύζευξη (natural join)** και συμβολίζεται με * (π.χ. πρωτεύον κλειδί – εξωτερικό κλειδί).
- Η φυσική σύζευξη είναι πράξη σύζευξης ισότητας, της οποίας το αποτέλεσμα περιέχει μόνο το κοινό πεδίο (το αριστερό).

Παράδειγμα equijoin

TEACHERS

TEACHERID	TFNAME	FLNAME	ΕΙΔΙΚΟΤΗΤΑ	SCHOOL	NOMOS
1	ΑΝΔΡΕΑΣ	ΑΝΔΡΕΟΥ	ΦΥΣΙΚΟΣ	ΕΜΠ	ΑΤΤΙΚΗ
2	ΓΕΩΡΓΙΟΣ	ΓΕΩΡΓΙΟΥ	ΜΑΘΗΜΑΤΙΚΟΣ	ΕΜΠ	ΒΟΙΩΤΙΑ
3	ΒΑΣΙΛΕΙΟΣ	ΒΑΣΙΛΕΙΟΥ	ΗΥ	ΔΠΘ	ΘΕΣΣΑΛΟΝΙΚΗ
4	ΠΕΤΡΟΥ	ΠΕΤΡΟΣ	ΜΑΘΗΜΑΤΙΚΟΣ	ΑΠΘ	ΚΑΤΕΡΙΝΗ
5	ΔΗΜΗΤΡΙΟΥ	ΔΗΜΗΤΡΙΟΣ	ΗΥ	ΕΜΠ	ΑΤΤΙΚΗ
6	ΕΛΕΥΘΕΡΙΟΥ	ΕΛΕΥΘΕΡΙΟΣ	ΤΟΠΟΓΡΑΦΟΣ	ΕΜΠ	ΚΟΡΙΝΘΙΑ
9	ΠΑΝΑΓΙΩΤΟΥ	ΠΑΝΑΓΙΩΤΗΣ	ΗΥ	ΑΠΘ	ΘΡΑΚΗ

UNIVERSITIES

UCODE	UNAME	USTUDENTS	USTAFF	LOCATION
ΑΠΘ	ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ	8900	450	ΘΕΣΣΑΛΟΝΙΚΗ
ΔΠΘ	ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ	5000	389	ΘΡΑΚΗ
ΕΜΠ	ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ	10000	1000	ΑΤΤΙΚΗ
ΧΠΑ	ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ	3000	100	ΑΤΤΙΚΗ

TEACHERS * TEACHERS.SCHOOL=UNIVERSITIES.UCODE UNIVERSITIES

TEACHE RID	TFNAME	FLNAME	ΕΙΔΙΚΟΤΗΤΑ	SCH OOL	NOMOS	UCODE	UNAME	USTUDENTS	USTAFF	LOCATION
1	ΑΝΔΡΕΑΣ	ΑΝΔΡΕΟΥ	ΦΥΣΙΚΟΣ	ΕΜΠ	ΑΤΤΙΚΗ	ΕΜΠ	ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ	10000	1000	ΑΤΤΙΚΗ
2	ΓΕΩΡΓΙΟΣ	ΓΕΩΡΓΙΟΥ	ΜΑΘΗΜΑΤΙΚΟΣ	ΕΜΠ	ΒΟΙΩΤΙΑ	ΕΜΠ	ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ	10000	1000	ΑΤΤΙΚΗ
3	ΒΑΣΙΛΕΙΟΣ	ΒΑΣΙΛΕΙΟΥ	ΗΥ	ΔΠΘ	ΘΕΣΣΑΛΟΝΙΚΗ	ΔΠΘ	ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ	5000	389	ΘΡΑΚΗ
4	ΠΕΤΡΟΥ	ΠΕΤΡΟΣ	ΜΑΘΗΜΑΤΙΚΟΣ	ΑΠΘ	ΚΑΤΕΡΙΝΗ	ΑΠΘ	ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ	8900	450	ΘΕΣΣΑΛΟΝΙΚΗ
5	ΔΗΜΗΤΡΙΟΥ	ΔΗΜΗΤΡΙΟΣ	ΗΥ	ΕΜΠ	ΑΤΤΙΚΗ	ΕΜΠ	ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ	10000	1000	ΑΤΤΙΚΗ
6	ΕΛΕΥΘΕΡΙΟΥ	ΕΛΕΥΘΕΡΙΟΣ	ΤΟΠΟΓΡΑΦΟΣ	ΕΜΠ	ΚΟΡΙΝΘΙΑ	ΕΜΠ	ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ	10000	1000	ΑΤΤΙΚΗ
9	ΠΑΝΑΓΙΩΤΟΥ	ΠΑΝΑΓΙΩΤΗΣ	ΗΥ	ΑΠΘ	ΘΡΑΚΗ	ΑΠΘ	ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ	8900	450	ΘΕΣΣΑΛΟΝΙΚΗ

Παράδειγμα natural/equi join

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

DEPT_MRG ← DEPARTMENT ⋈_{SSN=MGRSSN} EMPLOYEE

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE	FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
Research	5	333445555	22-May-78	Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5
Administration	4	987654321	01-Jan-85	Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Headquarters	1	888665555	19-Jun-71	James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPT_MRG ← DEPARTMENT *_{(MGRSSN),(SSN)} EMPLOYEE

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE	FNAME	MINIT	LNAME	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
Research	5	333445555	22-May-78	Franklin	T	Wong	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5
Administration	4	987654321	01-Jan-85	Jennifer	S	Wallace	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Headquarters	1	888665555	19-Jun-71	James	E	Borg	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPT_MRG ← EMPLOYEE *_{(SSN),(MGRSSN)} DEPARTMENT

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO	DNAME	DNUMBER	MGRSTARTDATE
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5	Research	5	22-May-78
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4	Administration	4	01-Jan-85
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1	Headquarters	1	19-Jun-71

Εξωτερική σύζευξη (outer join)

- Αντίθετα με την απλή σύζευξη, στην **εξωτερική σύζευξη (outer join +)** διατηρούνται ΟΛΕΣ οι πλειάδες που προκύπτουν, ακόμη και εκείνες που δεν πληρούν τον κανόνα επιλογής
- Οι πλειάδες του πρώτου πίνακα που δεν πληρούν το κριτήριο επιλογής συνδυάζονται με «κενές» πλειάδες του άλλου πίνακα (NULL)
- Η εξωτερική σύζευξη μπορεί να είναι είτε **αριστερή (outer left join)** είτε **δεξιά (outer right join)**

Παράδειγμα outer join (δες XLS)

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Research	5	333445555	22-May-78
Administration	4	987654321	01-Jan-85
Headquarters	1	888665555	19-Jun-71

TEMP ← **EMPLOYEE (+)** MGRSSN=SSN **DEPARTMENT**

RESULT ← **π** FNAME, MINIT, LNAME, DNAME (**TEMP**)

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO	DNAME	DNUMBER	MGRSTARTDATE
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5	NULL	NULL	NULL
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5	Research	5	22-May-78
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4	NULL	NULL	NULL
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4	Administration	4	01-Jan-85
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5	NULL	NULL	NULL
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5	NULL	NULL	NULL
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4	NULL	NULL	NULL
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1	Headquarters	1	19-Jun-71

Λοιποί τελεστές

- **Καταμέτρηση (COUNT)** των εγγραφών ενός πίνακα
- Υπολογισμός **αθροίσματος (SUM)**
- Υπολογισμός **μέσου όρου (AVERAGE)** των τιμών μιας στήλης
- Εύρεση της **μέγιστης/ελάχιστης τιμής (MAXIMUM/MINIMUM)**
- Οι παραπάνω συναρτήσεις μπορούν να εφαρμοσθούν πάνω σε όλες τις γραμμές του πίνακα ή σε μέρος του :

$\langle \text{grouping attributes} \rangle \mathcal{F} \langle \text{function list} \rangle (\langle \text{relation name} \rangle)$

Παράδειγμα τελεστών

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1

DNO \mathcal{F} (COUNT SSN), (AVERAGE SALARY) (EMPLOYEE)

John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	333445555	5	30750
Franklin	T	Wong	333445555	08-Dec-45	638 Vass, Houston, TX	M	30000	888665555	5	
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5	
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5	

DNO	EMP_NO	AVG_SAL
5	4	30750
4	3	31000
1	1	55000

Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, Spring, TX	F	25000	987654321	4	31000
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4	
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4	

James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000	NULL	1	55000
-------	---	------	-----------	-----------	------------------------	---	-------	------	---	-------

Διαίρεση

Έστω $T (r_1, \dots, r_n, s_1, \dots, s_m)$, $S (s_1, \dots, s_m)$

Η διαίρεση $T \div S$ των T και S είναι ένας πίνακας R με σχήμα $R (r_1, \dots, r_n)$

Ο R περιέχει τις πλειάδες r_i οι οποίες είναι τέτοιες ώστε, για κάθε πλειάδα s_i του S , η πλειάδα που προκύπτει από την παράθεση των r_i και s_i ανήκει στον R .

Εν γένει, αν $R = T \div S$, τότε η R είναι το μεγαλύτερο δυνατό σύνολο πλειάδων ώστε $R \times S \subseteq T$

Παράδειγμα διαίρεσης I

T

A	B	C
a1	b1	c1
a2	b1	c1
a1	b2	c1
a1	b2	c2
a2	b1	c2
a1	b2	c3
a1	b2	c4
a1	b1	c5

S

C
c1

$$R := T \div S$$

A	B
a1	b1
a2	b1
a1	b2

$$R \times S \subseteq T$$

Παράδειγμα διαίρεσης II

T

A	B	C
a_1	b_1	c_1
a_2	b_1	c_1
a_1	b_2	c_1
a_1	b_2	c_2
a_2	b_1	c_2
a_1	b_2	c_3
a_1	b_2	c_4
a_1	b_1	c_5

S

C
c_1
c_2

$R = T : S$

A	B
a_1	b_2
a_2	b_1

A	B
a_1	b_1
a_1	b_2
a_2	b_1

Προτεραιότητα τελεστών

Προτεραιότητα	Τελεστής	Σύμβολο
Υψηλότερη	Προβολή	π
↓	Επιλογή	σ
↓	Καρτεσιανό Γινόμενο	\times
↓	Σύζευξη, Διαίρεση	$ X , \div$
↓	Διαφορά	$-$
Χαμηλότερη	Ένωση, Τομή	\cup, \cap

Περισσότερα...

- http://en.wikipedia.org/wiki/Relational_algebra

Structured Query Language (SQL)

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Structured Query Language

- Κοινή γλώσσα για όλα τα σχεσιακά ΣΔΒΔ: η γλώσσα SQL.
- Κάθε σχεσιακό ΣΔΒΔ έχει τη δυνατότητα να δεχθεί και να εκτελέσει εντολές της γλώσσας SQL.
- Βασίζεται στη σχεσιακή άλγεβρα, αν και δεν είναι αυστηρή υλοποίηση των πράξεών της
- Είναι μια δηλωτική γλώσσα (μη-διαδικαστική) στην οποία περιγράφουμε τι θέλουμε κι όχι πώς θα το αποκτήσουμε
- Μπορεί να ενσωματωθεί σε γλώσσες προγραμματισμού
- Το λογισμικό του σχεσιακού ΣΔΒΔ αναλαμβάνει την εύρεση του τρόπου ανάκτησης των δεδομένων από τη βάση

Σύντομο ιστορικό της SQL

- Η SQL αναπτύχθηκε το 1974 από τον Donald D. Chamberlin Raymond F. Boyce, που εργαζόταν στην **IBM** και αρχικά χρησιμοποιήθηκε στα δικά της ΣΔΒΔ.
- Στα τέλη της δεκαετίας του '70, η **ORACLE**[®] δημιούργησε το δικό της ΣΔΒΔ, αλλά υιοθέτησε την SQL.
- Σύντομα ακολούθησαν και άλλες εταιρείες λογισμικού που ανέπτυξαν τα δικά τους συστήματα τα οποία υιοθέτησαν την SQL.
- Η ευρεία διάδοση της SQL οφείλεται στο γεγονός ότι τόσο το Αμερικανικό Ινστιτούτο Τυποποίησης (American National Standards Institute — ANSI) όσο και ο Παγκόσμιος Οργανισμός Τυποποίησης (International Organization for Standardization — ISO) όρισαν την SQL ως την κοινή γλώσσα των ΣΔΒΔ.

Πρότυπα της SQL

- Παρά τις σημαντικές προσπάθειες που καταβλήθηκαν από τους διεθνείς οργανισμούς, ο στόχος των προτύπων της SQL δεν επιτεύχθηκε απόλυτα:
- Τα σχεσιακά ΣΔΒΔ υποστηρίζουν εκδοχές της SQL με κεντρικό τους πυρήνα το πρότυπο, παρέχοντας όμως πολλές πρόσθετες δυνατότητες (επεκτάσεις) που δεν συμπεριλαμβάνονται στο πρότυπο.

Εντολές SQL

- Οι εντολές της SQL μπορεί να πληκτρολογηθούν από το χρήστη και να υποβληθούν προς άμεση εκτέλεση στο σχεσιακό ΣΔΒΔ και τα αποτελέσματα των εντολών SQL εμφανίζονται στην οθόνη του χρήστη.
- Η SQL είναι μία γλώσσα για τη διαχείριση δεδομένων – δεν μία πλήρης γλώσσα προγραμματισμού όπως η C++ ή η Java
- Πολλές φορές είναι απαραίτητο η SQL να ενσωματωθεί μέσα στον κώδικα μίας άλλης γλώσσας προγραμματισμού. Τότε, η SQL αναφέρεται ως ενσωματωμένη SQL (embedded SQL).
- Στην περίπτωση αυτή, τα αποτελέσματα των SQL εντολών γνωστοποιούνται στο πρόγραμμα που είναι γραμμένο στη γενική γλώσσα προγραμματισμού, το οποίο και τα χειρίζεται.

Εισαγωγικά για τις εντολές SQL

- Οι εντολές της SQL θυμίζουν έντονα απλές προτάσεις της Αγγλικής γλώσσας — άλλωστε έγινε προσπάθεια γι' αυτό, ώστε να είναι εύκολη η εκμάθησή της.
- Κάθε εντολή περιέχει ένα σύνολο **δεσμευμένων λέξεων**, δηλαδή λέξεων που έχουν συγκεκριμένο νόημα που δεν μπορεί να αλλάξει, και ένα σύνολο λέξεων που ορίζονται από το χρήστη.
- Ο τρόπος γραφής μίας εντολής SQL είναι αυστηρά καθορισμένος και ακολουθεί συγκεκριμένους συντακτικούς κανόνες.
- Στην περίπτωση που οι κανόνες αυτοί παραβιάζονται, η εντολή δεν μπορεί να γίνει κατανοητή από το ΣΔΒΔ.

Δομικές μονάδες SQL I

- Δομικές μονάδες μιας τυπικής εκδοχής της SQL :
 - **Γλώσσα ορισμού δεδομένων (Data Definition Language, DDL):** εντολές που μας επιτρέπουν να υλοποιήσουμε σχέσεις και γενικά όλη τη δομή μιας βάσης δεδομένων.
 - **Γλώσσα χειρισμού δεδομένων (Data Manipulation Language, DML):** διαχείριση των δεδομένων της εφαρμογής, όπως την εισαγωγή, διαγραφή, ανάκτηση και τροποποίηση δεδομένων.

Δομικές μονάδες SQL II

- **Ορισμός όψεων της βάσης (View Definition):** δημιουργία εικονικών πινάκων (views -virtual tables) που περιέχουν δεδομένα από έναν ή περισσότερους πίνακες της βάσης.
- **Ορισμός εξουσιοδοτήσεων (Authorization):** δημιουργία ομάδων χρηστών και απόδοση διαφορετικών δικαιωμάτων πρόσβασης σε καθένα
- **Διαχείριση ακεραιότητας (Integrity):** επιτρέπει το λεπτομερή έλεγχο των δεδομένων που καταχωρούνται ώστε να μην παραβιάζονται οι **κανόνες** ακεραιότητας οι οποίοι, όταν τηρούνται, απομακρύνουν τον κίνδυνο **ασυνεπών δεδομένων**

Γλώσσα ορισμού δεδομένων

Η διαχείριση μέσω της DDL περιλαμβάνει τον **ορισμό** τη **μεταβολή της δομής** και τη **διαγραφή** των παρακάτω αντικειμένων:

- **Πίνακες (Tables):** το δομικό χαρακτηριστικό μιας σχεσιακής βάσης δεδομένων, καθώς περιέχουν τα δεδομένα που καταχωρούνται σε αυτή
- **Όψεις (Views):** εικονικοί πίνακες (virtual tables) οι οποίοι περιέχουν δεδομένα από έναν ή περισσότερους πίνακες της βάσης.
- **Δείκτες (Indices):** ειδικές δομές δεδομένων που επιταχύνουν τη διαδικασία της αναζήτησης πληροφοριών από τη βάση.

Εντολές ορισμού δεδομένων

- CREATE TABLE (δημιουργία πίνακα)
- CREATE INDEX (δημιουργία ευρετηρίου)
- CREATE VIEW (δημιουργία όψης)

- DROP TABLE (διαγραφή πίνακα)
- DROP INDEX (διαγραφή ευρετηρίου)
- DROP VIEW (διαγραφή όψης)

- ALTER (τροποποίηση της δομής των αντικειμένων της βάσης, δηλ, πινάκων, ευρετηρίων, όψεων)

CREATE TABLE

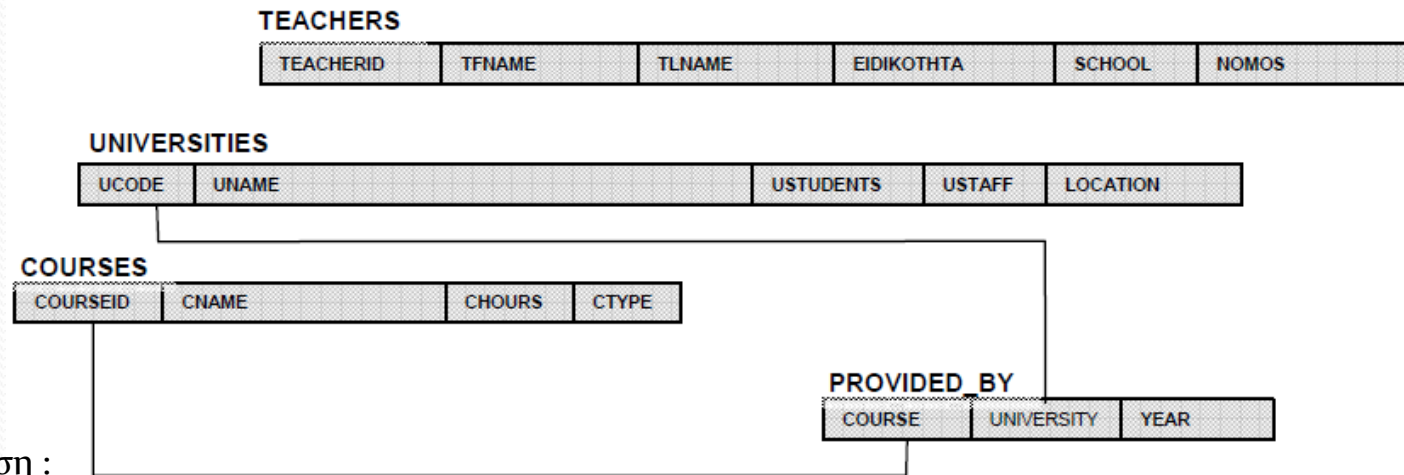
- CREATE TABLE πίνακας (όνομα_στήλης τύπος_στήλης, ...);

```
CREATE TABLE EMPLOYEE (  
  FNAME      VARCHAR (15)  NOT NULL,  
  MINIT      CHAR (1),  
  LNAME      VARCHAR (15)  NOT NULL,  
  SSN        CHAR (9)     NOT NULL,  
  BDATE      DATE,  
  ADDRESS    VARCHAR (30),  
  SEX        VARCHAR (30),  
  SALARY     INTEGER,  
  SUPERSSN   CHAR (9),  
  DNO        INTEGER ,  
  PRIMARY KEY (`SSN`) );
```

Παρατηρήσεις για CREATE

- Ο ορισμός του πρωτεύοντος κλειδιού αλλάζει ανάμεσα στις παραλλαγές της SQL
- Οι πίνακες που δημιουργούνται με την CREATE TABLE, είναι οι **βασικοί (base tables)**, διότι τα δεδομένα τους αποθηκεύονται σε αρχεία του συστήματος.
- Οι όψεις είναι **εικονικοί πίνακες (virtual tables)**, που μπορεί να αποθηκεύονται και αυτοί σε αρχείο, χωρίς όμως κάτι τέτοιο να είναι υποχρεωτικό.
- Η σειρά που αποθηκεύονται οι τιμές των πεδίων στους πίνακες, είναι ίδια με τη σειρά που έχουν δηλωθεί τα πεδία των πινάκων κατά την κλήση της εντολής.

Παράδειγμα CREATE για ΒΔ



Υλοποίηση :

```
CREATE TABLE TEACHERS(TEACHERID INT, TFNAME TEXT(50), TLNAME TEXT(50), EIDIKOTHTA  
TEXT(30), SCHOOL TEXT(10), NOMOS TEXT(50));
```

```
CREATE TABLE UNIVERSITIES(UCODE TEXT(10), UNAME TEXT(50), USTUDENTS INT, USTAFF INT,  
LOCATION TEXT(50));
```

```
CREATE TABLE COURSES(COURSEID INT, CNAME TEXT(50), CHOURS INT, CTYPE TEXT(1));
```

```
CREATE TABLE PROVIDED_BY(COURSE INT, UNIVERSITY TEXT(10), YEAR TEXT(10));
```

Παραδείγματα DROP – ALTER

DROP TABLE DEPENDENTS

Διαγράφει τον πίνακα DEPENDENTS και όλα τα δεδομένα

ALTER TABLE table_name **ADD** column_name datatype

ALTER TABLE EMPLOYEE **ADD** JOB VARCHAR(12)

προσθέτει στον πίνακα EMPLOYEE το πεδίο JOB που είναι συμβολοσειρά με μέγιστο μήκος ίσο με 12 χαρακτήρες.

ALTER TABLE table_name **DROP COLUMN** column_name

ALTER TABLE table_name **MODIFY** column_name datatype

ALTER TABLE table_name **RENAME COLUMN** old_name to new_name

Διαχείριση όψεων (Views)

- **Όψη (view):** ένας απλός πίνακας που προκύπτει από το συνδυασμό των πεδίων ενός ή περισσότερων πινάκων, οι οποίοι μπορεί να είναι είτε βασικοί πίνακες ή άλλες όψεις (εικονικοί πίνακες)
- Οι βασικοί πίνακες από τους οποίους δημιουργείται μια όψη, αναφέρονται και ως **πίνακες ορισμού της όψης (defining tables)**.
- Για να δημιουργήσουμε μια όψη χρησιμοποιούμε την εντολή **CREATE VIEW**
- Για να διαγράψουμε μια όψη χρησιμοποιούμε την εντολή **DROP VIEW**

Παραδείγματα VIEW

```
CREATE VIEW WORKS_ON1
  AS SELECT FNAME, LNAME, PNAME, HOURS
  FROM EMPLOYEE, PROJECT, WORKS_ON
  WHERE SSN = ESSN
  AND PNO = PNUMBER
```

```
CREATE VIEW DEPT_INFO (DNAME, EMP_NO, AVG_SAL)
  AS SELECT DNAME, COUNT (*), AVERAGE (SALARY)
  FROM DEPARTMENT, EMPLOYEE
  WHERE DNUMBER=DNO
  GROUP BY DNAME
```

Διαχείριση δεικτών I

- **Δείκτης (index):** μια δομή δεδομένων που αποθηκεύεται σε ειδικά αρχεία της βάσης (index files) και έχει στόχο να επιταχύνει τη διαδικασία αναζήτησης πληροφορίας από τους πίνακες της ΒΔ
- Οι δείκτες συνήθως ορίζονται για συγκεκριμένα πεδία αυτών των πινάκων (**indexing fields**) και για κάθε τιμή των εν λόγω πεδίων, αποθηκεύουν ένα σύνολο από pointers που δείχνουν τα δεδομένα της κάθε εγγραφής.

Διαχείριση δεικτών II

- Τα αρχεία δεικτών είναι αρκετά μικρότερα σε μέγεθος σε σχέση με τα αρχεία δεδομένων και η διαδικασία αναζήτησης της πληροφορίας γίνεται μέσω γνωστών τεχνικών αναζήτησης, όπως η **δυναμική αναζήτηση**

http://en.wikipedia.org/wiki/Binary_search_algorithm

- Η δημιουργία ενός **δείκτη** γίνεται χρησιμοποιώντας την εντολή **CREATE INDEX**
- Η διαγραφή ενός δείκτη γίνεται με την εντολή **DROP INDEX**

Παραδείγματα index

```
CREATE INDEX LNAME_INDEX  
ON EMPLOYEE (LNAME)
```

```
CREATE INDEX NAMES_INDEX  
ON EMPLOYEE (LNAME ASC, FNAME DESC, MINIT)
```

```
CREATE UNIQUE INDEX SSN_INDEX  
ON EMPLOYEE (SSN)
```

```
DROP INDEX LNAME_INDEX
```

Γλώσσα χειρισμού δεδομένων

- Η **γλώσσα χειρισμού δεδομένων (Data Manipulation Language, DML)**, επιτρέπει τη διαχείριση των δεδομένων των πινάκων της ΒΔ, και πιο συγκεκριμένα, την **εισαγωγή, διαγραφή, και τροποποίηση των εγγραφών** των πινάκων.
- Μέσω της DML έχουμε τη δυνατότητα να ανακτήσουμε από τους πίνακες, δεδομένα, τα οποία πληρούν κάποια συγκεκριμένα κριτήρια.
- Η DML **δεν ελέγχει** αν οι εγγραφές που επηρεάζονται ικανοποιούν τους κανόνες ακεραιότητας του λογικού σχεδιασμού που αποτυπώνονται στο σχήμα της ΒΔ

SELECT

Τελεστής Επιλογής και Προβολής:

SELECT *<attribute list>*

FROM *<table list>*

WHERE *<condition>*

συνθήκη σύζευξης (join condition) / AND, OR, NOT

```
SELECT BDATE, ADDRESS  
FROM EMPLOYEE  
WHERE FNAME='John' AND MINIT='B' AND LNAME='Smith'
```

```
SELECT FNAME, LNAME, ADDRESS  
FROM EMPLOYEE , DEPARTMENT  
WHERE DNAME = 'Research' AND  
DEPARTMENT.DNUMBER = EMPLOYEE.DNO
```

Λέξεις-κλειδιά (keywords)

- **DISTINCT**

SELECT **DISTINCT** SALARY FROM EMPLOYEE

- **UNION**

SELECT PNUMBER FROM PROJECT, DEPARTMENT, EMPLOYEE
WHERE DNUM = DNUMBER AND MGRSSN = SSN AND LNAME = 'Smith'

UNION

SELECT PNUMBER FROM PROJECT, WORKS_ON, EMPLOYEE
WHERE PNUMBER = PNO AND ESSN = SSN AND LNAME = 'Smith'

- **IN (NESTED QUERIES)**

SELECT DISTINCT PNAME FROM PROJECT
WHERE PNUMBER **IN**

(SELECT PNUMBER FROM PROJECT, DEPARTMENT, EMPLOYEE
WHERE DNUM = DNUMBER AND MGRSSN = SSN AND LNAME = 'Smith')

OR PNUMBER **IN**

(SELECT PNO FROM WORKS_ON, EMPLOYEE WHERE ESSN = SSN AND LNAME = 'Smith')

INSERT

- **INSERT INTO** *<TABLE NAME>*

VALUES (*<v₁>*, *<v₂>*, , *<v_n>*)

INSERT INTO **EMPLOYEE**

VALUES ('Richard',
 'K', 'Marini',
 '653298653',
 '30-Dec-52',
 '98 Oak Forest, Katy, TX',
 'M',
 37000,
 '987654321',
 4)

INSERT INTO **EMPLOYEE** (FNAME,LNAME, SSN)

VALUES ('Richard', 'Marini', '653298653') **NULLS!!!**

INSERT INTO **DEPTS_INFO** (DNAME, EMP_NO, TOTAL_SAL)

SELECT **DNAME**, **COUNT** (*), **SUM** (SALARY)
FROM **DEPARTMENT**, **EMPLOYEE**
WHERE **DNUMBER** = **DNO**
GROUP BY **DNAME**

DELETE

- Χρησιμοποιείται για τη **διαγραφή εγγραφών** από **ένα πίνακα** κάθε φορά.
- DELETE FROM EMPLOYEE
WHERE LNAME='Brown'
- DELETE FROM EMPLOYEE
WHERE SSN = '123456789'
- DELETE FROM EMPLOYEE
WHERE DNO
IN (SELECT DNUMBER FROM DEPARTMENT WHERE DNAME='Research')

UPDATE

- **UPDATE** *<TABLE_NAME>* **SET** *<Field_Name₁>* = '*<value₁>*', *<Field_Name₂>* = '*<value₂>*', ..., *<Field_Name_n>* = '*<value_n>*'

```
UPDATE PROJECT SET PLOCATION = 'Bellaire', DNUM = 5
WHERE PNUMBER = 10
```

```
UPDATE EMPLOYEE SET SALARY = SALARY * 1.1
WHERE DNO
IN ( SELECT DNUMBER FROM DEPARTMENT WHERE DNAME = 'Research' )
```


Περισσότερα για εξάσκηση:

- http://w3schools.com/sql/sql_tryit.asp

Οι επόμενες διαφάνειες βασίζονται
στο υλικό που συνοδεύει το βιβλίο
του Ε. Κεχρή «Σχεσιακές Βάσεις
Δεδομένων» -2^η έκδοση

Συγκεντρωτικές συναρτήσεις

- **Συγκεντρωτικές συναρτήσεις** (aggregate functions): Ειδικές συναρτήσεις της SQL για σύνθετες αριθμητικές πράξεις
- Οι συγκεντρωτικές συναρτήσεις είναι οι εξής:
 - **COUNT** : απαριθμεί το πλήθος των εγγραφών
 - **SUM** : υπολογίζει το άθροισμα μίας στήλης ή μίας αριθμητικής παράστασης μεταξύ στηλών
 - **MIN** : υπολογίζει το ελάχιστο μίας αριθμητικής στήλης
 - **MAX** : υπολογίζει το μέγιστο μίας αριθμητικής στήλης
 - **AVG** : υπολογίζει τη μέση τιμή μίας αριθμητικής στήλης

Παράδειγμα συγκεντρωτικών συναρτήσεων

ΣΥΝΤΗΡΗΣΗ

ΑριθμόςΚυκλοφορίας	ΕτήσιοΚόστος
EPH 1234	4500
YAA 2345	1200
EPH 3455	900
KZA 4567	3510
PHZ 6789	2700

```
SELECT AVG(ΕτήσιοΚόστος)  
FROM ΣΥΝΤΗΡΗΣΗ
```

AVG(ΕτήσιοΚόστος)
2562

Ο μέσος όρος το κόστους της συντήρησης

```
SELECT MIN(ΕτήσιοΚόστος)  
FROM ΣΥΝΤΗΡΗΣΗ
```

MIN(ΕτήσιοΚόστος)
900

Το ελάχιστο κόστους της συντήρησης

```
SELECT SUM(ΕτήσιοΚόστος)  
FROM ΣΥΝΤΗΡΗΣΗ
```

SUM(ΕτήσιοΚόστος)
12810

Το συνολικό κόστος συντήρησης

```
SELECT COUNT (ΑριθμόςΚυκλοφορίας)  
FROM ΣΥΝΤΗΡΗΣΗ  
WHERE ΕτήσιοΚόστος > 3000
```

COUNT(ΑριθμόςΚυκλοφορίας)
2

Το πλήθος των αυτοκινήτων με κόστος συντήρησης πάνω από 3000 €

Ομαδοποίηση εγγραφών

- Πολλές φορές είναι αντιμετωπίζουμε έναν πίνακα ως αποτελούμενο από ομάδες εγγραφών (γραμμών).
- Η ομαδοποίηση των εγγραφών επιτρέπει την εξαγωγή αποτελεσμάτων ανά ομάδα.
- Η ομαδοποίηση γίνεται εφικτή μέσω της εντολής **SELECT** με προσθήκη της λέξης-κλειδί **GROUP BY**

Παράδειγμα ομαδοποίησης I

ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ

<u>Κωδικός Πωλητή</u>	<u>Κωδικός Προϊόντος</u>	<u>Πωλήσεις</u>
1	1001	55
1	1002	60
1	1003	25
2	1001	70
2	1002	15

- Ο πίνακας **ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ** δείχνει τις πωλήσεις σε κάθε προϊόν που έκαναν οι πωλητές μίας εταιρείας στη διάρκεια του προηγούμενου μήνα.
- Θεωρούμε ότι ο πίνακας αυτός αποτελείται από δύο ομάδες εγγραφών:
 - α' ομάδα: πωλήσεις του πωλητή με κωδικό 1 (**ομάδα A**), ενώ
 - β' ομάδα: πωλήσεις του πωλητή με κωδικό 2 (**ομάδα B**).
- Στην περίπτωση αυτή ο χωρισμός σε ομάδες γίνεται ανάλογα με την τιμή που έχει η στήλη **Κωδικός Πωλητή**.

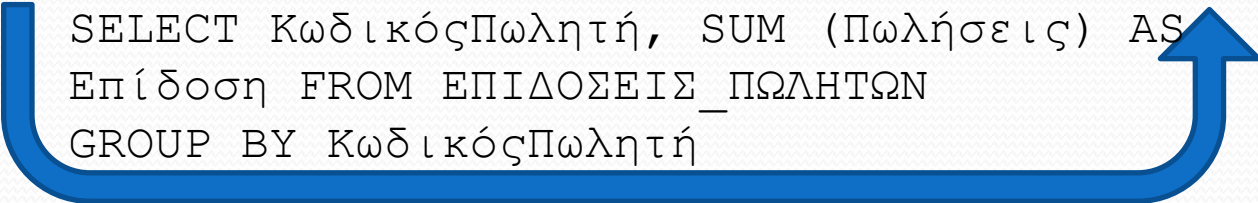
Ομαδοποίηση με το Group by

ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ

<u>Κωδικός Πωλητή</u>	<u>Κωδικός Προϊόντος</u>	<u>Πωλήσεις</u>
1	1001	55
1	1002	60
1	1003	25
2	1001	70
2	1002	15

<u>Κωδικός Πωλητή</u>	<u>Επίδοση</u>
1	130
2	85

```
SELECT ΚωδικόςΠωλητή, SUM (Πωλήσεις) AS  
Επίδοση FROM ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ  
GROUP BY ΚωδικόςΠωλητή
```



- Στην SQL είναι δυνατό οι συγκεντρωτικές συναρτήσεις να υπολογιστούν για κάθε ομάδα χωριστά.
- Αυτό γίνεται με χρήση των δεσμευμένων λέξεων **GROUP BY** με τη βοήθεια των οποίων δηλώνονται οι διάφορες ομάδες του πίνακα.
- Έτσι, το σύνολο των πωλήσεων για κάθε πωλητή χωριστά υπολογίζεται από την παραπάνω εντολή

Παράδειγμα ομαδοποίησης II

ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ

<u>Κωδικός Πωλητή</u>	<u>Κωδικός Προϊόντος</u>	<u>Πωλήσεις</u>
1	1001	55
1	1002	60
1	1003	25
2	1001	70
2	1002	15

- Οι εγγραφές ενός πίνακα μπορούν να ομαδοποιηθούν με διάφορους τρόπους, ανάλογα με τον υπολογισμό που απαιτείται.
- Έτσι, πίνακας **ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ** μπορεί να θεωρηθεί ότι αποτελείται από τρεις ομάδες εγγραφών:
 - η α' ομάδα αφορά τις πωλήσεις του προϊόντος 1001 (**ομάδα 1**),
 - η β' ομάδα αφορά τις πωλήσεις του προϊόντος 1002 (**ομάδα 2**) και
 - η γ' ομάδα αφορά τις πωλήσεις του προϊόντος 1003 (**ομάδα 3**).

Παράδειγμα ομαδοποίησης II

ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ

<u>Κωδικός Πωλητή</u>	Κωδικός Προϊόντος	Πωλήσεις
1	1001	55
1	1002	60
1	1003	25
2	1001	70
2	1002	15

Κωδικός Προϊόντος	ΠωλήσειςΠΡ
1001	125
1002	75
1003	25

```
SELECT ΚωδικόςΠροϊόντος, SUM (Πωλήσεις) AS  
ΠωλήσειςΠΡ FROM ΕΠΙΔΟΣΕΙΣ_ΠΩΛΗΤΩΝ  
GROUP BY ΚωδικόςΠροϊόντος
```



- Τώρα, η ομαδοποίηση γίνεται ανάλογα με την τιμή που έχει η στήλη **Κωδικός Προϊόντος**.
- Αυτή η ομαδοποίηση είναι απαραίτητη όταν θέλουμε να υπολογίσουμε το σύνολο των πωλήσεων για κάθε προϊόν.
- Η εντολή SQL που δίνει τον υπολογισμό αυτόν φαίνεται παραπάνω

Επιλογή με περισσότερους πίνακες

- Μέχρι τώρα είδαμε πώς μπορούμε να αναζητούμε με τη βοήθεια της SQL δεδομένα από ένα μόνο πίνακα.
- Ορισμένες φορές, ωστόσο, είναι απαραίτητο να αναζητήσουμε δεδομένα από δύο ή περισσότερους πίνακες, όπως εξηγείται στο επόμενο παράδειγμα.

Συνδυασμός δύο πινάκων

ΠΡΟΪΟΝΤΑ

Εγγραφή	<u>Κωδικός Προϊόντος</u>	<u>Περιγραφή</u>	<u>Τιμή Αγοράς</u>	<u>Κωδικός Προμηθευτή</u>	<u>Κατηγορία</u>
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	<u>Κωδικός Κατηγορίας</u>	<u>Ονομασία</u>
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής

Μπορούμε να συνδυάσουμε την εγγραφή π1 με την κ1 και να δημιουργήσουμε την εγγραφή:

εγγραφή	<u>Κωδικός Προϊόντος</u>	<u>Περιγραφή</u>	<u>Τιμή Αγοράς</u>	<u>Κωδικός Προμηθευτή</u>	<u>Κατηγορία</u>	<u>Κωδικός Κατηγορίας</u>	<u>Ονομασία</u>
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου

Με την ίδια λογική μπορούμε να συνδυάσουμε την εγγραφή π1 με την κ2 και να δημιουργήσουμε την εγγραφή:

εγγραφή	<u>Κωδικός Προϊόντος</u>	<u>Περιγραφή</u>	<u>Τιμή Αγοράς</u>	<u>Κωδικός Προμηθευτή</u>	<u>Κατηγορία</u>	<u>Κωδικός Κατηγορίας</u>	<u>Ονομασία</u>
πκ2: π1 με κ2	1001	Στυλό	0.75	Π200	1	2	Είδη υπολογιστή

Καρτεσιανό γινόμενο

Συνολικά μπορούμε να συνδυάσουμε κάθε μία από τις 4 εγγραφές του α' πίνακα με κάθε μία από τις 3 εγγραφές του β' πίνακα και έτσι να δημιουργήσουμε τις παρακάτω 12 εγγραφές:

εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Ονομασία
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
πκ2: π1 με κ2	1001	Στυλό	0.75	Π200	1	2	Είδη υπολογιστή
πκ3: π1 με κ3	1001	Στυλό	0.75	Π200	1	3	Υπολογιστής
πκ4: π2 με κ1	1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
πκ5: π2 με κ2	1002	Μολύβι	1	Π250	1	2	Είδη υπολογιστή
πκ6: π2 με κ3	1002	Μολύβι	1	Π250	1	3	Υπολογιστής
πκ7: π3 με κ1	1003	Μαρκαδόρος	1.3	Π130	5	1	Είδη γραφείου
πκ8: π3 με κ2	1003	Μαρκαδόρος	1.3	Π130	5	2	Είδη υπολογιστή
πκ9: π3 με κ3	1003	Μαρκαδόρος	1.3	Π130	5	3	Υπολογιστής
πκ10: π4 με κ1	1004	Δισκέτα 3,5"	0.8	Π120	2	1	Είδη γραφείου
πκ11: π4 με κ2	1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή
πκ12: π4 με κ3	1004	Δισκέτα 3,5"	0.8	Π120	2	3	Υπολογιστής

Καρτεσιανό γινόμενο στην SQL

Με τη βοήθεια της SQL μπορούμε να εμφανίσουμε το καρτεσιανό γινόμενο των πινάκων **ΠΡΟΪΟΝΤΑ**, **ΚΑΤΗΓΟΡΙΕΣ** με την εντολή:

```
SELECT ΠΡΟΪΟΝΤΑ.* , ΚΑΤΗΓΟΡΙΕΣ.*  
FROM ΠΡΟΪΟΝΤΑ, ΚΑΤΗΓΟΡΙΕΣ
```

Καρτεσιανό γινόμενο και μη αντιστοιχισμένες εγγραφές

- Η εγγραφή πκ2 συνδυάζει το προϊόν 1001 που είναι στην κατηγορία 1 με τα είδη υπολογιστή που έχουν κωδικό κατηγορίας 2. Επομένως για τη εγγραφή πκ2 δεν υπάρχει αντιστοιχία (μη αντιστοιχισμένη εγγραφή).
- Για τον ίδιο λόγο μη αντιστοιχισμένες είναι οι εγγραφές που εμφανίζονται με κόκκινα γράμματα.

εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Ονομασία
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
πκ2: π1 με κ2	1001	Στυλό	0.75	Π200	1	2	Είδη υπολογιστή
πκ3: π1 με κ3	1001	Στυλό	0.75	Π200	1	3	Υπολογιστής
πκ4: π2 με κ1	1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
πκ5: π2 με κ2	1002	Μολύβι	1	Π250	1	2	Είδη υπολογιστή
πκ6: π2 με κ3	1002	Μολύβι	1	Π250	1	3	Υπολογιστής
πκ7: π3 με κ1	1003	Μαρκαδόρος	1.3	Π130	5	1	Είδη γραφείου
πκ8: π3 με κ2	1003	Μαρκαδόρος	1.3	Π130	5	2	Είδη υπολογιστή
πκ9: π3 με κ3	1003	Μαρκαδόρος	1.3	Π130	5	3	Υπολογιστής
πκ10: π4 με κ1	1004	Δισκέτα 3,5"	0.8	Π120	2	1	Είδη γραφείου
πκ11: π4 με κ2	1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή
πκ12: π4 με κ3	1004	Δισκέτα 3,5"	0.8	Π120	2	3	Υπολογιστής

Αντιστοιχισμένες εγγραφές

- **Αντιστοιχισμένες εγγραφές** : αυτές που έχουν κάποιο κοινό πεδίο, βάση του οποίου μπορεί να γίνει αντιστοίχιση.
- Για να κρατήσουμε τις αντιστοιχισμένες εγγραφές του καρτεσιανού γινομένου χρησιμοποιούμε τον όρο **WHERE** και περιγράφουμε τη συνθήκη που θέλουμε να ικανοποιούν τα δεδομένα που εμφανίζονται (ΚΑΤΗΓΟΡΗΜΑ ΣΥΖΕΥΞΗΣ).
- Στο συγκεκριμένο παράδειγμα, οι γραμμές που θέλουμε να εμφανίζονται είναι αυτές στις οποίες κάθε προϊόν συνδυάζεται μόνο με την δική του κατηγορία:

```
SELECT ΠΡΟΪΟΝΤΑ.*, ΚΑΤΗΓΟΡΙΕΣ.*  
FROM ΠΡΟΪΟΝΤΑ, ΚΑΤΗΓΟΡΙΕΣ  
WHERE Κατηγορία = ΚωδικόςΚατηγορίας
```

Παράδειγμα αντιστοιχισμένων εγγραφών

Η εντολή:

```
SELECT ΠΡΟΪΟΝΤΑ.*, ΚΑΤΗΓΟΡΙΕΣ.*  
FROM ΠΡΟΪΟΝΤΑ, ΚΑΤΗΓΟΡΙΕΣ  
WHERE Κατηγορία = ΚωδικόςΚατηγορίας
```

Εμφανίζει τα παρακάτω δεδομένα:

εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Ονομασία
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
πκ4: π2 με κ1	1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
πκ11: π4 με κ2	1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή

- Στον προηγούμενο πίνακα η κατηγορία ενός προϊόντος εμφανίζεται δύο φορές: μία φορά στη στήλη Κατηγορία και μία φορά στη στήλη ΚωδικόςΚατηγορίας.
- Μπορούμε να εμφανίσουμε κάθε στήλη μία μόνο φορά χρησιμοποιώντας για παράδειγμα την παρακάτω εντολή:

```
SELECT ΠΡΟΪΟΝΤΑ.*, ΚΑΤΗΓΟΡΙΕΣ.Ονομασία  
FROM ΠΡΟΪΟΝΤΑ, ΚΑΤΗΓΟΡΙΕΣ  
WHERE Κατηγορία = ΚωδικόςΚατηγορίας
```

Πρακτική οδηγία

Όταν θέλουμε να συνδυάσουμε δύο ή περισσότερους πίνακες:

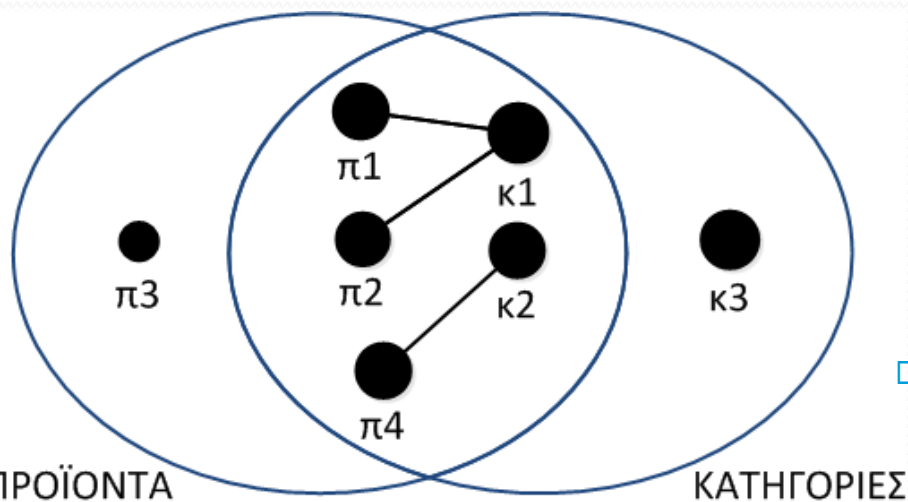
- στον όρο **SELECT** γράφουμε τα ονόματα των στηλών που θέλουμε να εμφανίζονται στο αποτέλεσμα
- στον όρο **FROM** αναφέρουμε τα ονόματα των πινάκων από τους οποίους θέλουμε να αντλήσουμε δεδομένα
- στον όρο **WHERE** περιγράφουμε με ποιόν τρόπο συνδέονται οι πίνακες από τους οποίους αντλούμε δεδομένα.

ΠΡΟΪΟΝΤΑ

Εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	Κωδικός Κατηγορίας	Ονομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



- Στη γραφική παράσταση Venn, οι εγγραφές των δύο πινάκων που συσχετίζονται αναπαριστούνται ως στοιχεία τα οποία
 - ▣ συνδέονται άμεσα με μία γραμμή
 - ▣ τοποθετούνται κεντρικά ώστε να ανήκουν και στα δύο σύνολα
- Οι εγγραφές πίνακα που δεν συσχετίζονται με άλλη εγγραφή, αναπαριστούνται ως στοιχεία τα οποία τοποθετούνται περιφερειακά στο σύνολο

Ο όρος JOIN της SQL

- Η SQL προσφέρει την δεσμευμένη λέξη **JOIN** η οποία μπορεί να χρησιμοποιηθεί για να συνδυάσουμε δεδομένα δύο ή περισσότερων πινάκων.
- Στην απλούστερη μορφή του ο όρος **JOIN** χρησιμοποιείται σε μία εντολή **SELECT** για να βρούμε τις εγγραφές δύο ή περισσότερων πινάκων που σε μία στήλη τους έχουν την ίδια τιμή.

Εσωτερική σύζευξη

- Η εσωτερική σύζευξη (inner join) ανάμεσα σε δύο πίνακες που έχουν μία κοινή στήλη είναι η πράξη με την οποία εμφανίζονται οι εγγραφές που στην κοινή τους στήλη έχουν την ίδια τιμή.

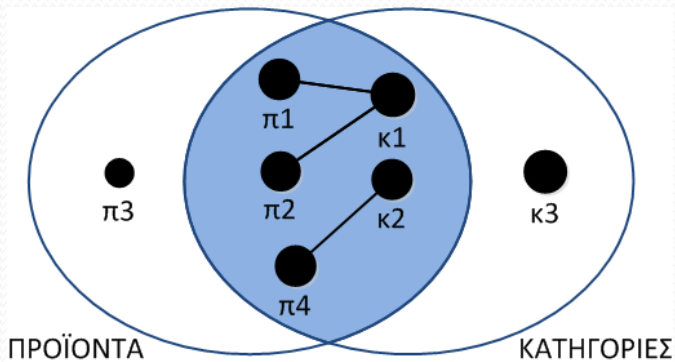
Παράδειγμα εσωτερικής σύζευξης πινάκων με τον όρο JOIN

ΠΡΟΪΟΝΤΑ

Εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	Κωδικός Κατηγορίας	Ονομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



SELECT *
FROM ΠΡΟΪΟΝΤΑ JOIN ΚΑΤΗΓΟΡΙΕΣ
ON Κατηγορία =
ΚωδικόςΚατηγορίας

	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Ονομασία
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
πκ4: π2 με κ1	1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
πκ11: π4 με κ2	1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή

Αριστερή σύζευξη

- Η αριστερή σύζευξη (left join) ανάμεσα σε δύο πίνακες που έχουν μία κοινή στήλη είναι η πράξη με την οποία εμφανίζονται **όλες** τις εγγραφές του αριστερού (δηλαδή του πρώτου) πίνακα.
- Δηλαδή η αριστερή σύζευξη περιλαμβάνει:
 - τις εγγραφές των δύο πινάκων που στη κοινή στήλη τους έχουν την ίδια τιμή καθώς και
 - τις εγγραφές του αριστερού (δηλαδή του πρώτου στη σειρά) πίνακα τα οποία στην κοινή στήλη έχουν τιμή που δεν υπάρχει στον δεξιό (δεύτερο) πίνακα.

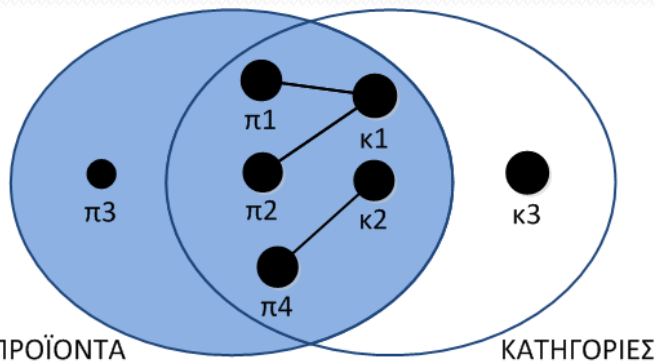
Παράδειγμα αριστερής σύζευξης πινάκων με τον όρο LEFT JOIN

ΠΡΟΪΟΝΤΑ

Εγγραφή	<u>Κωδικός Προϊόντος</u>	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	<u>Κωδικός Κατηγορίας</u>	Ονομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



SELECT *
FROM ΠΡΟΪΟΝΤΑ LEFT JOIN
ΚΑΤΗΓΟΡΙΕΣ
ON Κατηγορία = ΚωδικόςΚατηγορίας

πκ1: π1 με κ1
 πκ4: π2 με κ1
 πκ11: π4 με κ2

Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Ονομασία
1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή
1003	Μαρκαδόρος	1.3	Π130	5	NULL	NULL

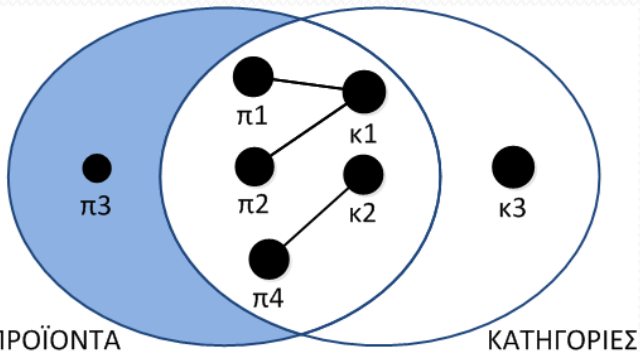
Παράδειγμα αριστερής σύζευξης πινάκων για μη σχετιζόμενες εγγραφές

ΠΡΟΪΟΝΤΑ

Εγγραφή	<u>Κωδικός Προϊόντος</u>	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	<u>Κωδικός Κατηγορίας</u>	Όνομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



```
SELECT *
FROM ΠΡΟΪΟΝΤΑ LEFT JOIN
ΚΑΤΗΓΟΡΙΕΣ
ON Κατηγορία = ΚωδικόςΚατηγορίας
WHERE ΚωδικόςΚατηγορίας IS NULL
```

Συνδυασμός
εγγραφών
Κ3 με καμία

Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Όνομασία
1003	Μαρκαδόρος	1.3	Π130	5	NULL	NULL

Δεξιά σύζευξη

- Η **δεξιά σύζευξη** (right join) ανάμεσα σε δύο πίνακες που έχουν μία κοινή στήλη είναι η πράξη με την οποία εμφανίζονται **όλες** τις εγγραφές του δεξιού (δηλαδή του δεύτερου) πίνακα.
- Η δεξιά σύζευξη περιλαμβάνει:
 - τις εγγραφές των δύο πινάκων που στη κοινή στήλη τους έχουν την ίδια τιμή καθώς και
 - τις εγγραφές του δεξιού (δηλαδή του δεύτερου στη σειρά) πίνακα οι οποίες στην κοινή στήλη έχουν τιμή που δεν υπάρχει στον αριστερό (πρώτο) πίνακα.

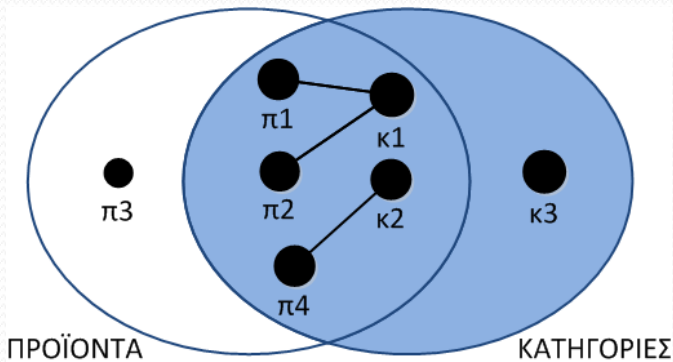
Παράδειγμα δεξιάς σύζευξης πινάκων με τον όρο RIGHT JOIN

ΠΡΟΪΟΝΤΑ

Εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	Κωδικός Κατηγορίας	Ονομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



SELECT *
FROM ΠΡΟΪΟΝΤΑ RIGHT JOIN
ΚΑΤΗΓΟΡΙΕΣ
ON Κατηγορία =
ΚωδικόςΚατηγορίας

	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Ονομασία
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
πκ4: π2 με κ1	1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
πκ11: π4 με κ2	1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή
	NULL	NULL	NULL	NULL	NULL	3	Υπολογιστή

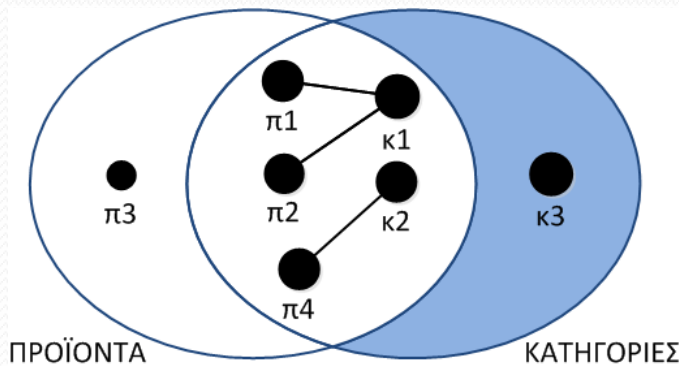
Παράδειγμα δεξιάς σύζευξης πινάκων για τις μη σχετιζόμενες εγγραφές

ΠΡΟΪΟΝΤΑ

Εγγραφή	<u>Κωδικός Προϊόντος</u>	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	<u>Κωδικός Κατηγορίας</u>	Όνομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



```
SELECT *
FROM ΠΡΟΪΟΝΤΑ RIGHT JOIN
ΚΑΤΗΓΟΡΙΕΣ
ON Κατηγορία = ΚωδικόςΚατηγορίας
WHERE Κατηγορία IS NULL
```

Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Όνομασία
NULL	NULL	NULL	NULL	NULL	3	Υπολογιστή

Πλήρης σύζευξη

Η πλήρης σύζευξη ανάμεσα σε δύο πίνακες που έχουν μία κοινή στήλη είναι η πράξη με την οποία εμφανίζονται όλα τα στοιχεία όλων των πινάκων:

- Εγγραφές που συνδυάζονται
- Εγγραφές του ενός πίνακα που δεν συνδέονται με καμία εγγραφή του άλλου. Τότε οι εγγραφές του άλλου πίνακα εμφανίζουν τιμές **NULL** στα πεδία που δεν υπάρχει αντιστοίχιση.

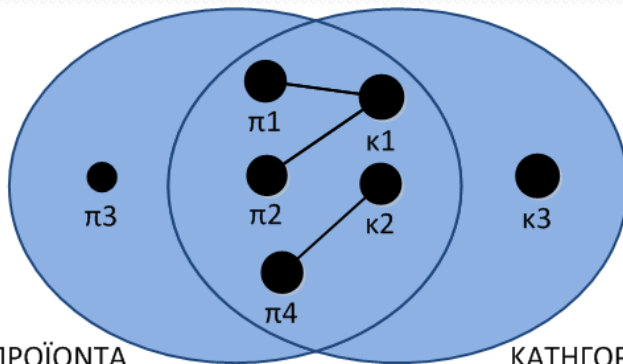
Παράδειγμα πλήρους σύζευξης πινάκων όταν υποστηρίζεται ο όρος FULL JOIN

ΠΡΟΪΟΝΤΑ

Εγγραφή	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία
π1	1001	Στυλό	0.75	Π200	1
π2	1002	Μολύβι	1	Π250	1
π3	1003	Μαρκαδόρος	1.3	Π130	5
π4	1004	Δισκέτα 3,5"	0.8	Π120	2

ΚΑΤΗΓΟΡΙΕΣ

Εγγραφή	Κωδικός Κατηγορίας	Όνομασία
κ1	1	Είδη γραφείου
κ2	2	Είδη υπολογιστή
κ3	3	Υπολογιστής



SELECT *

FROM ΠΡΟΪΟΝΤΑ FULL OUTER JOIN
ΚΑΤΗΓΟΡΙΕΣ

ON Κατηγορία = ΚωδικόςΚατηγορίας

	Κωδικός Προϊόντος	Περιγραφή	Τιμή Αγοράς	Κωδικός Προμηθευτή	Κατηγορία	Κωδικός Κατηγορίας	Όνομασία
πκ1: π1 με κ1	1001	Στυλό	0.75	Π200	1	1	Είδη γραφείου
πκ4: π2 με κ1	1002	Μολύβι	1	Π250	1	1	Είδη γραφείου
πκ11: π4 με κ2	1004	Δισκέτα 3,5"	0.8	Π120	2	2	Είδη υπολογιστή
	1003	Μαρκαδόρος	1.3	Π130	5	NULL	NULL
	NULL	NULL	NULL	NULL	NULL	3	Υπολογιστή

Πλήρης σύζευξη στην SQL όταν δεν υποστηρίζεται ο όρος FULL JOIN

- Σε συστήματα που δεν υποστηρίζεται η δεσμευμένη λέξη **FULL** χρησιμοποιείται η εντολή **UNION** :

```
SELECT * FROM ΠΡΟΪΟΝΤΑ LEFT JOIN ΚΑΤΗΓΟΡΙΕΣ  
ON Κατηγορία = ΚωδικόςΚατηγορίας  
UNION
```

```
SELECT * FROM ΠΡΟΪΟΝΤΑ RIGHT JOIN ΚΑΤΗΓΟΡΙΕΣ  
ON Κατηγορία = ΚωδικόςΚατηγορίας
```

LibreOffice Base

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Βάσεις Δεδομένων σε περιβάλλοντα εφαρμογών γραφείου

- Συνδυασμός ενός ΣΔΒΔ με μια διεπαφή χρήστη, που, ως εφαρμογή, εντάσσεται σε ένα περιβάλλον εφαρμογών γραφείου.
- Παραδείγματα:
 - **Access** στο Microsoft Office 2013
 - **Base** στο LibreOffice/OpenOffice



Τι είναι η LibreOffice Base;

- Η **LibreOffice Base (LO Base)** είναι μια εφαρμογή που συνδέεται διαφανώς με ένα ΣΔΒΔ. Μέχρι σήμερα (2015) η LO Base συνδέεται με το HyperSQL, που όμως προβλέπεται να αντικατασταθεί από το Firebird.
- Η LO Base διαθέτει χαρακτηριστικά που την καθιστούν φιλική προς το χρήστη και εξασφαλίζουν τη διαλειτουργικότητά της με τις υπόλοιπες εφαρμογές του Libre Office.
- Μπορεί να διαχειριστεί διαφορετικές «πηγές δεδομένων» (data sources): π.χ. βάση από MySQL ή MS Access, φύλλο υπολογισμών, αρχείο κειμένου κλπ.
- Μπορεί να χρησιμοποιηθεί σε όλα τα λειτουργικά συστήματα που εγκαθίσταται το LibreOffice (MS Windows, Linux, MacOS)
- Διανέμεται ΔΩΡΕΑΝ

HyperSQL

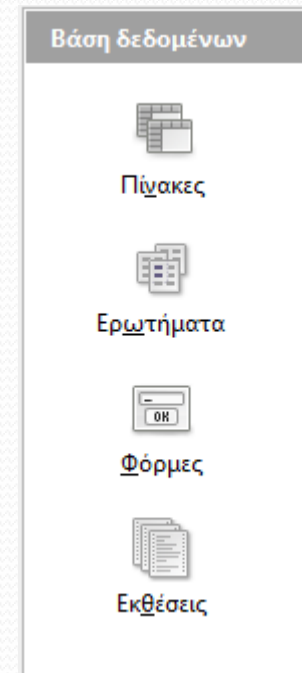
- Το **HyperSQL** ή HSQL είναι ένα σχεσιακό ΣΔΒΔ γραμμένο σε java, άρα ανεξάρτητο από λειτουργικά συστήματα.
- Είναι το υποκείμενο RDBMS, στο οποίο βασίζεται η LO Base μέχρι σήμερα, αλλά βρίσκεται ενσωματωμένο και σε εμπορικές εφαρμογές όπως το Mathematica
- Είναι ανοικτό λογισμικό.

Firebird

- Το Firebird είναι ένα σχεσιακό ΣΔΒΔ, με δυνατότητες υποστήριξης της SQL που τρέχει σε Linux, Windows, και μια πληθώρα συστημάτων Unix.
- Έχει χρησιμοποιηθεί σε συστήματα παραγωγής, κάτω από μια ποικιλία ονομάτων, από το 1981 και πρόκειται να υποστηρίξει τις επόμενες εκδόσεις της LO Base,
- Το Firebird Project υλοποιείται σε C και C++, που έχουν πολύ καλές επιδόσεις, και υποστηρίζεται από κοινότητα προγραμματιστών.
- Ο αρχικός πηγαίος κώδικάς προήλθε από την Inprise Corp (γνωστή σήμερα ως Borland Software Corp) στις 25 Ιουλίου, 2000.

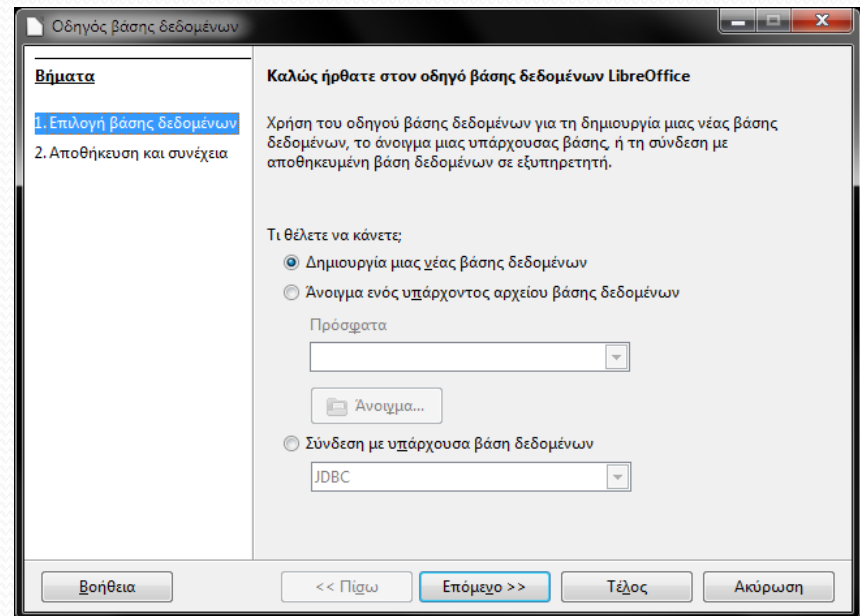
Τρόπος δόμησης στην LO Base

- Οι βάσεις δεδομένων της Base αποτελούνται από τα εξής αντικείμενα:
 - Οι **Πίνακες (Tables)** αποθηκεύουν τα δεδομένα σας σε γραμμές και στήλες.
 - Τα **Ερωτήματα (Queries)** ανακτούν και επεξεργάζονται τα δεδομένα σας.
 - Οι **Φόρμες (Forms)** ελέγχουν την εισαγωγή δεδομένων και τις προβολές δεδομένων.
 - Οι **Εκθέσεις (Reports)** συνοψίζουν και εκτυπώνουν τα δεδομένα.

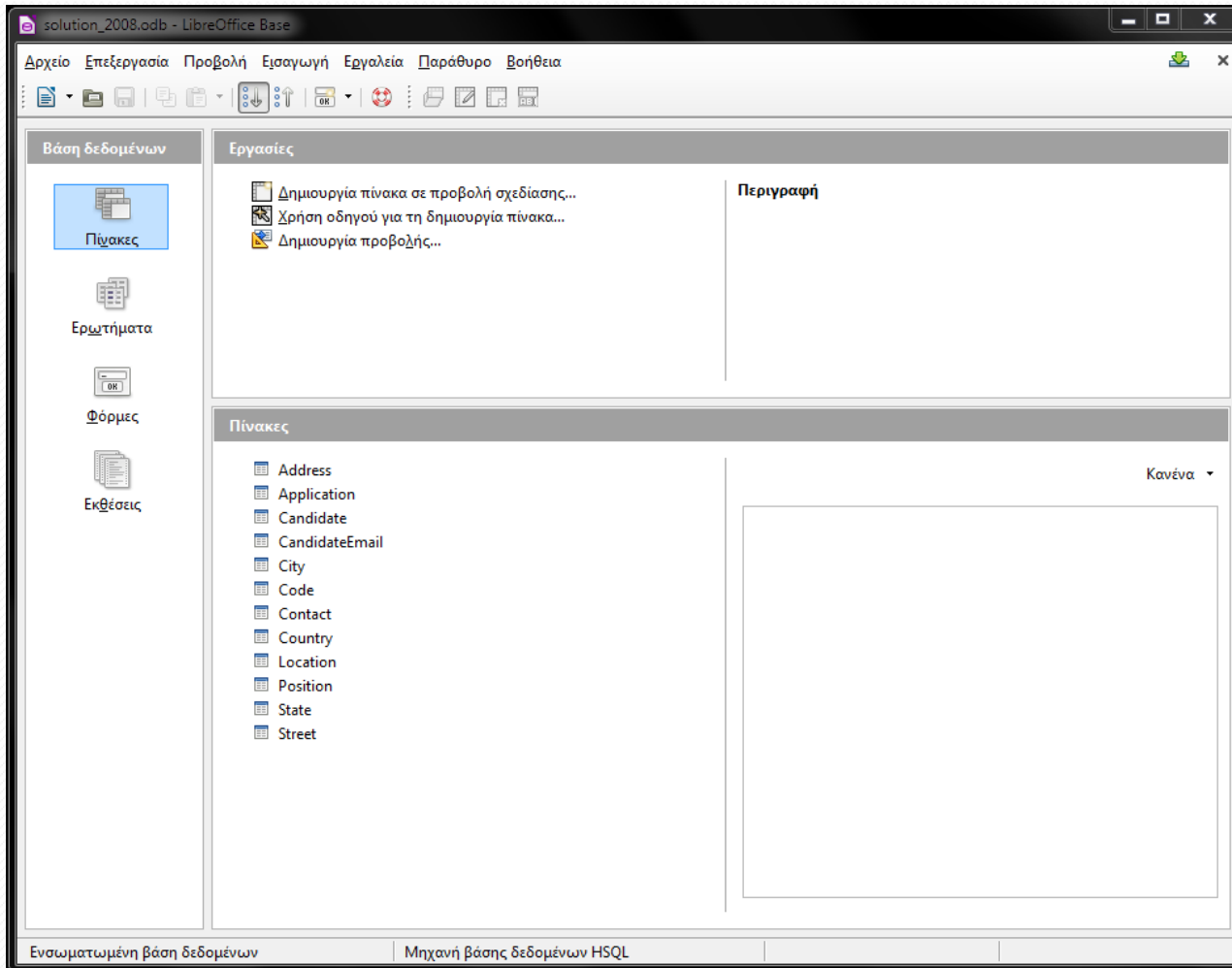


Έναρξη της εφαρμογής

- Κατά την έναρξη εμφανίζεται ο «Οδηγός βάσης δεδομένων»
- Μπορούμε
 - να δημιουργήσουμε μια νέα βάση
 - να ανοίξουμε μια υφιστάμενη βάση που φιλοξενείται στον τοπικό δίσκο
 - να συνδεθούμε με μια υφιστάμενη βάση σε ένα άλλο RDBMS, και
 - να προσδιορίσουμε τον τρόπο αποθήκευσης των δεδομένων μας



Περιβάλλον LO Base

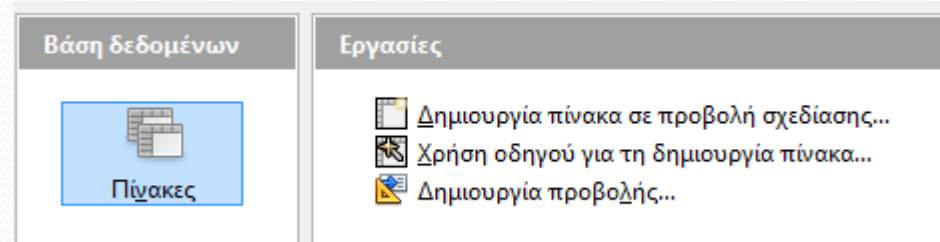


Τρόποι Εργασιών

- Οι δύο πιο παραγωγικοί τρόποι εργασιών με τα αντικείμενα της Base (*Πίνακες, Ερωτήματα, Φόρμες, Εκθέσεις*) είναι:
 - Η **προβολή σχεδίασης**: παρέχει ένα εποπτικό σχεδιαστικό περιβάλλον για να υλοποιηθεί η εργασία. Η προβολή σχεδίασης είναι προσαρμοσμένη σε κάθε εργασία.
 - Ο **οδηγός δημιουργίας (wizard)**: η εργασία υλοποιείται μέσω μιας σειράς βημάτων. Σε κάθε βήμα ο χρήστης κάνει μια σειρά επιλογών που τελικά καθορίζουν το επιθυμητό αποτέλεσμα.

Πίνακες

- Η δημιουργία πινάκων μπορεί να γίνει με τρεις τρόπους:
 - Μέσω της **προβολής σχεδίασης**
 - Με χρήση **οδηγού δημιουργίας**
 - Με τη **δημιουργία προβολής**, στην οποία χρησιμοποιούνται υφιστάμενοι πίνακες (δημιουργία «όψεων»)



Ερωτήματα

- Η LO Base προσφέρει τρεις βασικούς τρόπους για τη δημιουργία ερωτημάτων (queries):
 - **Προβολή σχεδίασης:** δίνει πλήρη έλεγχο κατά τη δημιουργία του ερωτήματος. Drag'n'drop επιλεγμένων πεδίων στο πλέγμα και καταχώρηση κριτηρίων για την επιλογή των δεδομένων.
 - **Οδηγός απλών ερωτημάτων:** στο ξεκίνημα όταν ορίζεται η δομή του ερωτήματος, εκτελούνται αυτόματα κάποιες εργασίες.
 - **Προβολή SQL :** Απ' ευθείας σύνταξη εντολών SQL

Φόρμες

- Οι Φόρμες χρησιμοποιούνται για εισαγωγή και προβολή δεδομένων από μια βάση.
- Οι Φόρμες δημιουργούνται με δύο τρόπους :
 - Μέσω της **προβολής σχεδίασης**
 - Με χρήση **οδηγού δημιουργίας**

Εκθέσεις

- Οι Εκθέσεις παράγονται από Πίνακες ή/και Ερωτήματα
- Προβάλλουν συγκεντρωτικά στοιχεία από τη βάση
- Είναι είτε στατικές, είτε δυναμικές.

- Οι Εκθέσεις δημιουργούνται με δύο τρόπους :
 - Μέσω της **προβολής σχεδίασης**
 - Με χρήση **οδηγού δημιουργίας**

Εκμάθηση του LibreOffice Base

- <http://www.youtube.com/playlist?list=PLy7Kah3WzqrEerJoVPNWVaR4CYHMr4wmV>
- <https://wiki.documentfoundation.org/images/e/e8/BH4o-BaseHandbook.pdf>