



EPL342 –Databases

Lecture 24: Functional Dependencies and Normalization and Introduction to EPL446/EPL646 (Chapter 14.5, Elmasri-Navathe 7ED)

Demetris Zeinalipour

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

Περιεχόμενο Διάλεξης



- **Ολοκλήρωση Διάλεξης 23**
- **Κεφάλαιο 10.5: SQL DML III**
 - Η Κανονική Μορφή Boyce-Codd (Boyce-Codd Normal Form (BCNF))
- **Εισαγωγή στο μάθημα ΕΠΛ446:
Προχωρημένες Βάσεις Δεδομένων**

Boyce-Codd Κανονική Μορφή (Εισαγωγή)



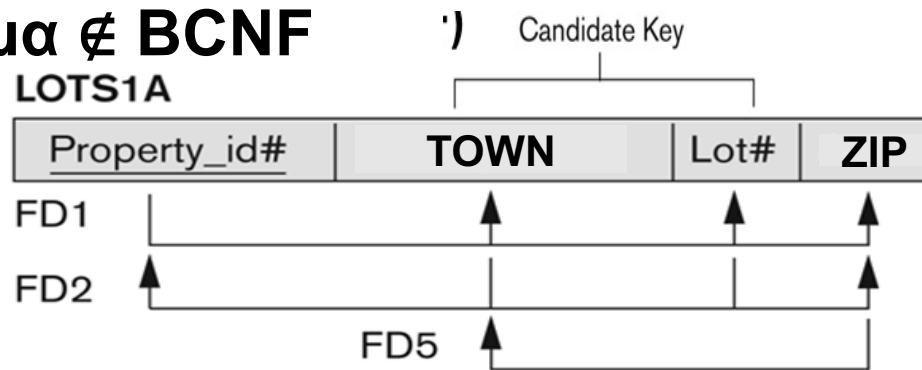
- Ο **Raymond F. Boyce**, ήταν ερευνητής της **IBM** όπου με άλλους συναδέλφους **Donald D. Chamberlin**, και κάτω από την καθοδήγηση του **Edgar F. Codd** (του δημιουργού του Σχεσιακού Μοντέλου), **σύγγραψε το 1974** την πρώτη έκδοση του προτύπου **SQL** (της **SEQUEL**).
 - *“SEQUEL: A Structured English Query Language” D.D. Chamberlin and R.F. Boyce, Proc. ACM SIGMOD Workshop on Data Description, Access and Control, Ann Arbor, Michigan (May 1974) pages 249-264.*
- Μεταξύ άλλων, **υπέδειξε** επίσης με τον Codd πως μπορεί να **βελτιστοποιηθεί** η **3NF**, σε κάποιες ειδικές περιπτώσεις, **πετυχαίνονται** ακόμη **χαμηλότερα επίπεδα πλεονασμού (redundancy)** σε ένα σχεσιακό σχήμα.

Boyce-Codd Κανονική Μορφή (BCNF: Boyce Codd Normal Form)



- **Ορισμός BCNF:** Ένα σχήμα R είναι σε **BCNF** εάν για κάθε μη-τετριμμένη (τετριμμένη: $X \rightarrow Y, X \supseteq Y$) FD $X \rightarrow A$ στο R, το X είναι υπερκλειδί της R

- **Παράδειγμα \notin BCNF**



Το ZIP δεν είναι Superkey! άρα δεν είναι BCNF (είναι όμως 3NF γιατί δεν υπάρχουν non-key σε non-key εξαρτήσεις)

- **Γιατί οι \notin BCNF σχέσεις έχουν πρόβλημα;**

– Γιατί εάν το X δεν είναι superkey, τότε δημιουργείται πλεονασμός

Redundancy!

Property_id	Town	Lot#	ZIP
1	Strovolos	1	2330
2	Lakatamia	2	2550
3	Strovolos	5	2330

Π.χ., πλεονασμός από ZIP → TOWN

Boyce-Codd Κανονική Μορφή (BCNF: Boyce Codd Normal Form)




- Εάν στο προηγούμενο παράδειγμα το **ZIP** ήταν **κλειδί (ή υπερκλειδί)** τότε δεν θα μπορούσαμε να είχαμε δυο **πλειάδες** με τιμή **2330**
- ... συνεπώς θα **αποφεύγαμε** τον πλεονασμό που πρόκυπτε πριν λόγω του **ZIP→TOWN**. **∈ BCNF**

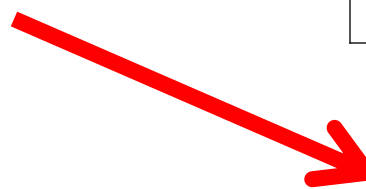
<u>Property_id</u>	Town	Lot#	ZIP
1	Strovolos	1	2330
2	Lakatamia	2	2550
3	Strovolos	5	2330

∉ BCNF

Διάσπαση



<u>Property_id</u>	Lot#	ZIP
1	1	2330
2	2	2550
3	5	2330



<u>ZIP</u>	Town
2330	Strovolos
2550	Lakatamia

∈ BCNF

Boyce-Codd Κανονική Μορφή (BCNF: Boyce Codd Normal Form)



Παράδειγμα Σχέσης $\in 3NF$ αλλά $\notin BCNF$

{Student, Course} \rightarrow Instructor
Instructor \rightarrow Course

TEACH

<u>Student</u>	<u>Course</u>	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar



Γιατί $R \in 3NF$;

Διότι δεν υπάρχουν μεταβατικές εξαρτήσεις από το κλειδί.

Γιατί $R \notin BCNF$;
Διότι ο "Instructor" (C)

ΔΕΝ είναι superkey. 24-7



Boyce-Codd Κανονική Μορφή (Λογική Διάσπασης σε BCNF)

Λογική Διάσπασης σε BCNF: Για κάθε FD που παραβιάζει την **BCNF** (δηλ., $X \rightarrow Y$, όπου X δεν είναι υπερκλειδί), δημιουργήσε μια νέα σχέση $R(X \rightarrow Y)$, διατηρώντας στην αρχική σχέση το X ως ξένο κλειδί.

LOTS1A \notin BCNF

<u>Property_id#</u>	County_name	Lot#	Area
---------------------	-------------	------	------



**Decomposition
(Διάσπαση)**

LOTS1AX

<u>Property_id#</u>	Area	Lot#
---------------------	------	------

\in BCNF

LOTS1AY

<u>Area</u>	County_name
-------------	-------------

\in BCNF

Boyce-Codd Κανονική Μορφή (Λογική Διάσπασης σε BCNF)



- Σας δίνεται η σχέση **BOOKING** η οποία περιγράφει τις κρατήσεις κάποιων γηπέδων αθλητισμού.
- Εξηγήστε εάν **BOOKING** ∈ **BCNF** δεδομένων **{Court, StartTime} → EndTime** και **{Court, EndTime} → StartTime**

Today's Court Bookings

Court	Start Time	End Time
1	9:30	10:30
1	11:00	12:00
1	14:00	15:30
2	10:00	11:30
2	11:30	13:30
2	15:00	16:30

Απάντηση

Εφόσον το αριστερό μέρος κάθε FD είναι superkey συνεπάγεται ότι η σχέση είναι σε BCNF

Κανονικοποίηση Ψηλότερου Βαθμού



- **2NF, 3NF, BCNF**: Στηρίζονται στα κλειδιά & συναρτησιακές εξαρτήσεις (FDs)
 - **4NF**: Στηρίζεται στα κλειδιά & Multi-valued dependencies (MVDs)
 - **5NF**: Στηρίζεται σε κλειδιά & Join Depend. (JDs)
- Πρακτικά, κάθε σχήμα βάσης **είναι καλό** να κανονικοποιήται σε **3NF** ή **BCNF** έτσι ώστε:
«Κάθε γνώρισμα να εξαρτάται από το κλειδί, πάνω σε ολόκληρο το κλειδί και όχι σε οτιδήποτε άλλο εκτός από το κλειδί» W. Kent (1992)
- Για να **διασφαλίσουμε** ένα καλό **σχεσιακό σχήμα** θα πρέπει να εγγυηθούμε και άλλες ιδιότητες (lossless join, dependency preservation; Κεφ. 11)



ΕΠΛ646 – Προχωρημένα Θέματα Βάσεων Δεδομένων

Παρουσίαση Γνωριμίας

Το ΕΠΛ646 Θα προσφερθεί το **Εαρινό
Εξάμηνο 2023**

Παρασκευή 15:00-18:00

Πανεπιστημιούπολη

<https://www.cs.ucy.ac.cy/~dzeina/courses/epl646/>

- Μάθημα Περιορισμένης Επιλογής
 - Αναγνώριση Εξειδίκευσης:



Προχωρημένες έννοιες Συστημάτων Βάσεων Δεδομένων

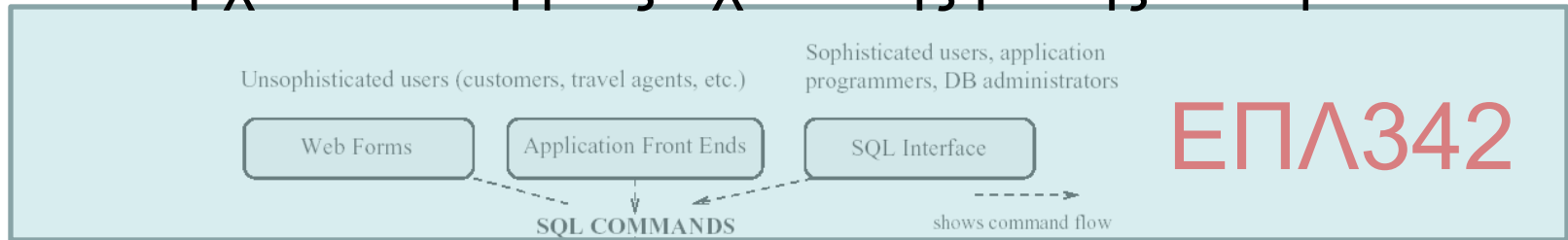
- **Στόχοι:**
 - Κατανόηση και Υλοποίηση προχωρημένων εννοιών που αφορούν την **εσωτερική λειτουργία** μιας σχεσιακής βάσης δεδομένων
 - Έκθεση σε **Προχωρημένα και Ανερχόμενα Θέματα** στο πεδίο των βάσεων δεδομένων.
 - Αναδυόμενες Τεχνολογίες και Ερευνητικές Κατευθύνσεις στο πεδίο των βάσεων δεδομένων



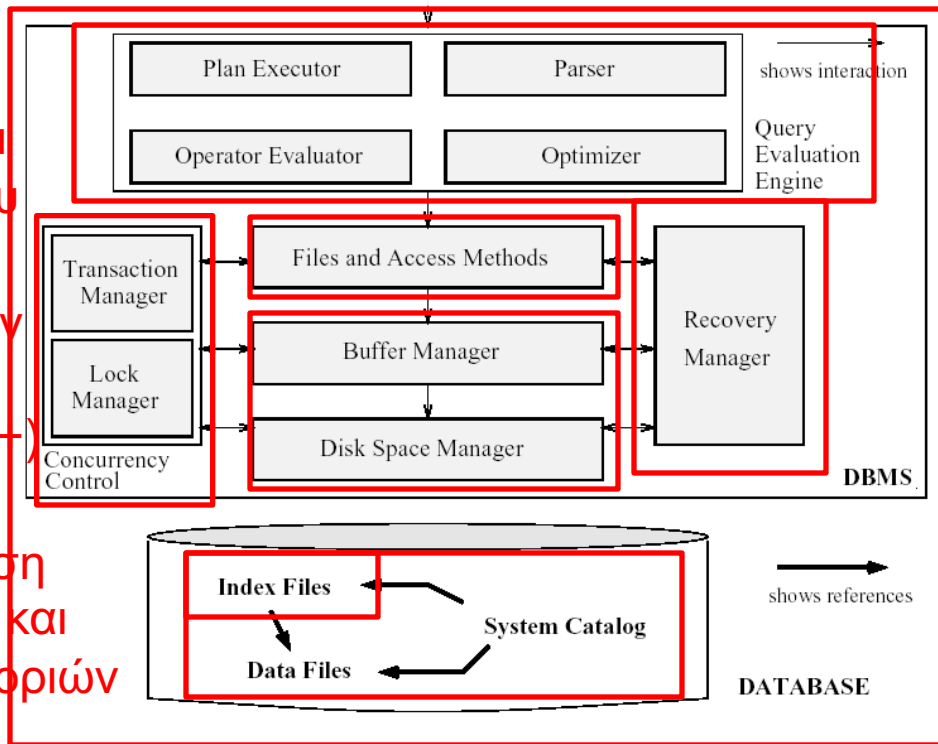
ΕΠΛ646: Τι καλύπτει;

A) Εσωτερική Λειτουργία μιας DB

Η αρχιτεκτονική μιας σχεσιακής βάσης δεδομένων



Έννοιες
 Δοσοληπιών και
 Τεχνικές Έλεγχου
 Ταυτοχρονίας
 Δομή Ευρετηρίων
 Δευτερευούσης
 Μνήμης (Hash, B+)



Αλγόριθμοι
 Βελτιστοποίησης
 Επερωτήσεων
 Τεχνικές
 Ανάκαμψης (σε
 περιπτώσεις
 σφαλμάτων)

Αποθήκευση
 Δεδομένων και
 Μετα-πληροφοριών

Ενδόμημη
 Διαχείριση
 Δεδομένων

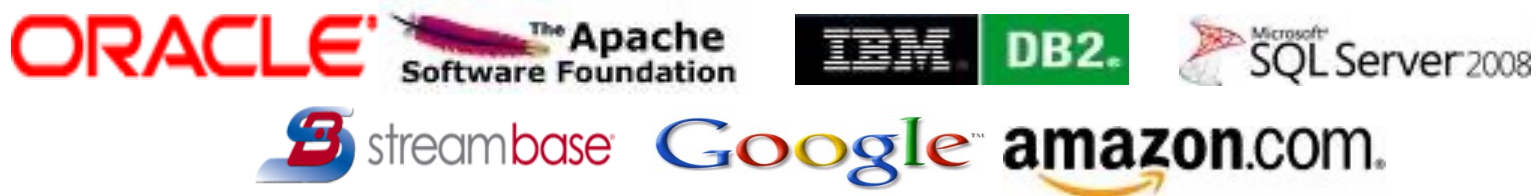
ΕΠΛ446

ΕΠΛ646: Τι καλύπτει;

B) Άλλα Προχωρημένα Θέματα



- **Κατανεμημένες Βάσεις Δεδομένων (Distributed Databases)**
- **Ημι-δομημένα Δεδομένα (XML/XPath/XQuery)**
- **Αναδυόμενες Τεχνολογίες Βάσεων Δεδομένων και Βάσεις σε Νέα Περιβάλλοντα:**
 - Streaming Databases, In-Memory Databases, Sensor Databases, Cloud Databases, ...
- **Χωρικές Βάσεις Δεδομένων (Spatial Databases)**



ΕΠΛ646: Σε ποιούς απευθύνεται;



- Άτομα που θέλουν να ασχοληθούν επαγγελματικά με τις βάσεις δεδομένων (π.χ., ως **DBA - Database Administrators** ή ακόμη ως **Database Analysts** και **Database Designers**).
- Άτομα που θέλουν να ακολουθήσουν μεταπτυχιακές σπουδές ή/και ακαδημαϊκή καριέρα.