



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Εισαγωγή στις Βάσεις Δεδομένων II

Ενότητα: Κανονικοποίηση

Διδάσκων: Πηγουνάκης Κωστής
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΣΧΟΛΗ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται στην άδεια χρήσης **Creative Commons** και ειδικότερα ***Αναφορά – Μη εμπορική Χρήση – Όχι Παράγωγο Έργο 3.0 Ελλάδα*** (***Attribution – Non Commercial – Non-derivatives 3.0 Greece***)



[ή επιλογή ενός άλλου από τους έξι συνδυασμούς]

[και αντικατάσταση λογότυπου άδειας όπου αυτό έχει μπει (σελ. 1, σελ. 2 και τελευταία)]

- Εξαιρείται από την ως άνω άδεια υλικό που περιλαμβάνεται στις διαφάνειες του μαθήματος, και υπόκειται σε άλλου τύπου άδεια χρήσης. Η άδεια χρήσης στην οποία υπόκειται το υλικό αυτό αναφέρεται ρητώς.

Χρηματοδότηση

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



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Κανονικοποίηση

Εισαγωγή στις Σχεσιακές Βάσεις Δεδομένων (ΟΙΚ3501)

Κωστής Πηγουνάκης

Τελευταία ενημέρωση : 01/04/2015

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

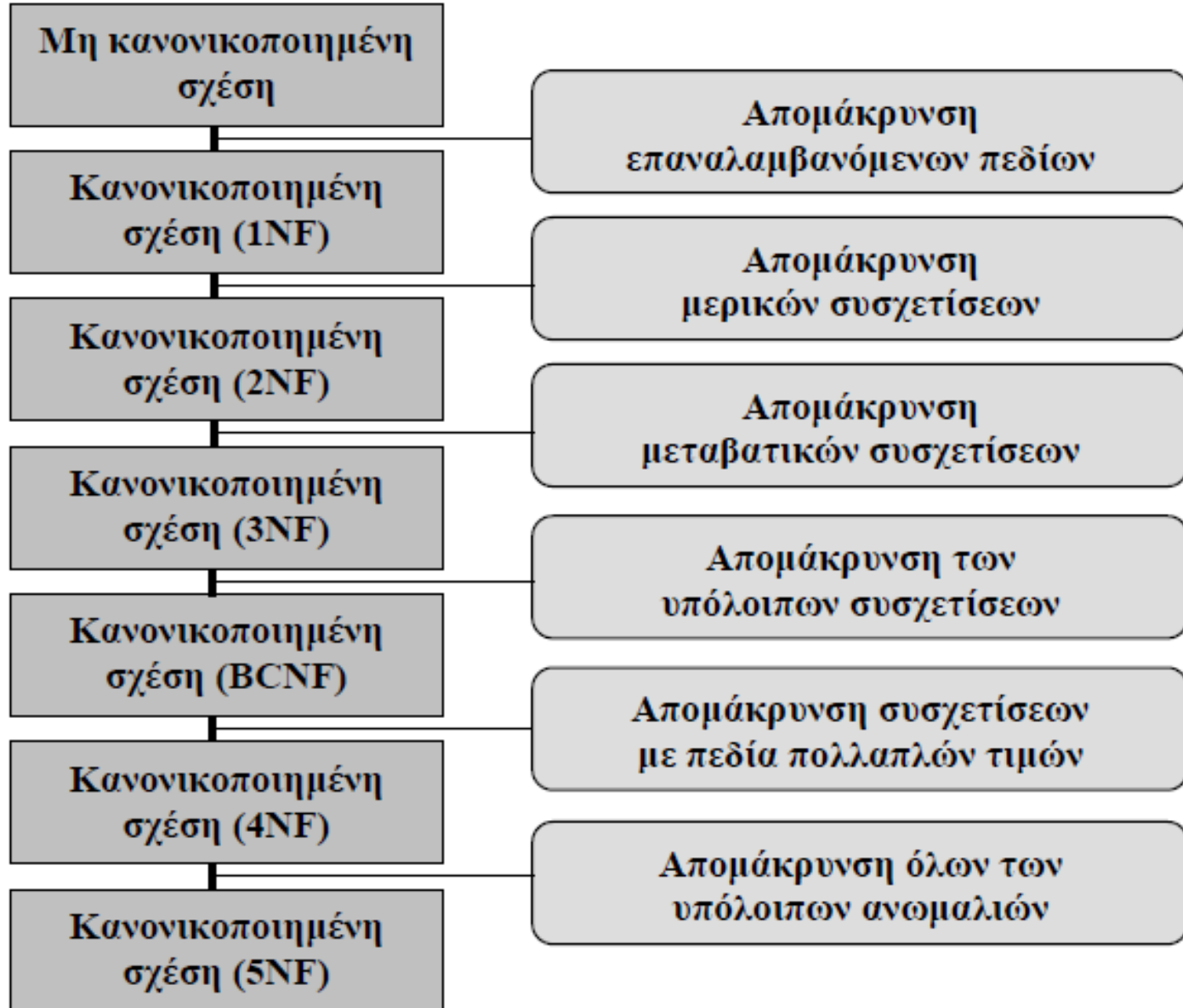
- **Κανονική μορφή (normal form):** ορίζεται ως η «κατάσταση» που βρίσκεται μια σχέση όσον αφορά τους τύπους των εξαρτήσεων που υφίστανται ανάμεσα στα πεδία της.
- Ο Codd εισήγαγε τρεις κανονικές μορφές, την *πρώτη (1NF)*, *δεύτερη (2NF)* και *τρίτη (3NF)* κανονική μορφή οι οποίες συσχετίζονται με τις 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 $\{\mathbf{StudentID}, \mathbf{Lesson}\}$: 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 $\{\mathbf{NULL}, \mathbf{biology}, \mathbf{Stewart}\}$, πράγμα που δεν επιτρέπεται όμως, καθώς το πεδίο **StudentID** είναι μέρος του πρωτεύοντος κλειδιού.
- Προτεινόμενες λύσεις :
 - $\mathbf{CLASS_1} \{\mathbf{StudentID}, \mathbf{Lesson}\}, \mathbf{CLASS_2} \{\mathbf{Teacher}, \mathbf{Lesson}\}$
 - $\mathbf{CLASS_1} \{\mathbf{StudentID}, \mathbf{Teacher}\}, \mathbf{CLASS_2} \{\mathbf{Teacher}, \mathbf{Lesson}\}$

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

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

