

# Κανονικές Μορφές

## Σχεδιασμός Σχεσιακών Σχημάτων

- Γενικές Οδηγίες
- Η Μέθοδος της Αποσύνθεσης
- Επιθυμητές Ιδιότητες της Αποσύνθεσης
  - Συνένωση Άνευ Απωλειών
  - Διατήρηση Εξαρτήσεων
  - Αποφυγή Επανάληψης Πληροφορίας

## Σχεδιασμός Σχεσιακών Σχημάτων

- Αποσύνθεση καθολικού σχήματος
- Επιθυμητές ιδιότητες
  - διατήρηση εξαρτήσεων
  - όχι απώλειες στη συνένωση
  - επανάληψη πληροφορίας λόγω ΣΕ
- Κανονικές Μορφές
  - BCNF
  - 3NF

## Σχεδιασμός Σχεσιακών Σχημάτων - Επανάληψη

### Αποσύνθεση (decomposition)

#### Αλγόριθμος σχεδιασμού

- Αρχικά ένα **καθολικό σχήμα σχέσης** που περιέχει όλα τα γνωρίσματα
- Προσδιορισμός των συναρτησιακών εξαρτήσεων
- **Διάσπαση** σε ένα σύνολο από σχήματα που ικανοποιούν κάποιες ιδιότητες

## Σχεδιασμός Σχεσιακών Σχημάτων - Επανάληψη

Έστω ένα σχεσιακό σχήμα R. Ένα σύνολο από σχεσιακά σχήματα  $\{R_1, R_2, \dots, R_n\}$  είναι μια **αποσύνθεση** του R αν **γνωρίσματα**

$$R = R_1 \cup R_2 \cup \dots \cup R_n$$

Δηλαδή,  $\forall i = 1, \dots, n \quad R_i \subseteq R$

Έστω  $r(R)$  και  $r_i = \pi_{R_i}(r)$ ,  $\forall i = 1, \dots, n$  **πλειάδες**

$$r \subseteq r_1 * r_2 * \dots * r_n$$

## Σχεδιασμός Σχεσιακών Σχημάτων - Επανάληψη

### Επιθυμητές Ιδιότητες Αποσύνθεσης

#### 1. Συνενώσεις Άνευ Απωλειών

Η φυσική συνένωση των σχέσεων που προκύπτουν μας δίνει **ακριβώς** την αρχική σχέση (χωρίς επιπρόσθετες πλειάδες):  $r = \pi_{R_1}(r) * \pi_{R_2}(r) * \dots * \pi_{R_n}(r)$

$R_1 \cap R_2 \rightarrow R_1$  ή  $R_1 \cap R_2 \rightarrow R_2$  ανήκει στο  $F^*$ , δηλαδή τα κοινά γνωρίσματα των δύο σχημάτων είναι κλειδί για τομλάχιστον ένα από τα δύο

#### 2. Διατήρηση Εξαρτήσεων

Στόχος: Έλεγχος διατήρησης εξαρτήσεων όταν γίνονται τροποποιήσεις χωρίς να υπολογίζουμε τις αρχικές σχέσεις (αποφυγή των συνενώσεων)

$$F^* = F_1 \cup F_2 \cup \dots \cup F_n, \text{ πέρνει } F^* = F^*$$

#### 3. Αποφυγή Επανάληψης Πληροφορίας, **πως; Κανονικές Μορφές**

## Boyce-Codd Κανονική Μορφή

Ένα σχεσιακό σχήμα R είναι σε **Κανονική Μορφή Boyce-Codd (BCNF)** σε σχέση με ένα σύνολο F συναρτησιακών εξαρτήσεων αν

για όλες τις ΣΕ στο F της μορφής  $X \rightarrow Y$  ισχύει τουλάχιστον ένα από τα παρακάτω:

- $X \rightarrow Y$  είναι μια τετριμμένη ΣΕ ή
- X είναι **υπερκλειδί** του σχήματος R

Δηλαδή το αριστερό μέρος κάθε μη τετριμμένης ΣΕ πρέπει να περιέχει ένα κλειδί

Το σχήμα μιας ΒΔ είναι σε BCNF αν το σχήμα **κάθε** σχέσης της είναι σε BCNF.

## Boyce-Codd Κανονική Μορφή

### Παράδειγμα 1

Ταινία (Τίτλος, Έτος, Διάρκεια, Είδος, Εταιρεία-Παραγωγής, Όνομα-Ηθοποιού)

Η σχέση Ταινία δεν είναι σε BCNF

(υποψήφιο) κλειδί: {Τίτλος, Έτος, Όνομα-Ηθοποιού}

Για παράδειγμα η ΣΕ Τίτλος Έτος  $\rightarrow$  Εταιρεία-Παραγωγής

## Boyce-Codd Κανονική Μορφή

### Παράδειγμα 2

Ταινία2 (Τίτλος, Έτος, Διάρκεια, Είδος, Εταιρεία-Παραγωγής)

Η σχέση Ταινία2 είναι σε BCNF

### Παράδειγμα 3

Οποιαδήποτε σχέση με δύο γνωρίσματα είναι σε BCNF

## Boyce-Codd Κανονική Μορφή

### Αλγόριθμος Αποσύνθεσης σε BCNF

- Βρες μια μη τετριμμένη ΣΕ που παραβιάζει τον BCNF ορισμό, έστω  $X \rightarrow Y$  και  $X \cap Y = \emptyset$
- Αποσύνθεση του αρχικού σχήματος R σε δύο σχήματα  
 $R_1$  με γνωρίσματα  $X \cup Y$   
 $R_2$  με γνωρίσματα  $R - Y$

Ευριστικός: στα δεξιά όσο το δυνατόν περισσότερα γνωρίσματα

Αποσύνθεση χωρίς απώλειες;

## Boyce-Codd Κανονική Μορφή

### Παράδειγμα 1

Ταινία (Τίτλος, Έτος, Διάρκεια, Είδος, Εταιρεία-Παραγωγής, Όνομα-Ηθοποιού)

Τίτλος Έτος  $\rightarrow$  Διάρκεια Είδος Εταιρεία-Παραγωγής

Ταινία1(Τίτλος, Έτος, Διάρκεια, Είδος, Εταιρεία-Παραγωγής)

Ταινία2(Τίτλος, Έτος, Όνομα-Ηθοποιού)

## Boyce-Codd Κανονική Μορφή

### Παράδειγμα 2

Ταινία-Εταιρεία (Τίτλος, Έτος, Διάρκεια, Είδος, Εταιρεία-Παραγωγής, Διεύθυνση-Εταιρείας)

Πρόβλημα: υπάρχει μια **μεταβατική** εξάρτηση

Τίτλος Έτος  $\rightarrow$  Εταιρεία-Παραγωγής

**Εταιρεία-Παραγωγής  $\rightarrow$  Διεύθυνση-Εταιρείας**

Τίτλος Έτος  $\rightarrow$  Διεύθυνση-Εταιρείας

Ταινία-Εταιρεία1 (Εταιρεία-Παραγωγής, Διεύθυνση-Εταιρείας)

Ταινία-Εταιρεία2 (Τίτλος, Έτος, Διάρκεια, Είδος, Εταιρεία-Παραγωγής)

### Boyce-Codd Κανονική Μορφή

- Μπορεί να χρειαστεί παραπάνω από μία αποσύνθεση

Αποσύνθεση του αρχικού σχήματος R σε δύο σχήματα -  $R_1$  με γνωρίσματα  $X \cup Y$  και  $R_2$  με γνωρίσματα  $R - Y$

$R_2$  μπορεί να μην είναι σε BCNF

### Boyce-Codd Κανονική Μορφή

- Δεν είναι πάντα δυνατή η αποσύνθεση σε μια BCNF που να διατηρεί τις εξαρτήσεις

Παράδειγμα

Έστω η σχέση Παίζει(Έργο, Κινηματογράφος, Πόλη) με τους περιορισμούς ότι (i) δεν υπάρχουν κινηματογράφοι με το ίδιο όνομα, (ii) κάθε κινηματογράφος έχει πολλές αίθουσες (παίζει πολλά έργα) και (iii) κάθε έργο παίζεται μόνο σε ένα κινηματογράφο σε κάθε πόλη

Κινηματογράφος → Πόλη

Κλειδιά:

Έργο Πόλη → Κινηματογράφος

{Έργο, Πόλη}  
{Κινηματογράφος, Έργο}

### Boyce-Codd Κανονική Μορφή

Παίζει(Έργο, Κινηματογράφος, Πόλη)

Κινηματογράφος → Πόλη

Έργο Πόλη → Κινηματογράφος

Κλειδιά

{Έργο, Πόλη}

{Κινηματογράφος, Έργο}

Αποσύνθεση σε: {Κινηματογράφος, Πόλη} και {Κινηματογράφος, Έργο}

Κινηματογράφος	Πόλη	Κινηματογράφος	Έργο
Οντεόν	Αθήνα	Οντεόν	Safe Sex
Village Center	Αθήνα	Village Center	Safe Sex

### Τρίτη Κανονική Μορφή

Ένα σχεσιακό σχήμα R είναι σε **τρίτη κανονική μορφή (3NF)** σε σχέση με ένα σύνολο F συναρτησιακών εξαρτήσεων αν για όλες τις ΣΕ στο  $F^+$  της μορφής  $X \rightarrow Y$  ισχύει τουλάχιστον ένα από τα παρακάτω:

- $X \rightarrow Y$  είναι μια τετριμμένη ΣΕ ή
- X είναι υπερκλειδί του σχήματος R
- κάθε γνώρισμα A του Y - X περιέχεται σε κάποιο υποψήφιο κλειδί

BCNF πιο περιοριστική -- αν σε BCNF  $\Rightarrow$  3NF

### Τρίτη Κανονική Μορφή

Παράδειγμα

Παίζει(Έργο, Κινηματογράφος, Πόλη)

Κινηματογράφος → Πόλη

Έργο Πόλη → Κινηματογράφος

Κλειδιά {Έργο, Πόλη}  
{Κινηματογράφος, Έργο}

Η σχέση είναι σε 3NF

Είδος ΣΕ: μοναδικά  
αντικείμενα ή με βάση  
πρακτικές

### Τρίτη Κανονική Μορφή

#### Αλγόριθμος Αποσύνθεσης σε 3NF

- Υπολόγισε το ελάχιστο κάλυμμα  $F_c$  της F
- Για κάθε ΣΕ  $X \rightarrow Y$  της  $F_c$  αν κανένα από τα  $R_i$  δεν περιέχει τα X, Y νέα σχέση με γνωρίσματα  $X \cup Y$
- Αν καμία από τις σχέσεις που προέκυψαν δεν περιέχει κάποιο υποψήφιο κλειδί του R δημιουργήσε μια νέα σχέση με γνωρίσματα τα γνωρίσματα του κλειδιού

Αλγόριθμος Αποσύνθεσης σε 3NF

- Απώλειες στη συνένωση;
- Διατήρηση εξαρτήσεων;

Παράδειγμα

Τραπεζίτης(Όνομα-Υποκαταστήματος, Όνομα-Πελάτη, Όνομα-Τραπεζίτη, Αριθμός Γραφείου)

Όνομα-Τραπεζίτη → Όνομα-Υποκαταστήματος Αριθμός-Γραφείου  
 Όνομα-Πελάτη Όνομα-Υποκαταστήματος → Όνομα-Τραπεζίτη

Κλειδιά {Όνομα-Πελάτη, Όνομα-Υποκαταστήματος}

3NF;

Τραπεζίτης1(Όνομα-Τραπεζίτη, Όνομα-Υποκαταστήματος Αριθμός-Γραφείου)

Τραπεζίτης2(Όνομα-Πελάτη, Όνομα-Υποκαταστήματος, Όνομα-Τραπεζίτη)

BCNF;

Κανονική Μορφή Boyce-Codd

Ένα σχεσιακό σχήμα R είναι σε BCNF σε σχέση με ένα σύνολο F συναρτησιακών εξαρτήσεων αν για όλες τις ΣΕ στο F\* της μορφής X → Y ισχύει τουλάχιστον ένα από τα παρακάτω:

- X → Y είναι μια τετριμμένη ΣΕ ή
- X είναι υπερκλειδί του σχήματος R

Τρίτη Κανονική Μορφή

-- κάθε γνώρισμα A του Y - X περιέχεται σε κάποιο υποψήφιο κλειδί

	BCNF	3NF
• Αποφυγή επανάληψης πληροφορίας	ναι	όχι πάντα
• Αποσύνθεση χωρίς απώλειες στη συνένωση	ναι	ναι
• Διατήρηση εξαρτήσεων	όχι πάντα	ναι

Ιστορικοί Λόγοι

Πρώτη Κανονική Μορφή (1NF)

Δεν επιτρέπονται πλειότιμα ή σύνθετα γνωρίσματα

Παράδειγμα με σύνθετα (εμφωλευόμενες σχέσεις)

ΕΡΓ\_ΕΡΓΟ(ΑΡ\_ΤΑΥΤ, ΕΡ\_ΟΝΟΜΑ, {ΕΡΓΑ(ΚΩΔ\_ΕΡΓΟΥ, ΩΡΕΣ)}))

ΕΡΓ\_ΕΡΓΟ(ΑΡ\_ΤΑΥΤ, ΕΡ\_ΟΝΟΜΑ, ΚΩΔ\_ΕΡΓΟΥ, ΩΡΕΣ)

ΑΡ\_ΤΑΥΤ → ΕΡ\_ΟΝΟΜΑ      ΑΡ\_ΤΑΥΤ ΚΩΔ\_ΕΡΓΟΥ → ΩΡΕΣ

Αποσύνθεση (unnesting) σε ΕΡΓ\_ΕΡΓΟ1(ΑΡ\_ΤΑΥΤ, ΕΡ\_ΟΝΟΜΑ)  
 ΕΡΓ\_ΕΡΓΟ2(ΑΡ\_ΤΑΥΤ, ΚΩΔ\_ΕΡΓΟΥ, ΩΡΕΣ)

Παράδειγμα με πλειότιμα

ΤΜΗΜΑ(Τ\_ΟΝΟΜΑ, ΚΩΔ\_ΤΜΗΜ, ΔΙΕΥΘΥΝΤΗΣ\_ΤΜΗΜ, ΤΟΠΟΘΕΣΙΕΣ)

Ένα τμήμα σε πολλές τοποθεσίες. Κλειδί:

ΚΩΔ\_ΤΜΗΜΑΤΟΣ →→ ΤΟΠΟΘΕΣΙΕΣ

Αποσύνθεση σε ΤΜΗΜΑ1(ΚΩΔ\_ΤΜΗΜΑΤΟΣ, ΤΟΠΟΘΕΣΙΕΣ)  
 ΤΜΗΜΑ2(Τ\_ΟΝΟΜΑ, ΚΩΔ\_ΤΜΗΜ, ΔΙΕΥΘΥΝΤΗΣ\_ΤΜΗΜ)

Πλήρης Συναρτησιακή Εξάρτηση

Μια συναρτησιακή εξάρτηση  $X \rightarrow Y$  είναι **πλήρης συναρτησιακή εξάρτηση** αν για κάθε  $A \in X$ , δεν ισχύει  $(X - \{A\}) \rightarrow Y$

δηλαδή η αφαίρεση οποιουδήποτε γνωρίσματος  $A$  από το  $X$  σημαίνει ότι η εξάρτηση δεν ισχύει πλέον

δεν έχει περιττά γνωρίσματα στο αριστερό μέρος

Μερική εξάρτηση αν όχι πλήρης

**Πρωτεύον Γνώρισμα:** Ένα γνώρισμα που είναι μέλος κάποιου υποψήφιου κλειδιού

Ένα σχήμα σχέσης είναι σε **Δεύτερη Κανονική Μορφή (2NF)** αν κάθε **μη πρωτεύον** γνώρισμα  $A$  (δηλαδή γνώρισμα που δεν εμφανίζεται σε κανένα υποψήφιο κλειδί) του  $R$  είναι **πλήρως** συναρτησιακά εξαρτώμενο από οποιοδήποτε κλειδί του  $R$ .

Παράδειγμα παραβίασης της 2NF

ΕΡΓ\_ΕΡΓΟ(ΑΡ\_ΤΑΥΤ, ΕΡ\_ΟΝΟΜΑ, ΚΩΔ\_ΕΡΓΟΥ, ΩΡΕΣ)

ΑΡ\_ΤΑΥΤ  $\rightarrow$  ΕΡ\_ΟΝΟΜΑ      ΑΡ\_ΤΑΥΤ ΚΩΔ\_ΕΡΓΟΥ  $\rightarrow$  ΩΡΕΣ

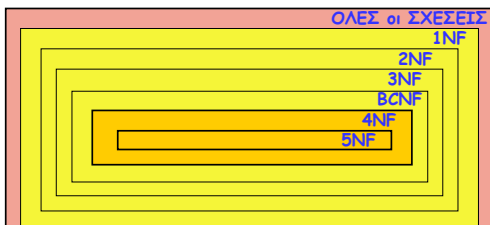
Μια συναρτησιακή εξάρτηση  $X \rightarrow Y$  σε ένα σχήμα σχέσης  $R$  είναι μια **μεταβατική εξάρτηση** αν υπάρχει ένα σύνολο γνωρισμάτων  $Z$  που δεν είναι υποσύνολο οποιουδήποτε κλειδιού της  $R$  τέτοιο ώστε να ισχύουν  $X \rightarrow Z$  και  $Z \rightarrow Y$ .

Τρίτη Κανονική Μορφή - Εναλλακτικός Ορισμός

Ένα σχήμα σχέσης είναι σε 3NF αν κάθε **μη πρωτεύον** γνώρισμα  $A$  του  $R$  είναι

-- πλήρως συναρτησιακά εξαρτώμενο από κάθε κλειδί του  $R$  και

-- μη μεταβατικά εξαρτώμενο από κάθε κλειδί του  $R$



- Ένας μεγάλος αριθμός από εμπορικά εργαλεία, δοθέντων ενός συνόλου Σχημάτων Σχέσεων/Γνωρισμάτων και ενός συνόλου συναρτησιακών εξαρτήσεων **δημιουργούν αυτόματα** σχήματα σχέσεων σε μορφή 3NF (σπάνια πάνε σε BCNF, 4NF και 5NF)
- Μια άλλη χρήση τέτοιων εργαλείων είναι να **ελέγχουν το επίπεδο κανονικοποίησης** μιας σχέσης - γενικά, η χρήση ως ευριστικό εργαλείο επιλογής ενός σχεδιασμού έναντι κάποιου άλλου
- Υπάρχουν **πρακτικά αποτελέσματα** της θεωρίας που επιτρέπουν σε έναν σχεδιαστή να κάνει ανάλυση της μορφής:

Αν μια σχέση είναι σε 3NF και κάθε υποψήφιο κλειδί αποτελείται ακριβώς από ένα γνώρισμα, τότε είναι και σε 5NF (Fagin, 1991)

- Η διαδικασία Κανονικοποίησης έχει και *μειονεκτήματα*:
  - ο Δεν είναι δημιουργική -- με στόχο τα κριτήρια που αναφέρθηκαν προηγουμένως, δεν υπάρχει τρόπος να δημιουργηθεί μια «καλή» βάση δεδομένων
  - ο **Συνήθως η κανονικοποίηση γίνεται αφού έχουμε κάποιο σχήμα (μας λέει αν είναι «καλό» ή «κακό»)**
  - ο Δεν προσφέρει ένα εννοιολογικό σχήμα (ασχολείται μόνο με σχέσεις και γνωρίσματα)

Όμως, είναι μια ενδιαφέρουσα και πρακτικά χρήσιμη προσπάθεια να γίνουν με τυπικό και συστηματικό τρόπο πράγματα που τα κάνουμε συνήθως διαισθητικά.

### Η Διαδικασία Σχεδιασμού

1. Συλλογή και ανάλυση απαιτήσεων
2. Εννοιολογικός σχεδιασμός
3. Επιλογή ΣΔΒΔ
4. Απεικόνιση στο μοντέλο δεδομένων (λογικός σχεδιασμός)
5. Φυσικός σχεδιασμός
6. Υλοποίηση