# Λογικός Σχεδιασμός Σχεσιακών Σχημάτων: Αποσύνθεση

Θα εξετάσουμε πότε ένα σχεσιακό σχήμα για μια βάση δεδομένων είναι *«καλό»*

- Γενικές Οδηγίες
- Η Μέθοδος της Αποσύνθεσης (γενική μεθοδολογία)
- Επιθυμητές Ιδιότητες της Αποσύνθεσης

  Συνένωση Άνευ Απωλειών

  Διατήρηση Εξαρτήσεων

  Αποφυγή Επανάληψης Πληροφορίας

# Σχεδιασμός *καλών* σχεσιακών σχημάτων

- Μη τυπικές - γενικές κατευθύνσεις
- Θεωρία κανονικών μορφών που βασίζεται στις συναρτησιακές εξαρτήσεις

# Γενικές Κατευθύνσεις

1. Σημασιολογία
2. Ελάττωση πλεονασμού
3. Ελάττωση τιμών null
4. Μη πλασματικές πλειάδες

## 1. Σημασιολογία

- Εύκολη η εξήγηση της σημασίας του

- Αποφυγή συνδυασμού γνωρισμάτων από πολλές οντότητες και συσχετίσεις στην ίδια σχέση

**Ταινία**

| Τίτλος | Έτος | Διάρκεια | Είδος |
|---|---|---|---|

**Παίζει**

| Όνομα | Τίτλος | Έτος |
|---|---|---|

**Ηθοποιός**

| Όνομα | Διεύθυνση | Έτος-Γέννησης |
|---|---|---|

# 2. Πλεονασμός (επανάληψη πληροφορίας)

*εδώ βοηθούν οι συναρτησιακές εξαρτήσεις*

**Ταινία**

| Τίτλος | Έτος | Διάρκεια | Είδος | Όνομα-Ηθοποιού |
|---|---|---|---|---|

*Εισαγωγή*

- Για την εισαγωγή μιας νέας ταινίας πρέπει να εισάγουμε τουλάχιστον έναν ηθοποιό (τιμή null;)
- Για την εισαγωγή ενός ηθοποιού στην ταινία πρέπει να επαναλάβουμε τα γνωρίσματα (διάρκεια, είδος) της ταινίας

**Ταινία**

| <u>Τίτλος</u> | <u>Έτος</u> | Διάρκεια | Είδος | <u>Όνομα-Ηθοποιού</u> |
|---|---|---|---|---|

*Διαγραφή*

- Τι γίνεται αν διαγράψουμε και τον τελευταίο ηθοποιό
- Διαγραφή μιας ταινίας;

**Ταινία**

| Τίτλος | Έτος | Διάρκεια | Είδος | Όνομα-Ηθοποιού |
|---|---|---|---|---|

*Τροποποίηση*

- Τι γίνεται αν θελήσουμε να τροποποιήσουμε τη διάρκεια μιας ταινίας;

**Σύνοψη Προβλημάτων Λόγω Πλεονασμού**

- Πλεονασμός στην αποθήκευση
- Προβληματική ενημέρωση
- Προβληματική εισαγωγή
- Προβληματική διαγραφή

# 3. Αποφυγή τιμών null

**Ηθοποιός**

| Όνομα | Διεύθυνση | Έτος-Γέννησης | Σύζυγος-Ηθοποιού |
|---|---|---|---|

**Ηθοποιός**

| Όνομα | Διεύθυνση | Έτος-Γέννησης |
|---|---|---|

**Ζευγάρι-Ηθοποιών**

| Όνομα | Σύζυγος-Ηθοποιού |
|---|---|

# 4. Αποφυγή δημιουργίας πλασματικών πλειάδων

(αδυναμία αναπαράστασης συγκεκριμένης πληροφορίας)

**Ταινία**

| Τίτλος | Έτος | Διάρκεια | Είδος |
|---|---|---|---|

**Παίζει**

| Τίτλος | Όνομα-Ηθοποιού |
|---|---|

*Χάνουμε πληροφορία δεν μπορούμε να βρούμε ποιος ηθοποιός σε ποια ταινία*

**Ταινία**

| Τίτλος | Έτος | Διάρκεια | Είδος | Όνομα-Ηθοποιού |
|---|---|---|---|---|

## Αλγόριθμος Σχεδιασμού - Εισαγωγή

Ο τρόπος που σχεδιάζαμε ένα σχήμα ΒΔ μέχρι τώρα:

από το εννοιολογικό στο σχεσιακό μοντέλο

Θα δούμε ένα *γενικό* ***θεωρητικό*** *(formal) τρόπο* κατασκευής του σχήματος

Γενικά:

- Ξεκινάμε από το καθολικό σχήμα (που περιέχει όλα τα γνωρίσματα)
- Διαδοχικές διασπάσεις έτσι ώστε τα σχήματα που προκύπτουν να ικανοποιούν κάποιες ιδιότητες (να είναι σε κάποιες κανονικές μορφές)

-- top-down τεχνική

## Αλγόριθμος Σχεδιασμού - Εισαγωγή

Μπορούμε να το εφαρμόσουμε και για να «διασπάσουμε» μια «προβληματική» σχέση που προέκυψε από την μετατροπή του εννοιολογικού σχεδιασμού

Μειονέκτημα των διασπάσεων:

μπορεί να απαιτεί συνενώσεις (join) για να απαντηθούν ερωτήματα η να ελεγχθούν εξαρτήσεις -> αποδοτικότητα του συστήματος

Ένας γενικός (θεωρητικός) τρόπος κατασκευής του σχήματος

Αποσύνθεση (decomposition)

**Αλγόριθμος σχεδιασμού**

1. Αρχικά ένα καθολικό (universal) σχήμα σχέσης που περιέχει όλα τα γνωρίσματα
2. Προσδιορισμός των συναρτησιακών εξαρτήσεων
3. Διάσπαση σε ένα σύνολο από σχήματα σχέσεων που ικανοποιούν κάποιες ιδιότητες

- Αποσύνθεση καθολικού σχήματος

Επιθυμητές ιδιότητες

1. διατήρηση εξαρτήσεων
2. όχι απώλειες στη συνένωση
3. όχι επανάληψη πληροφορίας λόγω ΣΕ

Κανονικές μορφές

*Παράδειγμα*

Καθολικό Σχήμα: R = {Τίτλος, Έτος, Διάρκεια, Είδος, Όνομα-Ηθοποιού, Διεύθυνση, Έτος-Γέννησης}

Σύνολο ΣΕ που ισχύουν στο πρόβλημα:

| |
|---|
| Τίτλος Έτος → Είδος |
| Τίτλος Έτος → Διάρκεια |
| Όνομα Ηθοποιού → Διεύθυνση |
| Όνομα-Ηθοποιού → Έτος Γέννησης |

Πιθανή διάσπαση:

$R_1$ = {Τίτλος, Έτος, Διάρκεια, Είδος}

$R_2$ = {Τίτλος, Έτος, Όνομα-Ηθοποιού, Διεύθυνση, Έτος-Γέννησης}

- *Ποια είναι μια καλή διάσπαση; Πως μπορούμε να πάρουμε την αρχική σχέση;*
- *Μπορούμε να διασπάσουμε την $R_2$ με τον ίδιο τρόπο.*

# Τυπικός ορισμός

Αρχικά ένα *καθολικό σχήμα* **R = {$A_1$, $A_2$, …, $A_n$}** αποσύνθεση (decomposition) σε δύο σχήματα

**$R_1$ = {$B_1$, $B_2$, …, $B_m$}** και **$R_2$ = {$C_1$, $C_2$, …, $C_k$}**

τέτοια ώστε:

**1.** {$A_1$, $A_2$, …, $A_n$} = {$B_1$, $B_2$, …, $B_m$} $\cup$ {$C_1$, $C_2$, …, $C_k$} *(διατήρηση γνωρισμάτων)* *[γνωρίσματα]*

**2.** Οι πλειάδες της $r_1(R_1)$ είναι η *προβολή των πλειάδων* της r(R) στα {$B_1$, $B_2$, …, $B_m$} *[πλειάδες]*

**3.** Οι πλειάδες της $r_2(R_2)$ είναι η *προβολή των πλειάδων* της r(R) στα {$C_1$, $C_2$, …, $C_k$} *[πλειάδες]*

## Αλγόριθμος Σχεδιασμού (παράδειγμα)

Έστω το (καθολικό) σχήμα R(A, B, C) αποσύνθεση σε $R_1(A, B)$ και $R_2(B, C)$

Τι γίνεται με τα στιγμιότυπα (σχέσεις) που ανήκουν στο R (συμβολισμός r(R) ή **r**)?

**r(R) - r**

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 5 |

**$r_1(R_1)$**

| A | B |
|---|---|
| 1 | 2 |
| 4 | 5 |

**$r_2(R_2)$**

| B | C |
|---|---|
| 2 | 3 |
| 5 | 5 |

Μπορούμε να πάρουμε το αρχικό στιγμιότυπο;

Φυσική συνένωση $r_1 * r_2$ ;

γνωρίσματα

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

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

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

στιγμιότυπα

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

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

Έστω το σχήμα R(A, B, C) αποσύνθεση σε $R_1$(A, B) και $R_2$(B, C)

Τι γίνεται με τα στιγμιότυπα (σχέσεις) που ανήκουν στο R?

Έστω r(R) και $r_i = \pi_{R_i}(r)$, $\forall$ i = 1,..,n ---- $r \subseteq r_1 * r2 * \ldots * r_n$

*Παράδειγμα*

**r**

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 2 | 5 |

**$r_1$**

| A | B |
|---|---|
| 1 | 2 |
| 4 | 2 |

**$r_2$**

| B | C |
|---|---|
| 2 | 3 |
| 2 | 5 |

**$r_1 * r_2$**

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 1 | 2 | 5 |
| 4 | 2 | 3 |
| 4 | 2 | 5 |

$R_1 \cap R_2 = B$

- *Δεν μπορούμε να πάρουμε την αρχική σχέση r από τα $r_1$ και $r_2$*

Έστω το σχήμα R(A, B, C) αποσύνθεση σε $R_1$(A, C) και $R_2$(B, C)

Τι γίνεται με τα στιγμιότυπα (σχέσεις) που ανήκουν στο R;

**r(R)**

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 2 | 5 |

$$R_1 \cap R_2 = C$$

Μπορούμε να πάρουμε το αρχικό στιγμιότυπο;

Φυσική συνένωση $r_1$ * $r_2$

**$r_1(R_1)$**

| A | C |
|---|---|
| 1 | 3 |
| 4 | 5 |

**$r_2(R_2)$**

| B | C |
|---|---|
| 2 | 3 |
| 2 | 5 |

# Επιθυμητές Ιδιότητες για την Αποσύνθεση

## 1. Συνενώσεις Άνευ Απωλειών

Έστω *C* το σύνολο περιορισμών. Μια αποσύνθεση του R σε $\{R_1, R_2, .., R_n\}$ είναι μια **αποσύνθεση άνευ απωλειών στη συνένωση** (lossless join decomposition) αν για όλες τις σχέσεις r(R) που είναι νόμιμες στο *C* ισχύει

$$r = \pi_{R_1}(r) * \pi_{R_2}(r) * \ldots \pi_{R_n}(r)$$

Ονομάζεται και **μη προσθετική συνένωση** (non-additive join)

*Παράδειγμα*

**r**

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 2 | 5 |

$\mathbf{r_1}$

| A | B |
|---|---|
| 1 | 2 |
| 4 | 2 |

$\mathbf{r_2}$

| B | C |
|---|---|
| 2 | 3 |
| 2 | 5 |

$\mathbf{r_1 * r_2}$

| A | B | C |
|---|---|---|
| 1 | 2 | 3 |
| 1 | 2 | 5 |
| 4 | 2 | 3 |
| 4 | 2 | 5 |

$\mathbf{r'_1}$

| A | C |
|---|---|
| 1 | 3 |
| 4 | 5 |

$\mathbf{r'_2}$

| B | C |
|---|---|
| 2 | 3 |
| 2 | 5 |

$\mathbf{r'_1 * r'_2 =}$;

## Θεώρημα

Έστω R ένα σχεσιακό σχήμα και F ένα σύνολο από συναρτησιακές εξαρτήσεις στο R. Έστω $R_1$ και $R_2$ μια αποσύνθεση του R. Αν μια τουλάχιστον από τις ΣΕ

$$R_1 \cap R_2 \rightarrow R_1 \text{ ή } R_1 \cap R_2 \rightarrow R_2 \text{ ανήκει στο } F^+$$

τότε η διάσπαση είναι χωρίς απώλειες στη συνένωση.

**Δηλαδή τα κοινά γνωρίσματα των δύο σχημάτων είναι κλειδί για τουλάχιστον ένα από τα δύο σχήματα**

*Παράδειγμα:* R = {Τίτλος, Έτος, Διάρκεια, Είδος, Όνομα-Ηθοποιού, Διεύθυνση, Έτος-Γέννησης}

Τίτλος Έτος → Διάρκεια

Τίτλος Έτος → Είδος

Όνομα Ηθοποιού → Διεύθυνση

Όνομα-Ηθοποιού → Έτος Γέννησης

$R_1$ = {Τίτλος, Έτος, Διάρκεια, Είδος}

$R_2$ = {Τίτλος, Έτος, Όνομα-Ηθοποιού, Διεύθυνση, Έτος-Γέννησης}

$R_1 \cap R_2$ = {Τίτλος, Έτος}

# Επιθυμητές Ιδιότητες για την Αποσύνθεση

## 2. Διατήρηση Εξαρτήσεων

Στόχος:

Για να ελέγξουμε ότι διατηρούνται οι Σ.Ε. όταν γίνονται τροποποιήσεις σε μία από τις σχέσεις $r_i(R_i)$,

να αρκεί να ελέγξουμε μόνο τη συγκεκριμένη σχέση (δηλαδή, να μη χρειάζεται να υπολογίσουμε τις αρχικές σχέσεις - αποφυγή των συνενώσεων)

Έστω F ένα σύνολο από ΣΕ στο σχήμα R και $\{R_1, R_2, .., R_n\}$ μια αποσύνθεση του R.

**$F_i$ περιορισμός του F στο $R_i$** είναι το σύνολο όλων των συναρτησιακών εξαρτήσεων του $\mathbf{F^+}$ που περιέχουν μόνο γνωρίσματα του $R_i$.

Προσοχή: $F^+$ όχι F

Παράδειγμα: Υπολογισμός του περιορισμού του F σε ένα σχήμα

Παράδειγμα 1: Έστω R(A, B, C, D), F = {A $\rightarrow$B, B $\rightarrow$ C}. Περιορισμός του F στο S(A, C) *(δηλαδή ποιες ΣΕ του $F^+$ ισχύουν στο S)*

**Παράδειγμα 2**: Έστω R(A, B, C, D, E), F = {A $\rightarrow$ D, B $\rightarrow$ E, DE $\rightarrow$ C}. Περιορισμός του F στο S(A, B, C)

## Διατήρηση Εξαρτήσεων

Έστω F ένα σύνολο από ΣΕ στο σχήμα R και $\{R_1, R_2, .., R_n\}$ μια αποσύνθεση του R.

$$\text{Έστω } F' = F_1 \cup F_2 \ldots \cup F_n$$

Η αποσύνθεση είναι μια **αποσύνθεση που διατηρεί τις εξαρτήσεις** (dependency preserving) αν $F'^+ = F^+$

Παράδειγμα: Πως δείχνουμε αν μια αποσύνθεση διατηρεί τις εξαρτήσεις

Έστω R(A, B, C, D), $F = \{C \rightarrow D,\ A \rightarrow B,\ B \rightarrow D,\ C \rightarrow A\}$.
Έστω η αποσύνθεση S(A, C) και T(A, B, D)

*Μερικά ακόμα παραδείγματα:*

Έστω R(A, B, C, D, E), $F = \{A \rightarrow D,\ B \rightarrow E,\ DE \rightarrow C,\ B \rightarrow C\}$.

(α) Η αποσύνθεση του R σε S(A, B, C) και T(A, B, D, E) διατηρεί τις εξαρτήσεις;

(β) Είναι χωρίς απώλειες (lossless join);

## Παράδειγμα

Έστω $R(A, B, C, D)$, $F = \{A \rightarrow C, \ B \rightarrow C, BD \rightarrow A\}$ και η αποσύνθεση του R σε $R_1(A, C)$ και $R_2(A, B, D)$.

(α) Διατηρεί τις εξαρτήσεις;

(β) Είναι χωρίς απώλειες (lossless join);

## Επιθυμητές Ιδιότητες Αποσύνθεσης

### 1. Συνενώσεις Άνευ Απωλειών

Η φυσική συνένωση των σχέσεων που προκύπτουν μας δίνει *ακριβώς* την αρχική σχέση (χωρίς επιπρόσθετες πλειάδες): $r = \pi_{R1}(r) * \pi_{R2}(r) * \ldots \pi_{Rn}(r)$

**$R1 \cap R2 \rightarrow R1$ ή $R1 \cap R2 \rightarrow R2$ ανήκει στο $F^+$,** *δηλαδή τα κοινά γνωρίσματα των δύο σχημάτων είναι κλειδί για τουλάχιστον ένα από τα δύο*

### 2. Διατήρηση Εξαρτήσεων

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

**$F' = F1 \cup F2 \ldots \cup Fn$, πρέπει $F'^+ = F^+$**

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