

Λογικός Σχεδιασμός Σχεσιακών Σχημάτων

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

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

Σχεδιασμός Σχεσιακών Σχημάτων

Σχεδιασμός καλών σχεσιακών σχημάτων

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

Σχεδιασμός Σχεσιακών Σχημάτων

Γενικές Κατευθύνσεις

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

Γενικές Κατευθύνσεις

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
-------	-----------	---------------

Γενικές Κατευθύνσεις

2. Πλεονασμός (επανάληψη πληροφορίας)

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 της ταινίας

Ταινία

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

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;

3. Αποφυγή τιμών null

Ηθοποιός

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

Ηθοποιός

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

Συζυγί-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

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

Ταινία

Ένας γενικός (θεωρητικός) τρόπος κατασκευής του σχήματος

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

Παράδειγμα

$R = \{\text{Τίτλος, Έτος, Διάρκεια, Είδος, Όνομα-Ηθοποιού, Διεύθυνση, Έτος-Γέννησης}\}$

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

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

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

- Πως μπορούμε να πάρουμε την αρχική σχέση; Μπορούμε να διασπάσουμε την R_2 με τον ίδιο τρόπο.

Τυπικός ορισμός

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

$R_1 = \{B_1, B_2, \dots, B_m\}$ και $R_2 = \{C_1, C_2, \dots, C_k\}$ τέτοια ώστε

- $\{A_1, A_2, \dots, A_n\} = \{B_1, B_2, \dots, B_m\} \cup \{C_1, C_2, \dots, C_k\}$ (διατήρηση γνωρισμάτων) *γνωρίσματα*
- Οι πλειάδες της $r_1(R_1)$ είναι η *προβολή των πλειάδων της $r(R)$* στα $\{B_1, B_2, \dots, B_m\}$ *πλειάδες*
- Οι πλειάδες της $r_2(R_2)$ είναι η *προβολή των πλειάδων της $r(R)$* στα $\{C_1, C_2, \dots, C_k\}$ *πλειάδες*

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 , συμβολισμός $R(r)$

$R(r)$	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

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 , συμβολισμός $r(R)$

$r(R)$	A	B	C
	1	2	3
	4	2	5

$r_1(R_1)$

A	B
1	2
4	2

$r_2(R_2)$

B	C
2	3
2	5

$R_1 \cap R_2 = B$

Μπορούμε να πάρουμε το αρχικό στιγμιότυπο;

Φυσική συνένωση $r_1 * r_2$

γνωρίσματα

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

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

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

στιγμιότυπα

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

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

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 , συμβολισμός $r(R)$ ή r

Έστω $r(R)$ και $r_i = \pi_{R_i}(r), \forall i = 1, \dots, n$ ---- $r \subseteq r_1 * r_2 * \dots * r_n$

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

• Δεν μπορούμε να πάρουμε την αρχική σχέση r από τα r_1 και r_2

Επιθυμητές Ιδιότητες για την Αποσύνθεση

1. Συνενώσεις Άνευ Απωλειών

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

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

Συνενώσεις Άνευ Απωλειών

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				
$r_1 \circ r_2$				A B C	1 2 3	1 2 5	4 2 3
r_1	A C	1 3	4 5				
r'_2	B C	2 3	2 5				
$r'_1 \circ r'_2 = ;$							

Συνενώσεις Άνευ Απωλειών

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

Συνενώσεις Άνευ Απωλειών

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

Τίτλος Έτος \rightarrow Διάρκεια

Τίτλος Έτος \rightarrow Είδος

Όνομα Ηθοποιού \rightarrow Διεύθυνση

Όνομα-Ηθοποιού \rightarrow Έτος

Γέννησης

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

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

$$R_1 \cap R_2 = \{\text{Τίτλος, Έτος}\}$$

Επιθυμητές Ιδιότητες Αποσύνθεσης

Επιθυμητές Ιδιότητες για την Αποσύνθεση

2. Διατήρηση Εξαρτήσεων

Στόχος: Για να ελέγχουμε ότι διατηρούνται οι Σ.Ε. όταν γίνονται τροποποιήσεις σε μία από τις σχέσεις $r_i(R_i)$ να αρκεί να ελέγξουμε τη συγκεκριμένη σχέση (δηλαδή, να μη χρειάζεται να υπολογίσουμε τις αρχικές σχέσεις - αποφυγή των συνενώσεων)

Διατήρηση Εξαρτήσεων

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

F_i περιορισμός του F στο R_i είναι το σύνολο όλων των συναρτησιακών εξαρτήσεων του F^* που περιέχουν μόνο γνωρίσματα του R_i .

Διατήρηση Εξαρτήσεων

Παράδειγμα: Υπολογισμός του περιορισμού του F σε ένα σχήμα

Εφαρμογή 1: Έστω $R(A, B, C, D)$, $F = \{A \rightarrow B, B \rightarrow C\}$. Περιορισμός του F στο $S(A, C)$

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

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

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

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

Παράδειγμα: Πως δείχνουμε αν μια διάσπαση διατηρεί τις εξαρτήσεις

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

Παράδειγματα

1. Έστω $R(A, B, C, D)$, $F = \{A \rightarrow C, B \rightarrow C, BD \rightarrow A\}$. Η αποσύνθεση του R σε $S(A, C)$ και $T(A, B, D)$ διατηρεί τις εξαρτήσεις :

2. Έστω $R(A, B, C, D, E)$, $F = \{A \rightarrow D, B \rightarrow E, DE \rightarrow C\}$. Η αποσύνθεση του R σε $S(A, B, C)$ και $T(A, B, D, E)$ διατηρεί τις εξαρτήσεις :

- Αποσύνθεση καθολικού σχήματος
- Επιθυμητές ιδιότητες
 - διατήρηση εξαρτήσεων
 - όχι απώλειες στη συνένωση
 - όχι επανάληψη πληροφορίας λόγω ΣΕ

- Συνέχεια: Κανονικές Μορφές
- BCNF
- 3NF

Κανονικές Μορφές

Σχεδιασμός καλών σχεσιακών σχημάτων

- Μη τυπικές - γενικές κατευθύνσεις
 1. Σημασιολογία
 2. Ελάττωση πλεονασμού
 3. Ελάττωση τιμών null
 4. Μη πλασματικές πλειάδες

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

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

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

← γνωρίσματα

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

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

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 τουλάχιστον ένα από τα δύο

2. Διατήρηση Εξαρτήσεων

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

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

3. Αποφυγή Επανάληψης Πληροφορίας, πως: Κανονικές Μορφές

Θα δούμε πάλι τη διατήρηση εξαρτήσεων:

Στόχος: Για να ελέγξουμε ότι διατηρούνται οι ΣΕ όταν γίνονται τροποποιήσεις σε μία από τις σχέσεις $r_i(R_i)$ να αρκεί να ελέγξουμε 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 $\{R_1, R_2, \dots, R_n\}$ μια αποσύνθεση του R .

F_i περιορισμός του F στο R_i : το σύνολο όλων των συναρτησιακών εξαρτήσεων του F^* που περιέχουν μόνο γνωρίσματα του R_i .

Παραδείγματα

1. 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 $R_2(A, B, D)$.

Θα υπολογίσουμε τα F_1 και F_2

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

Θα υπολογίσουμε τα F_1 και F_2

Διατήρηση Εξαρτήσεων (επανάληψη)

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

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

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

Διατήρηση Εξαρτήσεων

Παράδειγμα

1. 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;

Κανονικές Μορφές: Εισαγωγή

▪ Στόχος: Δοσμένου ενός σχήματος, αν είναι «καλό» ή χρειάζεται περαιτέρω διάσπαση.

Τις; Κανονικές μορφές.

▪ Ξέρουμε ότι αν ένα σχήμα είναι σε κάποια Κανονική Μορφή δεν υπάρχουν συγκεκριμένα προβλήματα

▪ Με φθίνουσα σειρά (από την πιο περιοριστική στη λιγότερο περιοριστική)

BCNF 3NF 2NF 1NF

▪ Βασίζεται σε Σ.Ε., οι Σ.Ε. έχουν σχέση με την επανάληψη πληροφορίας

Κανονικές Μορφές: Εισαγωγή

Πλεονασμός (επανάληψη πληροφορίας)

Ταινία

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

Τι συμβαίνει με το (πρωτεύον) κλειδί και τις συναρτησιακές εξαρτήσεις;

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, Όνομα-Ηθοποιού)

Η σχέση Ταινία δεν είναι σε BCNF

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

Για παράδειγμα η ΣΕ Τίτλος Έτος \rightarrow Διάρκεια

Παράδειγμα 2

Ταινία2 (Τίτλος, Έτος, Διάρκεια, Είδος)

Η σχέση Ταινία2 είναι σε BCNF

Παράδειγμα 3

Οποιαδήποτε σχέση με δύο γνωρίσματα είναι σε BCNF

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;

Παράδειγμα 1

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

Τίτλος Έτος \rightarrow Διάρκεια Είδος

Ταινία1(Τίτλος, Έτος, Διάρκεια, Είδος)

Ταινία2(Τίτλος, Έτος, Όνομα-Ηθοποιού)

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)

- 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 $R - Y$

η R_2 μπορεί να μην είναι σε BCNF

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 \rightarrow Πόλη

Κλειδιά:

Έργο Πόλη \rightarrow Κινηματογράφος

{Έργο, Πόλη}

{Κινηματογράφος, Έργο}

Boyce-Codd Κανονική Μορφή

Παίζει(Έργο, Κινηματογράφος, Πόλη)

Κλειδιά

Κινηματογράφος → Πόλη

{Έργο, Πόλη}

Έργο Πόλη → Κινηματογράφος

{Κινηματογράφος, Έργο}

Αποσύνθεση σε: R_1 {Κινηματογράφος, Πόλη} και R_2 {Κινηματογράφος, Έργο}

Κινηματογράφος	Πόλη	Κινηματογράφος	Έργο
Odeon-ABANA	Αθήνα	Odeon-ABANA	Μένας Αλέξαντρος
Village Center Μαρούσι	Αθήνα	Village Center Μαρούσι	Μένας Αλέξαντρος

Δε μπορώ κοιτάζοντας μόνο την R2 να δω ότι η εισαγωγή της δεύτερης πλειάδας παραβιάζει μια ΣΕ

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 του F
- Για κάθε α.μ. X μιας συναρτησιακής εξάρτησης του F_c έστω Y το σύνολο όλων των γνωρισμάτων A_i που εμφανίζονται στο δ.μ. μιας ΣΕ του F_c $X \rightarrow A_i$
νέα σχέση με γνωρίσματα $X \cup Y$
- Βάλε όλα τα γνωρίσματα που περισσεύουν (αν υπάρχουν) σε ένα νέο σχήμα

Τρίτη Κανονική Μορφή

Αλγόριθμος Αποσύνθεσης σε 3NF

- Απώλειες στη συνένωση;
- Διατήρηση εξαρτήσεων;

Τρίτη Κανονική Μορφή

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 $X \rightarrow Y$ ισχύει τουλάχιστον ένα από τα παρακάτω:

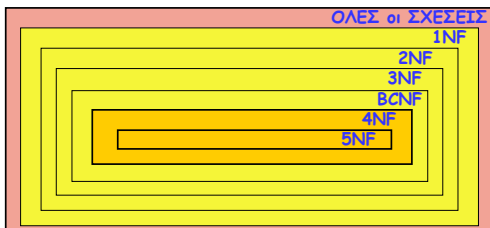
- $X \rightarrow Y$ είναι μια τετριμμένη ΣΕ ή
- X είναι υπερκλειδί του σχήματος R

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

Κανονικές Μορφές



Πρώτη Κανονική Μορφή

1NF (ιστορικοί λόγοι, κάθε γνώρισμα παίρνει ατομικές τιμές)

Πλειότιμες Εξαρτήσεις

Υπάρχει επανάληψη πληροφορίας που δεν μπορεί να εκφραστεί με απλές ΣΕ

Πλειότιμες Εξαρτήσεις

Πλειότιμες Εξαρτήσεις

Προκύπτουν όταν δυο γνώρισματα είναι ανεξάρτητα το ένα από το άλλο

Παράδειγμα

Ηθοποιός(Όνομα, Οδός, Πόλη, Τίτλος, Έτος)

Υποθέτουμε ότι για κάθε ηθοποιό είναι πιθανόν να υπάρχουν πολλές διευθύνσεις

Κανένα από τα 5 γνώρισματα δεν εξαρτάται συναρτησιακά από τα άλλα τέσσερα \Rightarrow δεν υπάρχουν μη τετριμμένες εξαρτήσεις \Rightarrow κλειδί ?

π.χ., Όνομα Οδός Τίτλος Έτος \rightarrow Πόλη δεν ισχύει

Πλειότεμες Εξαρτήσεις

Παράδειγμα (συνέχεια)

Ηθοποίος(Όνομα, Οδός, Πόλη, Τίτλος, Έτος)

Όλες οι εξαρτήσεις είναι τετριμμένες

Το σχήμα είναι σε BCNF αλλά υπάρχει επανάληψη πληροφορίας που δεν οφείλεται όμως σε συναρτησιακές εξαρτήσεις

Πλειότεμες Εξαρτήσεις

$$X \twoheadrightarrow Y$$

Για κάθε ζεύγος πλειάδων t_1 και t_2 της σχέσης R που συμφωνούν σε όλα τα γνωρίσματα του X μπορούμε να βρούμε στο R δυο πλειάδες t_3 και t_4 τέτοιες ώστε

• και οι δυο συμφωνούν με τις t_1 και t_2 στο X:

$$t_1[X] = t_2[X] = t_3[X] = t_4[X]$$

• η t_3 συμφωνεί με την t_1 στο Y: $t_3[Y] = t_1[Y]$

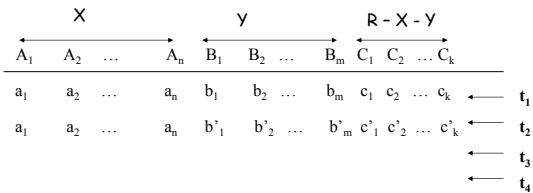
• η t_3 συμφωνεί με την t_2 στο R - X - Y: $t_3[R - X - Y] = t_2[R - X - Y]$

• η t_4 συμφωνεί με την t_2 στο Y: $t_4[Y] = t_2[Y]$

• η t_4 συμφωνεί με την t_1 στο R - X - Y: $t_4[R - X - Y] = t_1[R - X - Y]$

Πλειότεμες Εξαρτήσεις

$$A_1 A_2 \dots A_n \twoheadrightarrow B_1 B_2 \dots B_m$$



Πλειότεμες Εξαρτήσεις

Παράδειγμα

Ηθοποίος(Όνομα, Οδός, Πόλη, Τίτλος, Έτος)

Όνομα \twoheadrightarrow Οδός Πόλη

Όνομα	Οδός	Πόλη	Τίτλος	Έτος
C. Fisher	123 Mapple Str	Hollywood	Star Wars	1977
C. Fisher	5 Locust Ln	Malibu	Empire Strikes Back	1980
?				
?				

Σχεδιασμός Σχεσιακών Σχημάτων

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

Σχεδιασμός Σχεσιακών Σχημάτων

- 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 Σχεδιασμού

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