Σχεδιασμός Σχεσιακών Σχημάτων (3)

Αποσύνθεση (decomposition)

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

Επιθυμητές Ιδιότητες Αποσύνθεσης

1. Συνενώσεις Άνευ Απωλειών
2. Διατήρηση Εξαρτήσεων
3. Αποφυγή Επανάληψης Πληροφορίας

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

Πλειότιμες Εξαρτήσεις

Προκύπτουν όταν δυο γνωρίσματα είναι ανεξάρτητα το ένα από το άλλο

Παράδειγμα

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

Σημείωση: Για κάθε ηθοποιό είναι πιθανόν να υπάρχουν πολλές διευθύνσεις

Κανένα από τα 5 γνωρίσματα δεν εξαρτάται συναρτησιακά από τα άλλα τέσσερα ⇒ δεν υπάρχουν μη μη τετριμμένες εξαρτήσεις ⇒ κλειδί ?

π.χ., Όνομα Οδός Τίτλος Έτος → Πόλη δεν ισχύει

Παράδειγμα (συνέχεια)

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

Κανένα από τα 5 γνωρίσματα δεν εξαρτάται συναρτησιακά από τα άλλα τέσσερα ⇒ δεν υπάρχουν μη τετριμμένες εξαρτήσεις

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

Πλειότιμες 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:

  $t_1[X] = t_2[X] = t3[X] = t_4[X]$
- η t_3 συμφωνεί με την t_1 στο Y: **$t3[Y] = t_1[Y]$**
- η t_3 συμφωνεί με την t_2 στο R - X - Y: **$t3[R- X - Y] = t2[R - X - Y]$**
- η t_4 συμφωνεί με την t_2 στο Y: **$t4[Y] = t2[Y]$**
- η t_4 συμφωνεί με την t_1 στο R - X - Y: **$t4[R- X - Y] = t_1[R - X - Y]$**

Πλειότιμες Εξαρτήσεις

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

A_1	A_2	…	A_n	B_1	B_2	…	B_m	C_1	C_2	…	C_k	
a_1	a_2	…	a_n	b_1	b_2	…	b_m	c_1	c_2	…	c_k	← **t_1**
a_1	a_2	…	a_n	b'_1	b'_2	…	b'_m	c'_1	c'_2	…	c'_k	← **t_2**
												← **t_3**
												← **t_4**

Πλειότιμες Εξαρτήσεις

Παράδειγμα

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

Όνομα $\longrightarrow\!\!>$ 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

Κανόνες Συμπερασμού για Πλειότιμες Εξαρτήσεις

Κανόνες Συμπερασμού για Πλειότιμες Εξαρτήσεις

1. **Ανακλαστικός κανόνας** Αν $Y \subseteq X$, τότε $X \rightarrow Y$
2. **Επαυξητικός κανόνας** Αν $X \rightarrow Y$ τότε $XZ \rightarrow YZ$
3. **Μεταβατικός κανόνας** Αν $X \rightarrow Y$ και $Y \rightarrow Z$ τότε $X \rightarrow Z$
4. **Συμπληρωματικός κανόνας για ΠΕ** Αν $X \longrightarrow\!\!> Y$ τότε $X \longrightarrow\!\!>$ **R - X - Y**
5. **Επαυξητικός κανόνας για ΠΕ**

   Αν $X \longrightarrow\!\!> Y$ και $Z \subseteq W$ τότε $WX \longrightarrow\!\!>$ **Z** Y
6. **Μεταβατικός κανόνας για ΠΕ**

   Αν $X \longrightarrow\!\!> Y$ και $Y \longrightarrow\!\!> Z$ τότε $X \longrightarrow\!\!> Z - Y$

Κανόνες Συμπερασμού για Πλειότιμες Εξαρτήσεις

7. Κανόνας αντιγραφής

Αν $X \rightarrow Y$, τότε $X \twoheadrightarrow Y$

8. Κανόνας συγχώνευσης

Αν $X \twoheadrightarrow Y$, $Z \subseteq Y$ και

$\exists$ W τέτοιο ώστε: (α) $W \cap Y = \varnothing$ και (β) $W \rightarrow Z$

τότε $X \rightarrow Z$

Κανόνες Συμπερασμού για Πλειότιμες Εξαρτήσεις

- *βάσιμοι* (sound) δε δίνουν λανθασμένες εξαρτήσεις και *πλήρεις* (complete) μας δίνουν όλο το D^+

- *Απόδειξη του* **κανόνας αντιγραφής** (Αν $X \rightarrow Y$, τότε $X \twoheadrightarrow Y$)

  δηλαδή κάθε συναρτησιακή εξάρτηση είναι και πλειότιμη

  Οποτεδήποτε t1[X] = t2[X], υπάρχουν ...

Το ανάποδο ισχύει;

• *Παράδειγμα* **συμπληρωματικού κανόνα**

Αν $X \twoheadrightarrow Y$ τότε $X \twoheadrightarrow R - X - Y$

Όνομα $\twoheadrightarrow$ Οδός Πόλη

Όνομα $\twoheadrightarrow$ Τίτλος Έτος

• *Απόδειξη του συμπληρωματικού κανόνα*

Τέταρτη Κανονική Μορφή (4NF)

Ένα σχεσιακό σχήμα R είναι σε 4NF σε σχέση με ένα σύνολο F συναρτησιακών εξαρτήσεων αν για όλες τις ΣΕ και ΠΕ της μορφής

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

-- $X \twoheadrightarrow Y$ είναι μια τετριμμένη ΠΕ ή

-- X είναι υπερκλειδί του σχήματος R

• **τετριμμένη ΠΕ**; αν (α) το $Y \subseteq X$ ή (β) $X \cup Y = R$

• BCNF ? 3NF;

Αλγόριθμος Αποσύνθεσης σε 4NF

παρόμοιος με την BCNF αποσύνθεση

- Βρες μια μη τετριμμένη ΠΕ που παραβιάζει τον 4NF ορισμό, έστω $X \twoheadrightarrow Y$ και $X \cap Y = \varnothing$
- Αποσύνθεση του αρχικού σχήματος R σε δύο σχήματα

  R_1 με γνωρίσματα $X \cup Y$

  R_2 με γνωρίσματα $R - Y$

Παράδειγμα

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

Όνομα $\twoheadrightarrow$ Οδός Πόλη

Ηθοποιός1(Όνομα, Οδός, Πόλη)

Ηθοποιός2(Όνομα, Τίτλος. Έτος)

	3NF	BCNF	4NF
Αποφυγή επανάληψης πληροφορίας λόγω ΣΕ	συνήθως	ναι	ναι
Αποφυγή επανάληψης πληροφορίας λόγω ΠΕ	όχι	όχι	ναι
Διατήρηση ΣΕ	ναι	ίσως	ίσως
Διατήρηση ΠΕ	ίσως	ίσως	ίσως
Χωρίς απώλειες στη συνένωση	ναι	ναι	ναι

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)

Πρώτη Κανονική Μορφή

Παράδειγμα με **πλειότιμα**

ΤΜΗΜΑ(Τ_ΟΝΟΜΑ, ΚΩΔ_ΤΜΗΜ, ΔΙΕΥΘΥΝΤΗΣ_ΤΜΗΜ, ΤΟΠΟΘΕΣΙΕΣ)

Ένα τμήμα σε πολλές τοποθεσίες . Κλειδί;

ΚΩΔ_ΤΜΗΜΑΤΟΣ $\twoheadrightarrow$ ΤΟΠΟΘΕΣΙΕΣ

Αποσύνθεση σε ΤΜΗΜΑ1(ΚΩΔ_ΤΜΗΜΑΤΟΣ, ΤΟΠΟΘΕΣΙΕΣ)
ΤΜΗΜΑ2(Τ_ΟΝΟΜΑ, ΚΩΔ_ΤΜΗΜ, ΔΙΕΥΘΥΝΤΗΣ_ΤΜΗΜ)

Δεύτερη Κανονική Μορφή

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 Α από το Χ 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

Δεύτερη Κανονική Μορφή (2NF)

Ένα σχήμα σχέσης είναι σε 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 → ΕΡ_ΟΝΟΜΑ ΑΡ_ΤΑΥΤ ΚΩΔ_ΕΡΓΟΥ → ΩΡΕΣ

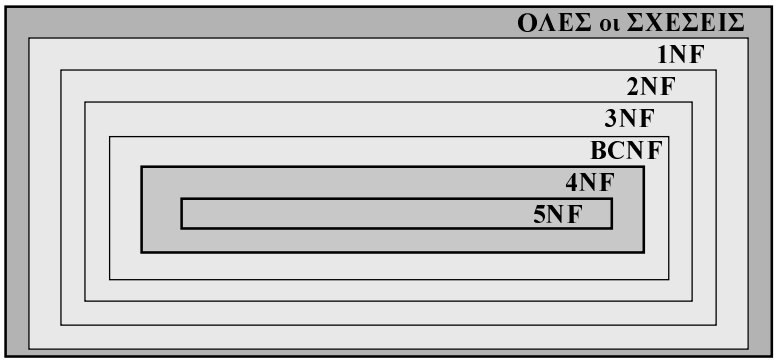
Μια συναρτησιακή εξάρτηση X → Y σε ένα σχήμα σχέσης R είναι μια **μεταβατική εξάρτηση** αν υπάρχει ένα σύνολο γνωρισμάτων **Z** που δεν είναι υποσύνολο οποιουδήποτε κλειδιού της R τέτοιο ώστε να ισχύουν X → **Z** και **Z** → Y.

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



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)

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 αξιέπαινη 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