



Κανονικοποίηση

Χρυσόστομος Καπέτης



- Ιδιότητες Σχεσιακού περιβάλλοντος
- Εξαρτήσεις και κανονικοποίηση
- Επαναλαμβανόμενα γνωρίσματα (γνωρίσματα πολλαπλών τιμών)
- Συναρτησιακή και μεταβατική εξάρτηση
- Κανονικές μορφές



Ιδιότητες Σχεσιακού Περιβάλλοντος

Κανονικοποίηση

- Κάθε στήλη ενός πίνακα αντιπροσωπεύει ένα και μόνο μοναδικό γνώρισμα μιας οντότητας του υπό ανάλυση οργανισμού.
- Κάθε στήλη έχει ένα μοναδικό όνομα μέσα στον ίδιο πίνακα, το οποίο αποτελεί το αναγνωριστικό αυτής.
- Κάθε στήλη περιλαμβάνει μόνο ομοιογενείς τιμές. Με άλλα λόγια κάτω από κάθε στήλη καταχωρίζονται τιμές που βασίζονται στο ίδιο εννοιολογικό αντικείμενο.
- Κάθε στήλη αντλεί τις τιμές της από τον ίδιο τύπο δεδομένων ο οποίος προσδιορίζεται με ακρίβεια και ισχύει για τον συγκεκριμένο πίνακα.
- Κάθε στήλη αντιπροσωπεύει ένα ανεξάρτητο αντικείμενο του πραγματικού κόσμου, το οποίο όμως δεν μπορεί να είναι αποτέλεσμα σύνθεσης, συγχώνευσης, υπολογισμού ή άλλης πράξης επεξεργασίας άλλων γνωρισμάτων.



Ιδιότητες Σχεσιακού Περιβάλλοντος

Κανονικοποίηση

- Κάθε πλειάδα ενός πίνακα αντιπροσωπεύει μια εγγραφή μιας οντότητας του υπό ανάλυση οργανισμού. Μια πλειάδα περιλαμβάνει τιμές κάτω από όλες τις στήλες του πίνακα και θεωρείται ακέραια όταν δεν παραβιάζει τους περιορισμούς που ισχύουν για τον πίνακα στον οποίο ανήκει.
- Μία εγγραφή στην τομή κάθε γραμμής και στήλης είναι ατομική (αποτελείται μόνο από μία τιμή). Σε μια συγκεκριμένη γραμμή ενός πίνακα μπορεί να υπάρχει μόνο μία τιμή που σχετίζεται με κάθε γνώρισμα (στήλη). Σε έναν πίνακα δεν επιτρέπονται γνωρίσματα πολλαπλών τιμών.
- Η σειρά των στηλών και των εγγραφών του πίνακα δεν επηρεάζουν τη βάση δεδομένων.
- Διπλότυπες εγγραφές σε ένα πίνακα βάσης δεν επιτρέπονται παρόλο που μπορεί ένας προσωρινός πίνακας ή όψη να περιέχει διπλότυπες εγγραφές.
- Κάθε πίνακας έχει ένα μοναδικό όνομα, καθώς επίσης και ένα μοναδικό πρωτεύον κλειδί.



Εξαρτήσεις και Κανονικοποίηση

Κανονικοποίηση

- Αφού ελέγχουμε τις παραπάνω δομικές ιδιότητες των οντοτήτων προχωρούμε στην εφαρμογή της κανονικοποίησης τα κριτήρια της οποίας είναι:
 - Η **συναρτησιακή εξάρτηση** (Functional Dependence – FD)
 - Η **μεταβατική εξάρτηση** (Transitive Dependence –TD)
 - Η **εξάρτηση πολλαπλών τιμών** (Multivalued Dependence – MD)



Εξαρτήσεις και Κανονικοποίηση

Κανονικοποίηση

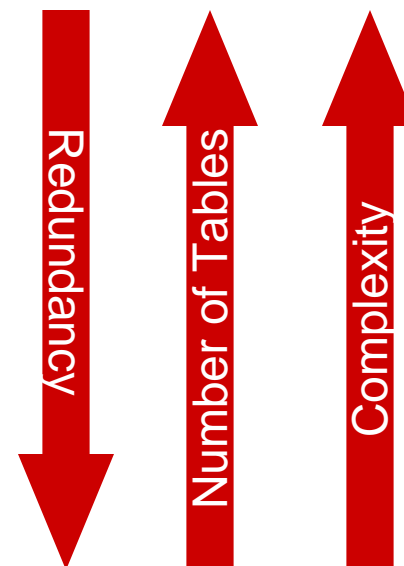
- Σε κάθε στάδιο εφαρμογής των προαναφερόμενων κριτηρίων έχουμε ένα σύνολο από σχέσεις (πίνακες) σε μία **κανονική μορφή (Normal Form –NF)**. Υπάρχουν πέντε κανονικές μορφές που μπορεί να έχει μία σχέση:
 - 1^η κανονική μορφή (1NF), η οποία ασχολείται κυρίως με τις δομικές ιδιότητες των σχέσεων και με τα επαναλαμβανόμενα γνωρίσματα.
 - 2^η κανονική μορφή (2NF), η οποία ασχολείται με την πλήρη συναρτησιακή εξάρτηση, τις περιπτώσεις δηλαδή που το πρωτεύον κλειδί είναι σύνθετο.
 - 3^η κανονική μορφή (3NF), η οποία ασχολείται με την μεταβατική εξάρτηση και τις περιπτώσεις όπου υπάρχει επικάλυψη μεταξύ εναλλακτικών συνδυασμών κλειδιών.
 - 4^η κανονική μορφή (4NF), η οποία ασχολείται με την εξάρτηση πολλαπλών τιμών.
 - 5^η κανονική μορφή (5NF) η οποία ασχολείται με την εξάρτηση πολλαπλών τιμών, καθώς και με τις περιπτώσεις που μία σχέση μπορεί να δημιουργηθεί με τον τελεστή σύνδεσης και όπου κάθε υποκείμενη σχέση έχει μικρότερο βαθμό από την υπό ανάλυση σχέση.



Κανονινές Μορφές

Κανονικοποίηση

- Οι κανονικές μορφές είναι οι εξής:
 - First Normal Form (1NF)
 - Second Normal Form (2NF)
 - Third Normal Form (3NF)
 - Boyce-Codd Normal Form (BCNF)
 - Fourth Normal Form (4NF)
 - Fifth Normal Form (5NF)



Οι περισσότερες βάσεις πρέπει να είναι στην 3NF (ή BCNF) προκειμένου να αποφευχθούν οι ανωμαλίες στην διαχείριση των δεδομένων.



Επαναλαμβανόμενα Γνωρίσματα

Κανονικοποίηση

- **Επαναλαμβανόμενο γνώρισμα ή γνώρισμα πολλαπλών τιμών:** Ένα γνώρισμα το οποίο λαμβάνει περισσότερες από μία τιμές για κάθε στιγμιότυπο μιας οντότητας.
- Ο σχεδιαστής της βάσης πρέπει να εντοπίζει και να απομακρύνει όλα τα επαναλαμβανόμενα γνωρίσματα από μία σχέση, μεταφέροντάς τα σε μια άλλη ανεξάρτητη σχέση, διατηρώντας παράλληλα μια συσχέτιση με τη μητρική σχέση.



Επαναλαμβανόμενα Γνωρίσματα - Παράδειγμα

Κανονικοποίηση

- Έστω η οντότητα ΒΙΒΛΙΟ, η οποία περιλαμβάνει τα ακόλουθα γνωρίσματα: Κωδικός_βιβλίου#, Τίτλος, Έτος, Κωδικός_εκδότη και Ονομα_συγγραφέα.
 - Το γνώρισμα Ονομα_συγγραφέα είναι επαναλαμβανόμενο γνώρισμα, επειδή για κάθε πλειάδα (δηλαδή βιβλίο) μπορεί να έχουμε ένα, δύο ή περισσότερους συγγραφείς.
 - Κατά συνέπεια, για κάθε σειρά τιμών κάτω από τα ακόλουθα γνωρίσματα: <Κωδικός_βιβλίου#, Τίτλος, Έτος, Ονομα_εκδότη, Ονομα_συγγραφέα>, τα οποία εκπροσωπούν ένα βιβλίο θα έχουμε υποχρεωτικά τόσες διαφορετικές τιμές κάτω από το γνώρισμα Ονομα_συγγραφέα, όσοι είναι οι συγγραφείς του βιβλίου



Συναρτησιακή Εξάρτηση

Κανονικοποίηση

- **Απλή Συναρτησιακή Εξάρτηση (FD):** είναι μία συσχέτιση μεταξύ δύο γνωρισμάτων – έστω α και β - μιας οντότητας, όπου σε οποιαδήποτε χρονική στιγμή υπάρχει μόνο μια λογικά αποδεκτή τιμή κάτω από το β , η οποία αντιστοιχεί σε μία δεδομένη τιμή κάτω από το α .
 - Συμβολισμός: $\alpha \rightarrow \beta$ (ισχύει απλή συναρτησιακή εξάρτηση)
 $\alpha \not\rightarrow \beta$ (δεν ισχύει απλή συναρτησιακή εξάρτηση)
- **πλήρη Συναρτησιακή Εξάρτηση (FD):** πλήρη συναρτησιακή εξάρτηση μεταξύ γνωρισμάτων ισχύει, εφόσον κανένα υποσύνολο του α δεν παρέχει συναρτησιακή εξάρτηση (απλή ή πλήρη) στο β .
 - Συμβολισμός: $\alpha+ \Rightarrow \beta$ (ισχύει πλήρη συναρτησιακή εξάρτηση)
 $\alpha+ \not\Rightarrow \beta$ (δεν ισχύει πλήρη συναρτησιακή εξάρτηση)
 - προφανώς, η πλήρη συναρτησιακή εξάρτηση εξετάζεται μόνο εφόσον το αριστερό σκέλος (α) είναι σύνθετο (συμβολικά $\alpha+$), δηλαδή αποτελείται από δύο ή περισσότερα γνωρίσματα.



Συναρτησιακή Εξάρτηση - Παραδείγματα

Κανονικοποίηση

- Με δεδομένο τον κωδικό ενός βιβλίου μπορούμε να εντοπίσουμε τον τίτλο αυτού:

Κωδικός_βιβλίου# \rightarrow Τίτλος

- Με δεδομένο το όνομα του συγγραφέα, δεν είμαστε σίγουροι ότι θα εντοπίσουμε τον επιθυμητό τίτλο, αφού ένας συγγραφέας μπορεί να έχει συγγράψει δύο ή περισσότερα βιβλία. Κατά συνέπεια έχουμε την ακόλουθη εξάρτηση:

Όνομα_συγγραφέα / \rightarrow Τίτλος



Συναρτησιακή Εξάρτηση - Παραδείγματα

Κανονικοποίηση

- Έστω η παρακάτω οντότητα:

παραγγελία(Κωδικός_παραγγελίας#, Κωδικός_αντικειμένου#,
περιγραφή_αντικειμένου, ποσότητα, Τιμή_μονάδας).

Υποθέτουμε ότι:

- Κάθε κωδικός αντικείμενου αντιστοιχεί πάντα σε συγκεκριμένο αντικείμενο και σε σχετική περιγραφή αυτού.
 - Σε κάθε κωδικό παραγγελίας είναι δυνατόν να αντιστοιχιστούν ένα ή περισσότερα διαφορετικά αντικείμενα – το καθένα με δικό του κωδικό, ποσότητα και τιμή μονάδας.
 - Κάθε φορά που παραγγέλλεται ένα αντικείμενο είναι δυνατό να ζητείται διαφορετική ποσότητα, αλλά και με διαφορετική τιμή .
- Από τις παραπάνω υποθέσεις προκύπτουν διάφορες εξαρτήσεις μεταξύ των οποίων και η ακόλουθη πλήρης συναρτησιακή εξάρτηση:
 - Κωδικός_παραγγελίας#, Κωδικός_αντικειμένου# ==> ποσότητα, Τιμή_μονάδαςαφού ισχύουν:
 - Κωδικός_παραγγελίας# /→ ποσότητα, τιμή_μονάδας
 - Κωδικός_αντικειμένου# /→ ποσότητα, Τιμή_μονάδας



Μεταβατική Εξάρτηση

Κανονικοποίηση

- **Μεταβατική Εξάρτηση (TD):** Μεταβατική εξάρτηση είναι μία συσχέτιση μεταξύ τουλάχιστον τριών διαφορετικών γνωρισμάτων – έστω α, β και γ – μιας οντότητας, όπου το γ είναι συναρτησιακά εξαρτημένο από το α μέσω του β , ενώ παράλληλα το α δεν είναι συναρτησιακά εξαρτημένο από το β .
 - Συμβολισμός: $\alpha \rightarrow \beta \rightarrow \gamma$ και $\beta \not\rightarrow \alpha$



Μεταβατική Εξάρτηση - Παραδείγματα

Κανονικοποίηση

- Έστω η παρακάτω σχέση:

παραγγελία (Κωδικός_παραγγελίας#, ποσότητα, Τιμή,
Κωδικός_αντικειμένου, περιγραφή_αντικειμένου)

Υποθέτουμε ότι:

- Σε κάθε κωδικό παραγγελίας αντιστοιχεί πάντοτε μία ποσότητα, μία τιμή ένας κωδικός αντικειμένου και μια περιγραφή αντικειμένου
 - Σε κάθε κωδικό αντικειμένου αντιστοιχεί πάντοτε μία συγκεκριμένη περιγραφή.
- Η παραπάνω σχέση εμφανίζει μεταξύ άλλων εξαρτήσεων και την παρακάτω μεταβατική εξάρτηση:
 - $\text{Κωδικός_παραγγελίας\#} \rightarrow \text{Κωδικός_αντικειμένου} \rightarrow \text{περιγραφή_αντικειμένου}$
ενώ παράλληλα έχουμε:
 $\text{Κωδικός_αντικειμένου} / \rightarrow \text{Κωδικό_παραγγελίας\#}$



Μεταβατική Εξάρτηση - Παράδειγμα

Κανονικοποίηση

- Έστω η παρακάτω σχέση:
Εργο (Κωδικός_έργου#, προϋπολογισμός, Κωδικός_τμήματος,
περιγραφή_τμήματος)
 - Σε κάθε κωδικό έργου αντιστοιχεί πάντοτε ένας προϋπολογισμός, ένας κωδικός τμήματος και μία περιγραφή
 - Σε κάθε κωδικό τμήματος αντιστοιχεί πάντοτε μία συγκεκριμένη περιγραφή.
- Η παραπάνω σχέση εμφανίζει μεταξύ άλλων εξαρτήσεων και την παρακάτω μεταβατική εξάρτηση:
 - Κωδικός_έργου# \rightarrow Κωδικός_τμήματος \rightarrow περιγραφή_τμήματοςενώ παράλληλα έχουμε:
Κωδικός_τμήματος $/\rightarrow$ Κωδικό_έργου#



Πρώτη Κανονική Μορφή (1NF)

Κανονικοποίηση

Μια σχέση R είναι στην **πρώτη κανονική μορφή (1NF)** αν δεν περιέχει ούτε σύνθετα ούτε επαναλαμβανόμενα γνωρίσματα (γνωρίσματα πολλαπλών τιμών) .

- **Ενέργειες κανονικοποίησης (1NF)**
 1. Αποσυνθέτουμε τα σύνθετα γνωρίσματα σε γνωρίσματα που δέχονται απλές τιμές.
 2. Διερευνούμε τη σχέση R για υποψήφια κλειδιά και επιλέγουμε ένα πρωτεύον κλειδί (PK).
 3. Εντοπίζουμε κάθε γνώρισμα X που εμφανίζει περισσότερες από μία τιμές για δεδομένη τιμή του PK.
 4. Για κάθε X δημιουργούμε μια επιπλέον σχέση R' που περιλαμβάνει το X και το PK της R (αποτελούν σύνθετο κλειδί).
 5. Διαγράφουμε το X από την R .



Πρώτη Κανονική Μορφή (1NF)

Κανονικοποίηση

Παράδειγμα (Not 1NF)

ISBN	Title	AuName	AuPhone	PubName	PubPhone	Price
0-321-32132-1	Balloon	Sleepy, Snoopy, Grumpy	321-321-1111, 232-234-1234, 665-235-6532	Small House	714-000-0000	\$34.00
0-55-123456-9	Main Street	Jones, Smith	123-333-3333, 654-223-3455	Small House	714-000-0000	\$22.95
0-123-45678-0	Ulysses	Joyce	666-666-6666	Alpha Press	999-999-9999	\$34.00
1-22-233700-0	Visual Basic	Roman	444-444-4444	Big House	123-456-7890	\$25.00

Επαναλαμβανόμενα γνωρίσματα: AuName, AuPhone



1NF-Αποσύνθεση

Κανονικοποίηση

1.

Παράδειγμα (1NF)

ISBN	Title	PubName	PubPhone	Price
0-321-32132-1	Balloon	Small House	714-000-0000	\$34.00
0-55-123456-9	Main Street	Small House	714-000-0000	\$22.95
0-123-45678-0	Ulysses	Alpha Press	999-999-9999	\$34.00
1-22-233700-0	Visual Basic	Big House	123-456-7890	\$25.00

ISBN	AuName	AuPhone
0-321-32132-1	Sleepy	321-321-1111
0-321-32132-1	Snoopy	232-234-1234
0-321-32132-1	Grumpy	665-235-6532
0-55-123456-9	Jones	123-333-3333
0-55-123456-9	Smith	654-223-3455
0-123-45678-0	Joyce	666-666-6666
1-22-233700-0	Roman	444-444-4444



1NF - Παραδείγματα

Κανονικοποίηση

- Έστω η ακόλουθη σχέση R2 με πρωτεύον κλειδί το γνώρισμα Personal_code#:

R2(Personal_code#, name, address, telephone, salary_history: date, salary, job_title)

Είναι εμφανές ότι η σχέση R2 περιέχει την επαναλαμβανομένη ομάδα γνωρισμάτων:

Salary_history(Date, Salary, job_title)

Μέσω της οποίας όλες οι αλλαγές στο μισθό ενός υπαλλήλου καταγράφονται κάτω από τις στήλες ημερομηνία, μισθός και θέση εργασίας.

- Για να απαλλάξουμε την R2 από την επαναλαμβανόμενη ομάδα, χωρίζουμε τη σχέση στις ακόλουθες δύο:

R21(Personal_code#, Name, Address, Telephone)

R22(Personal_code#, Date#, salary, job_title)



1NF - Παραδείγματα

- Στη R21 ισχύουν οι εξαρτήσεις:

Personal_code# \rightarrow Name, Address, Telephone

Άρα η R21 είναι στην 1NF

- Στη R22 ισχύουν οι εξαρτήσεις:

Personal_code# $/\rightarrow$ date, salary, job_title

Date# $/\rightarrow$ personal_code, salary, job_title

Personal_code#, date# \Rightarrow salary, job_title

Άρα η R22 είναι στην 1NF.



Δεύτερη Κανονική Μορφή (2NF)

Κανονικοποίηση

Μια σχέση R είναι στη δεύτερη κανονική μορφή (2NF) αν:

- είναι στην πρώτη κανονική μορφή και
- το πρωτεύον κλειδί (PK) της R παρέχει ΠΛΗΡΗ συναρτησιακή εξάρτηση σε καθένα από τα υπόλοιπα γνωρίσματα ($PK \Rightarrow$).

- **Ενέργειες κανονικοποίησης (1NF→2NF)**
 1. Αν το PK είναι απλό η R βρίσκεται ήδη στη δεύτερη κανονική μορφή.
 2. Αν το PK είναι σύνθετο αλλά για κάθε υποσύνολο του έστω $S \subset PK$ ισχύει $S \rightarrow A_i$ τότε η R βρίσκεται και πάλι στη δεύτερη κανονική μορφή (τηρείται η πλήρης συναρτησιακή εξάρτηση από το κλειδί).
 3. Αν υπάρχει $S \rightarrow A_i$ τότε δημιουργούμε μια επιπλέον σχέση R' που περιλαμβάνει μόνο το S και το A_i και έχει το S πρωτεύον κλειδί. Επίσης διαγράφουμε το A_i από την R .



Δεύτερη Κανονική Μορφή - Παραδείγματα

Κανονικοποίηση

Παράδειγμα 1 (Not 2NF)

Scheme \rightarrow {Title, PubId, AuId, Price, AuAddress}

1. Key \rightarrow {Title, PubId, AuId}
2. {Title, PubId, AuId} \rightarrow {Price}
3. {AuId} \rightarrow {AuAddress}
4. AuAddress, δεν αποτελεί μέρος του κλειδιού
5. AuAddress, εξαρτάται (συναρητσιακή εξάρτηση) από το πεδίο AuId το οποίο αποτελεί μέρος του κλειδιού

Παράδειγμα 1 (Convert to 2NF)

Old Scheme \rightarrow {Title, PubId, AuId, Price, AuAddress}

New Scheme \rightarrow {Title, PubId, AuId, Price}

New Scheme \rightarrow {AuId, AuAddress}



Δεύτερη Κανονική Μορφή - Παραδείγματα

Κανονικοποίηση

Παράδειγμα 2 (Not 2NF)

Scheme \rightarrow {City, Street, HouseNumber, HouseColor, CityPopulation}

1. key \rightarrow {City, Street, HouseNumber}
2. {City, Street, HouseNumber} \rightarrow {HouseColor}
3. {City} \rightarrow {CityPopulation}
4. CityPopulation δεν αποτελεί μέρος του κλειδιού.
5. CityPopulation εξαρτάται από το πεδίο City το οποίο αποτελεί μέρος του κλειδιού.

Παράδειγμα 2 (Convert to 2NF)

Old Scheme \rightarrow {City, Street, HouseNumber, HouseColor, CityPopulation}

New Scheme \rightarrow {City, Street, HouseNumber, HouseColor}

New Scheme \rightarrow {City, CityPopulation}



Δεύτερη Κανονική Μορφή - Παραδείγματα

Κανονικοποίηση

Example 3 (Not 2NF)

Scheme \rightarrow {studio, movie, budget, studio_city}

1. Key \rightarrow {studio, movie}
2. {studio, movie} \rightarrow {budget}
3. {studio} \rightarrow {studio_city}
4. studio_city δεν αποτελεί μέρος του κλειδίου
5. studio_city εξαρτάται από το πεδίο studio το οποίο αποτελεί μέρος του κλειδίου.

Example 3 (Convert to 2NF)

Old Scheme \rightarrow {Studio, Movie, Budget, StudioCity}

New Scheme \rightarrow {Movie, Studio, Budget}

New Scheme \rightarrow {Studio, City}



Τρίτη Κανονική Μορφή (3NF)

Κανονικοποίηση

Μια σχέση R είναι στην τρίτη κανονική μορφή (3NF) αν:

- είναι στη δεύτερη κανονική μορφή και
- για κάθε (μη ασήμαντη) συναρτησιακή εξάρτηση $X \rightarrow Y$ είτε το X είναι υπερκλειδί της R είτε το Y είναι μέρος κάποιου υποψήφιου κλειδιού της R .

- **Ενέργειες κανονικοποίησης (2NF \rightarrow 3NF)**

1. Αν για κάθε $X \rightarrow Y$ που ισχύει στη σχέση R το X είναι υπερκλειδί της, τότε η σχέση βρίσκεται ήδη στην τρίτη κανονική μορφή.
2. Αν για κάθε $X \rightarrow Y$ που ισχύει στη σχέση R το Y είναι μέρος κάποιου υποψήφιου κλειδιού της, τότε η σχέση βρίσκεται ήδη στην τρίτη κανονική μορφή.
3. Αν υπάρχει $X \rightarrow Y$ που δεν ισχύει τίποτα από τα παραπάνω τότε δημιουργούμε μια επιπλέον σχέση R' που περιλαμβάνει μόνο τα X και Y και έχει το X πρωτεύον κλειδί. Επίσης διαγράφουμε το Y από την R .



Τρίτη Κανονική Μορφή - Παραδείγματα

Κανονικοποίηση

Παράδειγμα 1 (Not in 3NF)

Scheme \rightarrow {Studio, StudioCity, CityTemp}

1. Primary Key \rightarrow {Studio}
2. {Studio} \rightarrow {StudioCity}
3. {StudioCity} \rightarrow {CityTemp}
4. {Studio} \rightarrow {CityTemp}
5. Τα πεδία StudioCity και CityTemp εξαρτώνται από το κλειδί (2NF).
6. Μεταξύ των πεδίων CityTemp και Studio υπάρχει μεταβατική εξάρτηση οπότε και παραβιάζεται η 3NF

Παράδειγμα 1 (Convert to 3NF)

Old Scheme \rightarrow {Studio, StudioCity, CityTemp}

New Scheme \rightarrow {Studio, StudioCity}

New Scheme \rightarrow {StudioCity, CityTemp}



Τρίτη Κανονική Μορφή - Παραδείγματα

Κανονικοποίηση

Παράδειγμα 2 (Not in 3NF)

Scheme \rightarrow {BuildingID, Contractor, Fee}

1. Primary Key \rightarrow {BuildingID}
2. {BuildingID} \rightarrow {Contractor}
3. {Contractor} \rightarrow {Fee}
4. {BuildingID} \rightarrow {Fee}
5. Μεταβατική εξάρτηση μεταξύ των πεδίων Fee και BuildingID
6. Τα πεδία Contractor και Fee εξαρτώνται από ολόκληρο το κλειδί (2NF).

BuildingID	Contractor	Fee
100	Randolph	1200
150	Ingersoll	1100
200	Randolph	1200
250	Pitkin	1100
300	Randolph	1200

Παράδειγμα 2 (Convert to 3NF)

Old Scheme \rightarrow {BuildingID, Contractor, Fee}

New Scheme \rightarrow {BuildingID, Contractor}

New Scheme \rightarrow {Contractor, Fee}



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

PVFC Customer Invoice

Customer ID	2	Order ID	1006
Customer Name	Value Furniture	Order Date	24/10/2010
Address	15145 S.W. 17th St. Plano TX 75022		

Product ID	Product Description	Finish	Quantity	Unit Price	Extended Price
7	Dining Table	Natural Ash	2	\$800,00	\$1.600,00
5	Writer's Desk	Cherry	2	\$325,00	\$650,00
4	Entertainment Center	Natural Maple	1	\$650,00	\$650,00
				Total	\$2.900,00



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

<u>OrderID</u>	Order Date	Customer ID	Customer Name	Customer Address	<u>ProductID</u>	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2010	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
					5	Writer's Desk	Cherry	325.00	2
					4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2010	6	Furniture Gallery	Boulder, CO	11	4-Dr Dresser	Oak	500.00	4
					4	Entertainment Center	Natural Maple	650.00	3

FIGURE 4-25 INVOICE data (Pine Valley Furniture Company)



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

Πίνακας με γνωρίσματα πολλαπλών τιμών, ο οποίος δεν είναι στην πρώτη κανονική μορφή.

<u>OrderID</u>	Order Date	Customer ID	Customer Name	Customer Address	<u>ProductID</u>	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2010	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
					5	Writer's Desk	Cherry	325.00	2
					4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2010	6	Furniture Gallery	Boulder, CO	11	4-Dr Dresser	Oak	500.00	4
					4	Entertainment Center	Natural Maple	650.00	3

FIGURE 4-25 INVOICE data (Pine Valley Furniture Company)



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

Ο παρακάτω πίνακας (INVOICE) είναι στην πρώτη κανονική μορφή

<u>OrderID</u>	Order Date	Customer ID	Customer Name	Customer Address	<u>ProductID</u>	Product Description	Product Finish	Product StandardPrice	Ordered Quantity
1006	10/24/2010	2	Value Furniture	Plano, TX	7	Dining Table	Natural Ash	800.00	2
1006	10/24/2010	2	Value Furniture	Plano, TX	5	Writer's Desk	Cherry	325.00	2
1006	10/24/2010	2	Value Furniture	Plano, TX	4	Entertainment Center	Natural Maple	650.00	1
1007	10/25/2010	6	Furniture Gallery	Boulder, CO	11	4-Dr Dresser	Oak	500.00	4
1007	10/25/2010	6	Furniture Gallery	Boulder, CO	4	Entertainment Center	Natural Maple	650.00	3

FIGURE 4-26 INVOICE relation (1NF) (Pine Valley Furniture Company)



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

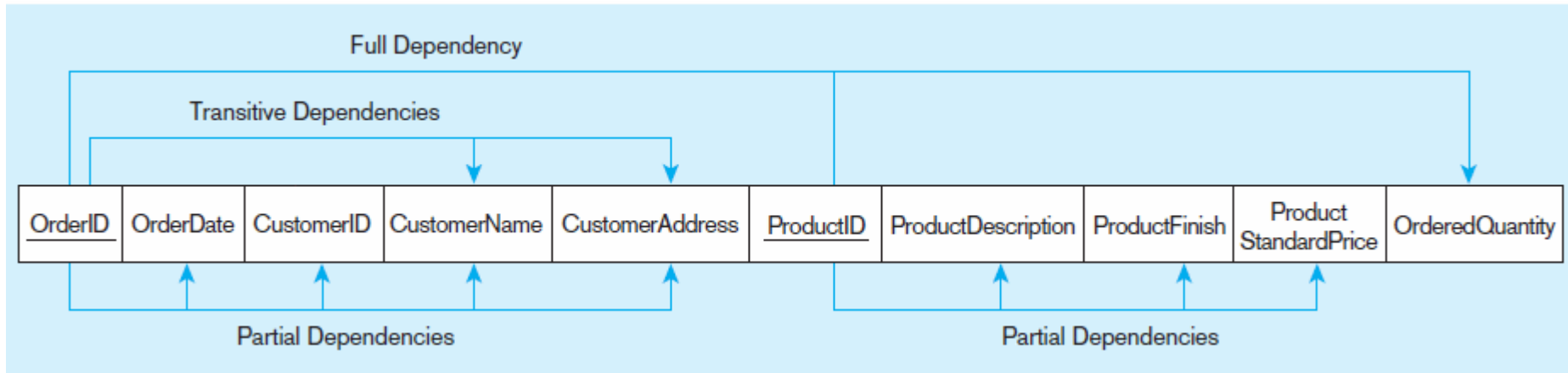
- Ο παραπάνω πίνακας περιέχει στοιχεία για παραπάνω από μία οντότητες. Το γεγονός αυτό οδηγεί σε πλεονασμό και μη επιθυμητές εξαρτήσεις μεταξύ των οντοτήτων. Ως αποτέλεσμα εμφανίζονται ορισμένες ανωμαλίες κατά τις λειτουργίες εισαγωγής, διαγραφής και ενημέρωσης των δεδομένων.
 - **Εισαγωγή:** Η παραγγελία ενός νέου προϊόντος από έναν υπάρχοντα πελάτη απαιτεί την επανα-εισαγωγή των στοιχείων του πελάτη.
 - **Διαγραφή:** Η διαγραφή ενός προϊόντος από μία παραγγελία μπορεί να προκαλέσει την απώλεια πληροφοριών (τιμή, φινίρισμα) για το συγκεκριμένο προϊόν.
 - **Ενημέρωση:** Η ενημέρωση της τιμής ενός προϊόντος απαιτεί την ενημέρωση πολλών εγγραφών του πίνακα.



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

Διάγραμμα Συναρτησιακών εξαρτήσεων του πίνακα INVOICE



OrderID → OrderDate, CustomerID, CustomerName, CustomerAddress

CustomerID → CustomerName, CustomerAddress

ProductID → ProductDescription, ProductFinish, ProductStandardPrice

OrderID, ProductID → OrderQuantity

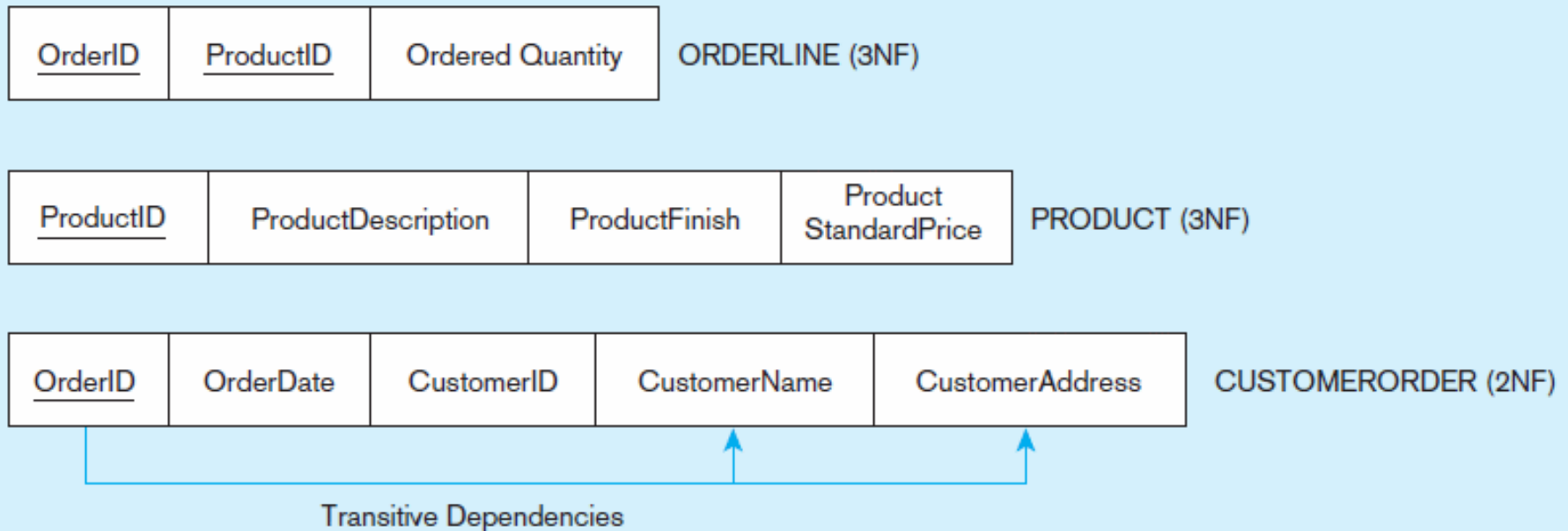
Ο πίνακας INVOICE δεν είναι στην δεύτερη Κανονική Μορφή (2NF)



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

Αφαίρεση μερικών εξαρτήσεων (partial Dependencies) ώστε να πάμε στην δεύτερη κανονική μορφή (2NF).



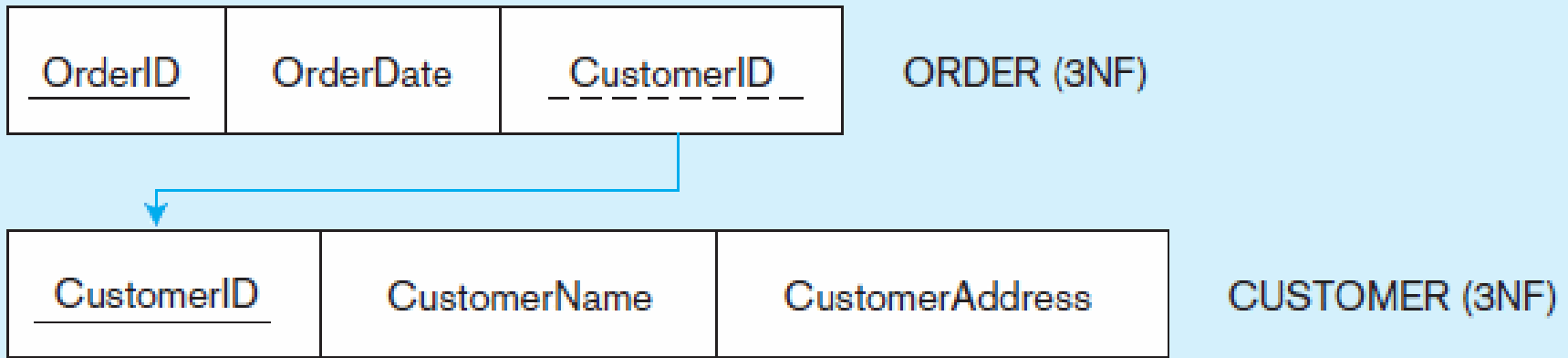
Παρατηρούμε ότι στον πίνακα CUSTOMERORDER υπάρχει μεταβατική εξάρτηση. Ο πίνακας δεν είναι στην τρίτη κανονική μορφή.



Κανονικοποίηση – Παράδειγμα (1)

Κανονικοποίηση

Αφαίρεση μεταβατικών εξαρτήσεων (partial Dependencies) ώστε να πάμε στην τρίτη κανονική μορφή (3NF).



ΤΕΛΙΚΟ ΣΧΗΜΑ

ORDER (OrderID, OrderDate, CustomerID)

CUSTOMER(CustomerID, CustomerName, CustomerAddress)

PRODUCT(ProductID, ProductDescription, ProductFinish, ProductStandardPrice)

ORDERLINE(OrderId, ProductId, OrderedQuantity)



Κανονικοποίηση – Παράδειγμα (2)

Κανονικοποίηση

Μία κατασκευαστική εταιρεία αναλαμβάνει διάφορα έργα. Κάθε έργο έχει έναν μοναδικό αριθμό (project number) και ένα όνομα (name). Σε κάθε έργο απασχολείται έναν αριθμός υπαλλήλων. Κάθε υπάλληλος έχει έναν μοναδικό κωδικό (employee number), όνομα και ειδικότητα (π.χ. μηχανικός, τεχνικός υπολογιστών). Η εταιρεία τιμολογεί τον πελάτη με βάση τις ώρες απασχόλησης των υπαλλήλων σε κάθε έργο. Η ωριαία αποζημίωση εξαρτάται από την θέση του υπαλλήλου.



Κανονικοποίηση – Παράδειγμα (2)

Κανονικοποίηση

TABLE
5.1

A Sample Report Layout

PROJECT NUMBER	PROJECT NAME	EMPLOYEE NUMBER	EMPLOYEE NAME	JOB CLASS	CHARGE/ HOUR	HOURS BILLED	TOTAL CHARGE
15	Evergreen	103	June E. Arbough	Elec. Engineer	\$ 85.50	23.8	\$ 2,034.90
		101	John G. News	Database Designer	\$105.00	19.4	\$ 2,037.00
		105	Alice K. Johnson *	Database Designer	\$105.00	35.7	\$ 3,748.50
		106	William Smithfield	Programmer	\$ 35.75	12.6	\$ 450.45
		102	David H. Senior	Systems Analyst	\$ 96.75	23.8	\$ 2,302.65
				Subtotal			\$10,573.50
18	Amber Wave	114	Annelise Jones	Applications Designer	\$ 48.10	25.6	\$ 1,183.26
		118	James J. Frommer	General Support	\$ 18.36	45.3	\$ 831.71
		104	Anne K. Ramoras *	Systems Analyst	\$ 96.75	32.4	\$ 3,134.70
		112	Darlene M. Smithson	DSS Analyst	\$ 45.95	45.0	\$ 2,067.75
				Subtotal			\$ 7,265.52
22	Rolling Tide	105	Alice K. Johnson	Database Designer	\$105.00	65.7	\$ 6,998.50
		104	Anne K. Ramoras	Systems Analyst	\$ 96.75	48.4	\$ 4,682.70
		113	Delbert K. Joenbrood	Applications Designer	\$ 48.10	23.6	\$ 1,135.16
		111	Geoff B. Wabash	Clerical Support	\$ 26.87	22.0	\$ 591.14
		106	William Smithfield	Programmer	\$ 35.75	12.8	\$ 457.60
				Subtotal			\$13,765.10
25	Starflight	107	Maria D. Alonzo	Programmer	\$ 35.75	25.6	\$ 915.20
		115	Travis B. Bawangi	Systems Analyst	\$ 96.75	45.8	\$ 4,431.15
		101	John G. News *	Database Designer	\$105.00	56.3	\$ 5,911.50
		114	Annelise Jones	Applications Designer	\$ 48.10	33.1	\$ 1,592.11
		108	Ralph B. Washington	Systems Analyst	\$ 96.75	23.6	\$ 2,283.30
		118	James J. Frommer	General Support	\$ 18.36	30.5	\$ 559.98
		112	Darlene M. Smithson	DSS Analyst	\$ 45.95	41.4	\$ 1,902.33
				Subtotal			\$17,595.57
				Total			\$49,199.69

Note: * indicates project leader



Μετατροπή σε 1NF – Παράδειγμα (2)

Κανονικοποίηση

1. Εξάλειψη επαναλαμβανόμενων γνωρισμάτων
2. Προσδιορισμός Πρωτεύοντος Κλειδιού
3. Προσδιορισμός εξαρτήσεων

FIGURE 5.2 A table in first normal form

Table name: DATA_ORG_1NF			Database name: Ch05_ConstructCo			
PROJ_NUM	PROJ_NAME	EMP_NUM	EMP_NAME	JOB_CLASS	CHG_HOUR	HOURS
15	Evergreen	103	June E. Arbough	Elect. Engineer	64.50	23.6
15	Evergreen	101	John G. News	Database Designer	105.00	19.4
15	Evergreen	105	Alice K. Johnson *	Database Designer	105.00	35.7
15	Evergreen	106	William Smithfield	Programmer	35.75	12.6
15	Evergreen	102	David H. Senior	Systems Analyst	96.75	23.8
18	Amber Wave	114	Annelise Jones	Applications Designer	48.10	24.6
18	Amber Wave	116	James J. Frommer	General Support	18.36	45.3
18	Amber Wave	104	Anne K. Ramoras *	Systems Analyst	96.75	32.4
18	Amber Wave	112	Darlene M. Smithson	DSS Analyst	45.95	44.0
22	Rolling Tide	105	Alice K. Johnson	Database Designer	105.00	64.7
22	Rolling Tide	104	Anne K. Ramoras	Systems Analyst	96.75	48.4
22	Rolling Tide	113	Delbert K. Joerbrood *	Applications Designer	48.10	23.6
22	Rolling Tide	111	Geoff B. Wabash	Clerical Support	26.87	22.0
22	Rolling Tide	106	William Smithfield	Programmer	35.75	12.6
25	Starflight	107	Maria D. Alonzo	Programmer	35.75	24.6
25	Starflight	115	Travis B. Bawangi	Systems Analyst	96.75	45.8
25	Starflight	101	John G. News *	Database Designer	105.00	56.3
25	Starflight	114	Annelise Jones	Applications Designer	48.10	33.1
25	Starflight	108	Ralph B. Washington	Systems Analyst	96.75	23.6
25	Starflight	116	James J. Frommer	General Support	18.36	30.5
25	Starflight	112	Darlene M. Smithson	DSS Analyst	45.95	41.4

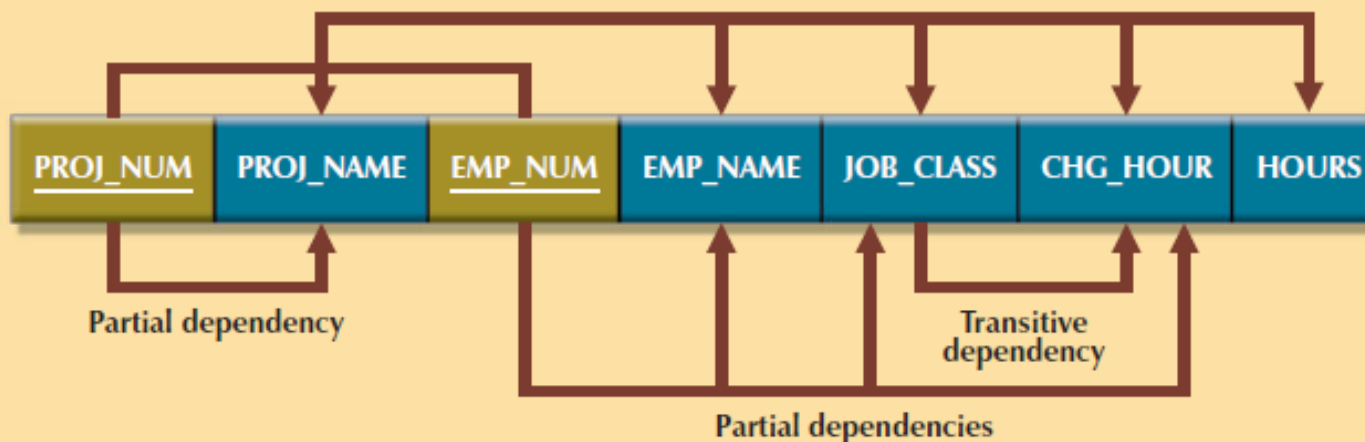


Μετατροπή σε 1NF – Παράδειγμα (2)

Κανονικοποίηση

FIGURE 5.3

First normal form (1NF) dependency diagram



1NF (PROJ_NUM, EMP_NUM, PROJ_NAME, EMP_NAME, JOB_CLASS, CHG_HOURS, HOURS)

PARTIAL DEPENDENCIES:

(PROJ_NUM \Rightarrow PROJ_NAME)

(EMP_NUM \Rightarrow EMP_NAME, JOB_CLASS, CHG_HOUR)

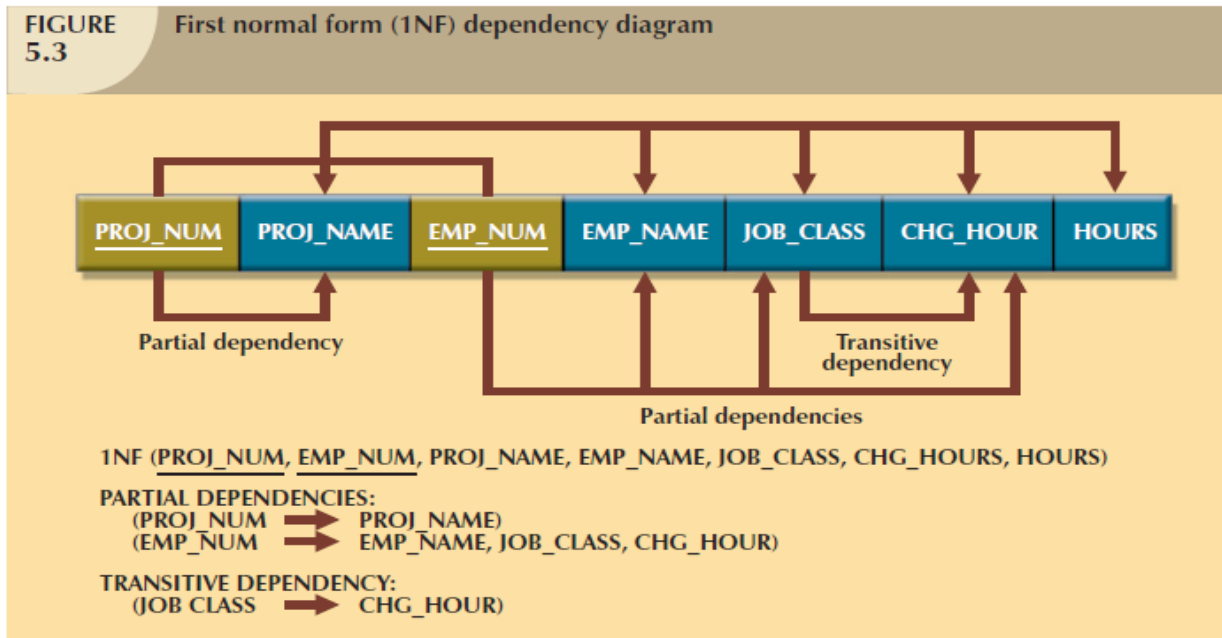
TRANSITIVE DEPENDENCY:

(JOB CLASS \Rightarrow CHG_HOUR)



Μετατροπή σε 2NF – Παράδειγμα (2)

Κανονικοποίηση



Πηγή: Database Systems: Design, Implementation, and Management, Eighth Edition by Peter Rob and Carlos Coronel

PROJ_NUM **PROJECT (PROJ_NUM, PROJ_NAME)**

EMP_NUM **EMPLOYEE (EMP_NUM, JOB_CLASS, CHG_HOUR)**

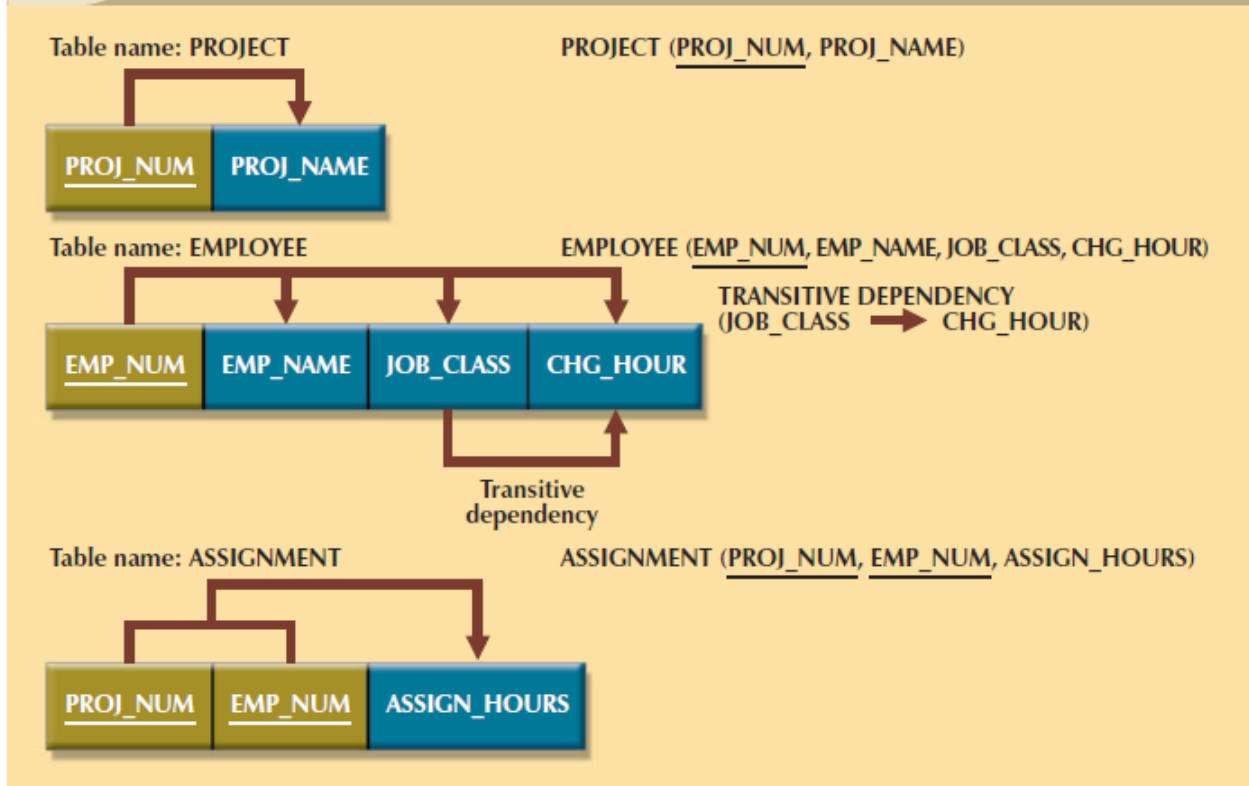
PROJ_NUM EMP_NUM **ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)**



Μετατροπή σε 2NF – Παράδειγμα (2)

Κανονικοποίηση

FIGURE 5.4 Second normal form (2NF) conversion results

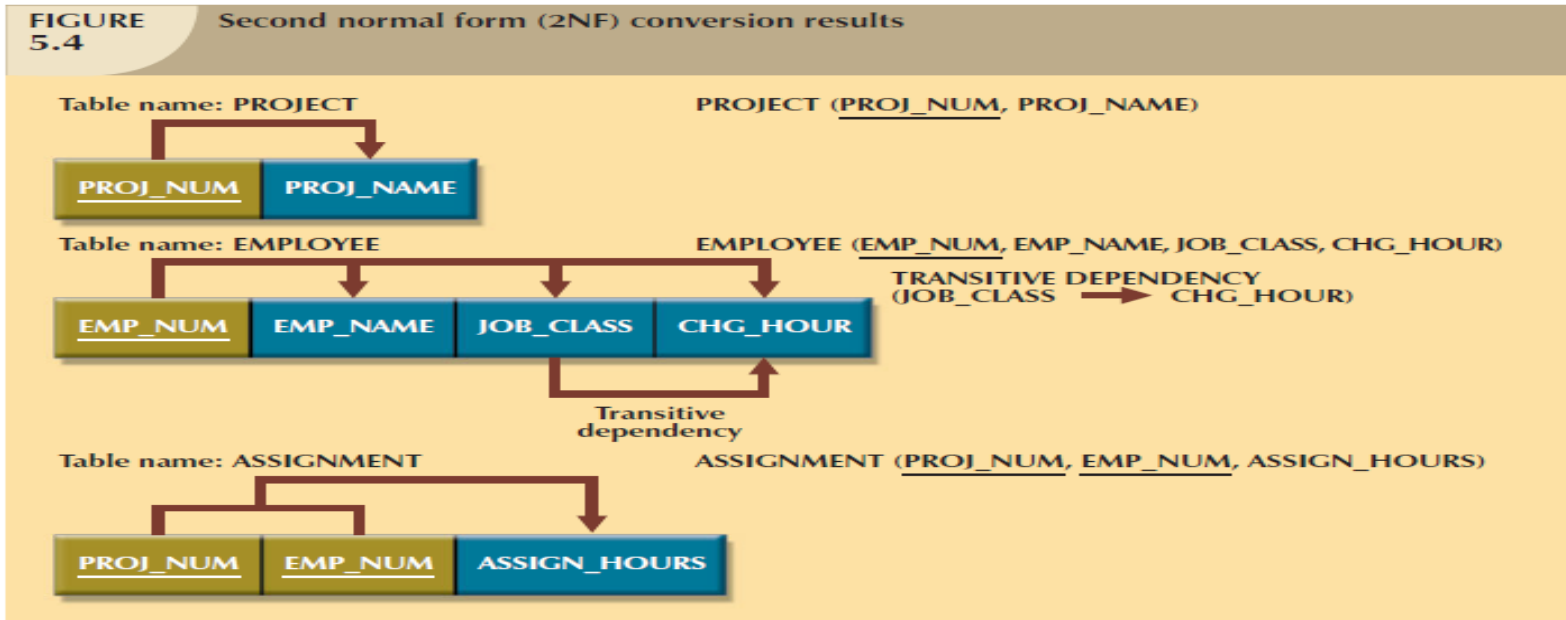


Πηγή: Database Systems: Design, Implementation, and Management, Eighth Edition by Peter Rob and Carlos Coronel



Μετατροπή σε 3NF – Παράδειγμα (2)

Κανονικοποίηση



Πηγή: Database Systems: Design, Implementation, and Management, Eighth Edition by Peter Rob and Carlos Coronel

Εντοπισμός Μεταβατικών Εξαρτήσεων

EMPLOYEE (EMP_NUM, JOB_CLASS)

JOB_CLASS → CHG_HOUR

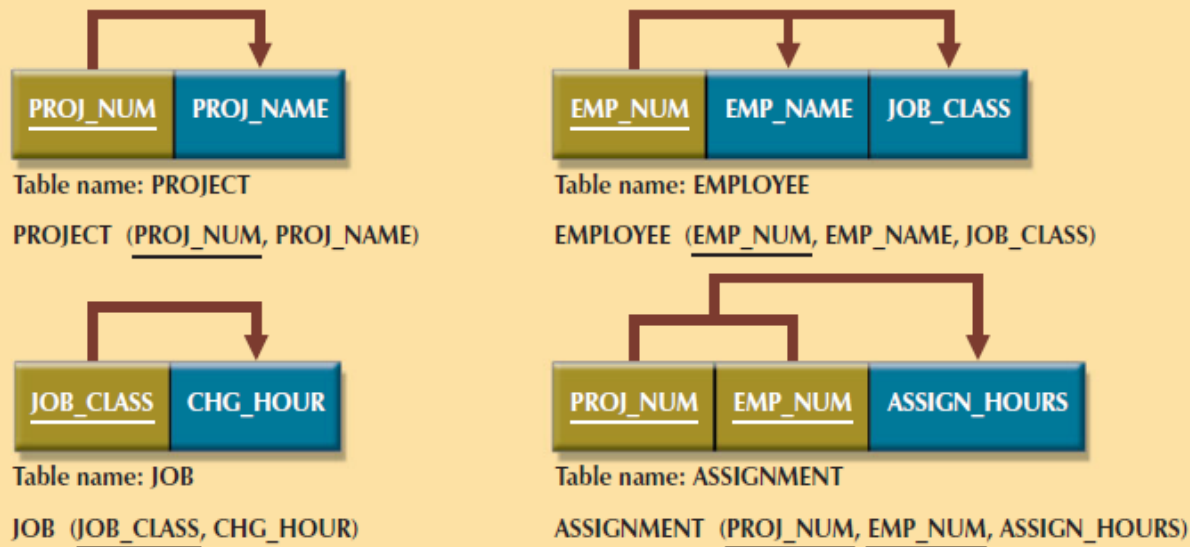
JOB (JOB_CLASS, CHG_HOUR)



Μετατροπή σε 3NF – Παράδειγμα (2)

Κανονικοποίηση

FIGURE 5.5 Third normal form (3NF) conversion results



In other words, after the 3NF conversion has been completed, your database contains four tables:

PROJECT (PROJ_NUM, PROJ_NAME)

EMPLOYEE (EMP_NUM, EMP_NAME, JOB_CLASS)

JOB (JOB_CLASS, CHG_HOUR)

ASSIGNMENT (PROJ_NUM, EMP_NUM, ASSIGN_HOURS)