



## **EPL342 –Databases**

# Lecture 21: Functional Dependencies and Normalization

## Informal Design Guidelines

(Chapter 10.1, Elmasri-Navathe 5ED)

**Διδάσκων: Παναγιώτης Ανδρέου**

<http://www.cs.ucy.ac.cy/courses/EPL342>



# Περιεχόμενο Διάλεξης

- Εισαγωγή στην **Κανονικοποίηση (Normalization)**
- **Άτυπες κατευθύνσεις** για Σχεδιασμό μιας Καλής” Βάσης
  - 1) **Σημασιολογία (Semantics)** των Γνωρισμάτων Σχέσεων
  - 2) Πλεονάζουσες (Redundant) Πληροφορίες σε Πλειάδες και **Ανωμαλίες Ενημερώσεων**
  - 3) **Τιμές NULL** σε πλειάδες
  - 4) Δημιουργία **Πλασματικών Πλειάδων (Spurious Tuples)**

## Επόμενες Διαλέξεις

- **Συναρτησιακές Εξαρτήσεις (Functional Dependencies):**  
Το μαθηματικό υπόβαθρο της Κανονικοποίησης.
- Κανονικοποίηση με Χρήση Συναρτησιακών Εξαρτήσεων:  
1NF, 2NF, 3NF και BCNF, κτλ.

# Εισαγωγή στην Κανονικοποίηση (Normalization)



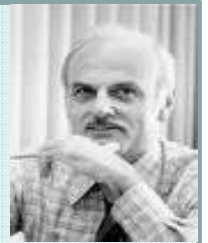
- Η Σχεδίαση του Σχεσιακού Σχήματος μιας βάσης δεδομένων μέχρι στιγμής γινόταν βάσει της **κοινής λογικής** του **σχεδιαστή** και όχι βάσει **αυστηρών κανόνων**.
  - Τα παραγόμενα σχήματα μπορούσαν να κριθούν **υποκειμενικά** ως προς την **Χρηστότητα (Goodness)** τους αλλά όχι **αντικειμενικά**.
- Παρόλο που αναφερθήκαμε στο παρελθόν σε κάποιες γενικές οδηγίες χαρακτηρισμού της **Χρηστότητας (Goodness)** ενός Σχεσιακού Σχήματος, π.χ.,
  - Να **ΜΗΝ** υπάρχει **επανάληψη πληροφορίας**
  - Να **ΜΗΝ** υπάρχουν **πολλά NULLs** σε γνωρίσματα.
  - Να **ΜΗΝ** απαιτούνται **πολλές συνενώσεις** για να καταλήξουμε σε μια οντότητα.

*Αυτές οι οδηγίες δεν ήταν τυπικές (δηλ, δεν στηρίζονταν σε κάποιο μαθηματικό υπόβαθρο οπότεν ήταν υποκειμενικές)!<sup>21-3</sup>*

# Εισαγωγή στην Κανονικοποίηση (Normalization)



- **Κανονικοποίηση (Normalization)** : Η συστηματική (τυπική) διαδικασία διασφάλισης ότι το σχήμα μιας βάσης δεδομένων έχει τις πιο κάτω πρόνοιες:
  - Είναι κατάλληλο για εκτέλεση γενικού είδους επερωτήσεων
  - Είναι ελεύθερο από ορισμένα μη-επιθυμητά χαρακτηριστικά (ανωμαλίες) εισαγωγής, ενημέρωσης, και διαγραφής τα οποία μπορεί να οδηγήσουν σε απώλεια δεδομένων



- Η Κανονικοποίηση ορίζεται το 1970 από τον **Edgar Frank "Ted" Codd**, τον δημιουργό του Σχεσιακού Μοντέλου.
  - Αυτό για να επιτευχθεί ευκολότερα η μετάβαση των τότε βάσεων δεδομένων (ιεραρχικό και δικτυακό μοντέλο) στο **σχεσιακό μοντέλο**.
- Ο Codd ορίζει μέχρι το 1974 διάφορα επίπεδα Κανονικοποίησης (**1NF, 2NF, 3NF και BCNF**)
  - ...τα οποία μπορούν να ακολουθηθούν βήμα-βήμα για προσδιορισμό ενός καλού σχεσιακού σχήματος.

# Εισαγωγή στην Κανονικοποίηση (Normalization)



- Η **Κανονικοποίηση** μιας Βάσης μπορεί να γίνει με τους ακόλουθους δυο τρόπους:

- **Από Πάνω Προς τα Κάτω (Top-Down Design)**

- Αρχικοποίηση μιας (1) σχέσης που περιλαμβάνει **ΌΛΑ** τα γνωρίσματα  **$R(A_1, \dots, A_N)$**  μιας εφαρμογής.
- Η σχέση **διασπάται αναδρομικά** από πάνω προς τα κάτω μέχρι να **φτάσουμε σε σχέσεις** οι οποίες πληρούν κάποιο επιθυμητό επίπεδο «**Χρηστότητας**».
- Ονομάζεται αλλιώς ***Design by Analysis***.

- **Από Κάτω Προς τα Πάνω (Bottom Up Design)**

- Ονομάζεται αλλιώς ***Design-by-Synthesis***
- *Δεν θα μελετηθεί στα πλαίσια αυτού του μαθήματος*

# Εισαγωγή στην Κανονικοποίηση (Normalization)



- Η Κανονικοποίηση στηρίζεται στην έννοια των **συναρτησιακών εξαρτήσεων (functional dependencies)**, τα οποία αποτελούν γενίκευση της έννοιας του κλειδιού.
  - Π.χ.,  $SSN \rightarrow \{Name, NumberGrade, LetterGrade, Dno\}$   
 $NumberGrade \rightarrow LetterGrade$
  - Αυτά θα μελετηθούν εκτενώς στην επόμενη διάλεξη.
- **Προτού δούμε** τις πιο πάνω έννοιες θα επιχειρήσουμε να **συνοψίσουμε με άτυπο τρόπο** κάποιες **γενικές οδηγίες** για το **σχεδιασμό** ενός καλού σχήματος μιας σχεσιακής βάσης.

# Οδηγία 1: Σημασιολογία (Semantics) των Γνωρισμάτων Σχέσεων



- **Οδηγία 1:** Τα απλά γνωρίσματα μιας σχέσης πρέπει να έχουν **κοινή σημασιολογία** (δηλαδή να αναφέρονται στην ίδια οντότητα)
- **Συγκεκριμένα:**
  - Απλά Γνωρίσματα από διαφορετικές οντότητες (π.χ., EMPLOYEES, DEPARTMENTS, PROJECTS) **ΔΕΝ** πρέπει να πλέκονται στην ίδια σχέση
  - ΜΟΝΟ ξένα κλειδιά πρέπει να χρησιμοποιούνται για την αναφορά μεταξύ οντοτήτων.
  - Γνωρίσματα Οντοτήτων και Γνωρίσματα Συσχετίσεων είναι καλό να διατηρούνται μακριά μεταξύ τους (δες επόμενη διαφάνεια)
    - Αυτά συχνά είναι πηγή για NULL τιμές.
- **Σημείωση:** Μετά την εκτενή απασχόληση σας με το Σχεσιακό (αλλα και το ER) μοντέλο, πολλά από τα πιο πάνω μπορεί να σας ακούγονται φυσιολογικά και να τα εφαρμόζεται ήδη.

# Οδηγία 1: Σημασιολογία (Semantics) των Γνωρισμάτων Σχέσεων



- **Παράδειγμα Ανάμειξης Γνωρισμάτων**

Πιο κάτω δείχνουμε δυο παραδείγματα σχέσεων που **αναμειγνύουν γνωρίσματα** μεταξύ οντοτήτων οδηγώντας μας σε ένα σχεσιακό σχήμα με φτωχό design

## Ανάμειξη Employee – Department

EMP\_DEPT

Ename	<u>Ssn</u>	Bdate	Address	Dnumber	Dname	Dmgr_ssn
-------	------------	-------	---------	---------	-------	----------

## Ανάμειξη Employee – Project

EMP\_PROJ

<u>Ssn</u>	<u>Pnumber</u>	Hours	Ename	Pname	Plocation
------------	----------------	-------	-------	-------	-----------

- *Σημείωση: Τα πιο πάνω έχουν φτωχό design ως φυσικές σχέσεις (base relations) όπως θα δούμε σε λίγο. Από την άλλη δεν έχουν ιδιαίτερο σχεδιαστικό πρόβλημα ως όψεις (νοητές σχέσεις)*



# Οδηγία 1: Σημασιολογία (Semantics) των Γνωρισμάτων Σχέσεων



## Συμπέρασμα:

- Θα πρέπει κανείς να μπορεί να περιγράψει μια-μια τις σχέσεις του σχήματος μιας βάσης χωρίς να προκύπτουν **σημασιολογικές αμφιβολίες (semantic ambiguities)**.
  - Π.χ., **Project.Dnum** εύκολα μας παραπέμπει στο number του department
  - Ενώ εάν ήταν απλά **Project.Num** τότε δεν θα ήταν ξεκάθαρη η σημασιολογία του εν λόγω γνωρίσματος.
- Αυτό επιτυγχάνεται με την χρήση **ξεκάθαρων ονομάτων στις σχέσεις και τα γνωρίσματα**.

EMPLOYEE				F.K.
Ename	<u>Ssn</u>	Bdate	Address	Dnumber

P.K.

DEPARTMENT		F.K.
Dname	<u>Dnumber</u>	Dmgr_ssn

P.K.

DEPT_LOCATIONS		F.K.
<u>Dnumber</u>	<u>Dlocation</u>	

P.K.

PROJECT			F.K.
Pname	<u>Pnumber</u>	Plocation	Dnum

P.K.

WORKS_ON			F.K.	F.K.
<u>Ssn</u>	<u>Pnumber</u>	Hours		

P.K.

# Πλεονάζουσα Πληροφορία (Προβλήματα)



- **Πλεονάζουσα Πληροφορία σε Σχέσεις:**  
Δεδομένα τα οποία αποθηκεύονται **επαναληπτικά (redundantly)** προκαλούν τα ακόλουθα προβλήματα:
  - A) Σπαταλούν Χώρο Αποθήκευσης**, εάν και αυτό το πρόβλημα είναι μικρό σε σχέση με το επόμενο
  - B) Προκαλούν προβλήματα με τις ενημερώσεις (update anomalies)**
    - Ανωμαλίες Εισαγωγής (Insertion anomalies)
    - Ανωμαλίες Διαγραφής (Deletion anomalies)
    - Ανωμαλίες Τροποποίησης (Modification anomalies)

*Ας δούμε κάποια παραδείγματα...*

# Πλεονάζουσα Πληροφορία (Σπατάλη Χώρου)



## A) Σπατάλη Χώρου Αποθήκευσης (Redundancy) σε δυο σχέσεις Emp\_Dept και Emp\_Proj

- Και οι δυο περιπτώσεις είναι αποτέλεσμα άκομψου σχεδιασμού του σχήματος των πινάκων

EMP_DEPT					Redundancy	
Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

EMP_PROJ					
Ssn	Pnumber	Hours	Ename	Pname	Plocation
123456789	1	32.5	Smith, John B.	ProductX	Bellaire
123456789	2	7.5	Smith, John B.	ProductY	Sugarland
666884444	3	40.0	Narayan, Ramesh K.	ProductZ	Houston
453453453	1	20.0	English, Joyce A.	ProductX	Bellaire
453453453	2	20.0	English, Joyce A.	ProductY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProductY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProductZ	Houston
333445555	10	10.0	Wong, Franklin T.	Computerization	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganization	Houston
999887777	30	30.0	Zelaya, Alicia J.	Newbenefits	Stafford
999887777	10	10.0	Zelaya, Alicia J.	Computerization	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Computerization	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Newbenefits	Stafford
987654321	30	20.0	Wallace, Jennifer S.	Newbenefits	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganization	Houston
888665555	20	Null	Borg, James E.	Reorganization	Houston

# Ανωμαλίες Ενημερώσεων (Insert Anomaly)



## • Ανωμαλία Εισαγωγής (Insert Anomaly)

– **Πρόβλημα 1:** Για εισαγωγή νέας πλειάδας πρέπει να βεβαιωθούμε ότι το (Dname, Dmgr\_ssn) είναι συνεπές με τις προηγούμενες εγγραφές.

• Π.χ., Dname="Research", Dmgr\_ssn=987654321 **λάθος. Γιατί;**

– **Πρόβλημα 2:** Δεν μπορούμε να εισάγουμε τα στοιχεία ενός Department το οποίο δεν έχει ακόμη employees!

• Δεν γίνεται ούτε εάν βάλουμε **NULL** στα ανύπαρκτα πεδία. **Γιατί;**

EMP\_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

# Ανωμαλίες Ενημερώσεων (Insert Anomaly)



- Το πιο κάτω στιγμιότυπο των Employee-Department δεν υποφέρει από τις ανωμαλίες εισαγωγής της προηγούμενης σελίδας

EMPLOYEE				
Ename	Ssn	Bdate	Address	Dnumber
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4
Narayan, Ramesh K.	666884444	1962-09-15	975 Fire Oak, Humble, TX	5
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1

DEPARTMENT		
Dname	Dnumber	Dmgr_ssn
Research	5	333445555
Administration	4	987654321
Headquarters	1	888665555

A) Δεν χρειάζεται έλεγχος συνέπειας για εισαγωγή στο Department ή Employee

B) Μπορούμε να εισάγουμε Departments, Employees ανεξάρτητα το ένα από την ύπαρξη του άλλου.

# Ανωμαλίες Ενημερώσεων (Delete Anomaly)



- **Ανωμαλία Διαγραφής (Delete Anomaly)**

- **Πρόβλημα:** Εάν σβήσουμε την τελευταία πλειάδα που περιέχει τα στοιχεία ενός Department, τότε θα χαθούν όλες οι πληροφορίες του εν λόγω Department.

- Π.χ., DELETE FROM EMP\_DEPT WHERE SSN=888665555

- Το πρόβλημα αυτό **ΔΕΝ** θα προκύψει στο **στιγμιότυπο της προηγούμενης διαφάνειας**

- όπου τα στοιχεία ενός department συνεχίζουν να διατηρούνται στον πίνακα Department

EMP\_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

Redundancy

# Ανωμαλίες Ενημερώσεων (Update Anomaly)



- **Ανωμαλία Ενημέρωσης (Update Anomaly)**

- **Πρόβλημα:** Εάν ενημερώσουμε το γνώρισμα ενός στοιχείου το οποίο επαναλαμβάνεται τότε πρέπει να ενημερώσουμε **ΟΛΕΣ** τις εμφανίσεις του.
- Αυτό το πρόβλημα είναι το μικρότερο από τα αναφερόμενα διότι λύνεται με μια έκφραση SQL .
  - Π.χ., UPDATE EMP\_DEPT SET Dname='RES' WHERE Dname='Research'
  - ...εάν και το query θα απαιτεί περισσότερο χρόνο από ότι πρέπει

Redundancy

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

# Οδηγία 2: Πλεονάζουσα Πληροφορία και Ανωμαλίες Ενημερώσεων



- **Οδηγία 2:**

- Σχεδίασε ένα σχήμα το οποίο **ΔΕΝ** θα υποφέρει από **ανωμαλίες εισαγωγών, διαγραφών και ενημερώσεων**.
- Εάν είναι **ανάγκη** να **υφίστανται** κάποιες **ανωμαλίες (π.χ., για λόγους επίδοσης\*)**, τότε αυτές πρέπει να **αναφερθούν** και **τεκμηριωθούν ρητά** για να λαμβάνονται υπόψη από την **εφαρμογή**.

- \* π.χ., σε εφαρμογές OLAP: Online Analytical Processing όπου ο φόρτος είναι κυρίως Readonly (χωρίς πολλές ενημερώσεις)



# Οδηγία 2: Πλεονάζουσα Πληροφορία και Ανωμαλίες Ενημερώσεων



- **Από-κανονικοποίηση (Denormalization)**

- Όπως αναφέραμε πριν, σε κάποιες περιπτώσεις ένας DBA μπορεί να κρατήσει το σχήμα μιας βάσης σε “κακή” δομή (μη-κανονικοποιημένη μορφή).

- Π.χ., Εάν έχουμε ένα **συνεχόμενο φόρτο επερωτήσεων** στον πίνακα EMP\_DEPT τότε ο μη-κανονικοποιημένος πίνακας αποφεύγει τη συνεχή συνένωση EMP με DEPT.

- Σημειώστε ότι οι απλές Όψεις δεν βοηθάνε εφόσον πρόκειται απλά για αποθηκευμένα SELECT ερωτήματα.

- Σε αυτές τις περιπτώσεις χρειάζεται η βάση να υλοποιεί, με χρήση **σκανδαλών (triggers), μηχανισμών** για αποφυγή των **διάφορων ανωμαλιών ενημέρωσης**.

- π.χ., κατά την ενημέρωση μιας πλειάδας αυτόματα ενημερώνονται όλες οι άλλες σχετικές πλειάδες

# Οδηγία 3: Τιμες NULL σε Πλειάδες



## • Οδηγία 3:

– Οι σχέσεις πρέπει να σχεδιάζονται έτσι ώστε να έχουν όσο το δυνατό **λιγότερες εμφανίσεις NULL**.

- Π.χ., **GOOD**: Employee(SSN, Dno) → Είναι αναμενόμενο ότι οι περισσότεροι Employees θα **ανήκουν** σε ένα Department οπότε δικαιολογείται η συμμετοχή του εν λόγω γνωρίσματος στην σχέση Employee.

– Γνωρίσματα που εμφανίζουν συχνά NULL πρέπει να **τοποθετούνται σε ξεχωριστές σχέσεις** (μαζί με το πρωτεύων κλειδί της σχέσης)

- Π.χ., *Μόνο το 10% των Employees έχουν γραφείο.*

**BAD**: EMPLOYEE(SSN, Name, **Officedetails**,...)

**GOOD**: EMPLOYEE(SSN, Name,....)

EMP\_OFFICE(SSN, Officedetails)

# Οδηγία 4: Πλασματικές Πλειάδες (Spurious Tuples)



- Κατά την φάση της top-down κανονικοποίησης, πολλές σχέσεις καλούνται αναδρομικά να διασπαστούν σε **μικρότερες σχέσεις**.
- **Διασπώντας (Decomposing)** λανθασμένα μια σχέση  $R$  κρύβει τους ακόλουθους δυο κινδύνους:
  - **A) Επανασύνθεση:** Είναι δυνατό να **ΜΗΝ** μπορούμε να επανασυνθέσουμε την  $R$  χάνοντας πολύτιμη σημασιολογία της εφαρμογής
    - Π.χ., εάν το κλειδί της  $R$  δεν περιέχεται τόσο στην  $A$  όσο και στην  $B$ .
  - **B) Πλασματικές Πλειάδες:** Είναι δυνατό να δημιουργούνται **Πλασματικές Πλειάδες (Spurious Tuples)**, δηλαδή πλειάδες που δεν άνηκαν στην  $R$ .
- *Ακολουθούν παραδείγματα για το  $B$ ....*

# Οδηγία 4: Πλασματικές Πλειάδες (Spurious Tuples)



EMP\_PROJ1

Ssn	Pnumber	Hours	Pname	Plocation
123456789	1	32.5	ProductX	Bellaire
123456789	2	7.5	ProductY	Sugarland

EMP\_LOCS

Ename	Plocation
Smith, John B.	Bellaire
Smith, John B.	Surgarland
Narayan, Ramesh K.	Houston
English, Joyce A.	Bellaire
English, Joyce A.	Surgarland
Wong, Franklin T.	Surgarland
Wong, Franklin T.	Houston
Wong, Franklin T.	Stafford
Zelaya, Alicia J.	Stafford
Jabbar, Ahmad V.	Stafford
Wallace, Jennifer S.	Stafford
Wallace, Jennifer S.	Houston
Borg, James E.	Houston



## Κανονικό Αποτέλεσμα (Χωρίς Πλασματικές Πλειάδες)

EMP\_PROJ

Ssn	Pnumber	Hours	Pname	Plocation	Ename
123456789	1	32.5	ProductX	Bellaire	Smith, John B.
123456789	2	7.5	ProductY	Sugarland	Smith, John B.

## Λάθος Αποτέλεσμα Με Πλασματικές Πλειάδες

Ssn	Pnumber	Hours	Pname	Plocation	Ename
123456789	1	32.5	ProductX	Bellaire	Smith, John B.
123456789	1	32.5	ProductX	Bellaire	English, Joyce A.
123456789	2	7.5	ProductY	Sugarland	Smith, John B.
123456789	2	7.5	ProductY	Sugarland	English, Joyce A.
123456789	2	7.5	ProductY	Sugarland	Wong, Franklin T.
⋮					

Το λάθος αποτέλεσμα είναι λόγω της κακής δομής του EMP\_LOCS που δεν έχει το SSN

# Οδηγία 4 Πλασματικές Πλειάδες (Spurious Tuples)



- **Οδηγία 4:**

- Οι σχέσεις πρέπει να σχεδιάζονται έτσι ώστε να μπορούν να επανσυνθετούν, τυπικά λέγεται ότι θα πρέπει να ισχύει το **lossless join condition**,  $^* (\pi_{R_1}(r), \dots, \pi_{R_m}(r)) = r$
- **ΔΕΝ** πρέπει να παράγονται πλασματικές πλειάδες κατά την φυσική συνένωση μεταξύ **ΟΠΟΙΩΝΔΗΠΟΤΕ** σχέσεων με κοινά γνωρίσματα συνένωσης.

- Υπάρχουν 2 σημαντικές **ιδιότητες** της **διάσπασης (decomposition)** μιας σχέσης:
  - a) Διαφύλαξη της Αρχικής Σχέσης: **Lossless Join Decomposition** → Δεν μπορεί να παραβλεφθεί.
  - b) Διαφύλαξη των Συναρτησιακών Εξαρτήσεων: **Functional dependency preservation** → Μπορεί κάποτε να παραβλεφθεί.
- **Περισσότερα σε ερχόμενες διαλέξεις**