

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΤΜΗΥΠ

ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ Ι

Β. Μεγαλοικονόμου
Δ. Χριστοδουλάκης

Σχεδιασμός Βάσεων Δεδομένων και Κανονικοποίηση

(παρουσίαση βασισμένη εν μέρη σε σημειώσεις των Silberchatz, Korth και Sudarshan και του C. Faloutsos)



Σύνοψη Ύλης

- Σχεσιακό μοντέλο
 - Τυπικές γλώσσες ερωτημάτων
 - Εμπορικές γλώσσες ερωτημάτων (SQL)
- Περιορισμοί ακεραιότητας
 - Πεδίο Περιορισμού ακεραιότητας, ξένα κλειδιά
 - Συναρτησιακές εξαρτήσεις
- Συναρτησιακές Εξαρτήσεις
- **Σχεδιασμός ΒΔ και κανονικοποίηση**



Σύνοψη Ύλης- αναλυτικά

- Σχεδιασμός ΒΔ και κανονικοποίηση
 - Παγίδες κακού σχεδιασμού
 - Αποσύνθεση (**decomposition**)
 - Κανονικές μορφές (**normal forms**)

ΣΤΟΧΟΣ

- Σχεδιασμός “καλών” πινάκων
 - Ειδικός στόχος 1: να οριστεί τι σημαίνει *καλός*
 - Ειδικός στόχος 2: να διορθωθούν οι “κακοί” πίνακες
- Συνοπτικά:

“θέλουμε πίνακες που τα γνωρίσματά τους να εξαρτώνται από το πρωτεύον κλειδί, *ολόκληρο* το κλειδί, και *τίποτα άλλο* εκτός από το κλειδί”
- Ας δούμε γιατί, και πώς:



Παγίδες

Παίρνει1 (AM, κωδ, βαθμός, όνομα, διεύθυνση)

AM	Κωδ	Βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας

Παγίδες

κακός πίνακας – Γιατί;

γιατί: AM \rightarrow διεύθυνση, όνομα

AM	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίο	Ήωνίας
123	cs211	A	Φωτίο	Ήωνίας



Παγίδες

- Πλεονασμός (επανάληψη πληροφορίας)
 - Χωρητικότητα
 - (ασυνέπειες)
 - Ανωμαλίες κατά την εισαγωγή/διαγραφή:

Παγίδες

- Ανωμαλίες κατά την εισαγωγή:
 - “Αλεξίου” γράφεται στο τμήμα αλλά δεν παίρνει κανένα μάθημα – δεν υπάρχει μέρος για την αποθήκευση της διεύθυνσής του!

ΑΜ	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331A		Φωτίου	Ιωνίας
...
234	null	null	Αλεξίου	Θράκης

Παγίδες

- Ανωμαλίες διαγραφής:
 - Με τη διαγραφή του τελευταίου record για τον 'Φωτίου' (χάνουμε τη διεύθυνσή του!)

AM	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
123	cs211	A	Φωτίου	Ιωνίας

Λύση: αποσύνθεση

Διάσπαση του πίνακα που προκαλεί το πρόβλημα

AM	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
123	cs211	A	Φωτίου	Ιωνίας

?

?



Επισκόπηση - Αναλυτικά

- Σχεδιασμός ΒΔ και κανονικοποίηση
 - Παγίδες κακού σχεδιασμού
 - Αποσύνθεση
 - Συνένωση άνευ απωλειών (lossless join)
 - Διατήρηση εξαρτήσεων (dependency preservation)
 - Κανονικές Μορφές



Αποσύνθεση

- Υπάρχουν και 'κακές' αποσυνθέσεις
- Θέλουμε να:
 - Είναι άνευ απωλειών και να
 - Διατηρούν τις εξαρτήσεις

Αποσύνθεση – με Απώλειες

R1(AM, βαθμός, όνομα, διεύθυνση) R2(Κωδ, βαθμός)

AM	βαθμός	όνομα	διεύθυνση
123	A	Φωτίου	Ιωνίας
123	B	Φωτίου	Ιωνίας
234	A	Αλεξίου	Θράκης

Κωδ	βαθμός
cs331	A
cs351	B
cs211	A

AM	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
234	cs211	A	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

AM, Κωδ → βαθμός

Αποσύνθεση – με Απώλειες

Δεν μπορούμε να πάρουμε το αρχικό στιγμιότυπο με μία πράξη συνένωσης (join)

ΑΜ	βαθμός	όνομα	διεύθυνση
123	A	Φωτίου	Ιωνίας
123	B	Φωτίου	Ιωνίας
234	A	Αλεξίου	Θράκης

Κωδ	βαθμός
cs331	A
cs351	B
cs211	A

ΑΜ	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
234	cs211	A	Αλεξίου	Θράκης

ΑΜ → όνομα, διεύθυνση

ΑΜ, Κωδ → βαθμός

Αποσύνθεση – με Απώλειες

Άλλο παράδειγμα

Αποσύνθεση της $R = (A, B)$ σε $R_1 = (A)$, $R_2 = (B)$

A	B
α	1
α	2
β	1

r

A
α
β

$\Pi_A(r)$

B
1
2

$\Pi_B(r)$

$\Pi_A(r) \bowtie \Pi_B(r)$

A	B
α	1
α	2
β	1
β	2

Αποσυνθέσεις

Αποσύνθεση που δε διατηρεί τις εξαρτήσεις

Κωδ	Πόλη	ήπειρος
123	Λονδίνο	E
125	Παρίσι	E
234	Πεκίνο	A

Κωδ → πόλη, ήπειρος

πόλη → ήπειρος

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

κωδ → πόλη

κωδ → ήπειρος

Κωδ	ήπειρος
123	E
125	E
234	A

Αποσυνθέσεις

Είναι άνευ απωλειών;

Κωδ	Πόλη	ήπειρος
123	Λονδίνο	E
125	Παρίσι	E
234	Πεκίνο	A

Κωδ → πόλη, ήπειρος

πόλη → ήπειρος

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

κωδ → πόλη

κωδ → ήπειρος

Κωδ	ήπειρος
123	E
125	E
234	A

Αποσυνθέσεις – άνευ απωλειών

Ορισμός:

Θεωρείστε το σχήμα R , με Σ.Ε. 'F'.

Η διάσπαση R_1, R_2 είναι διάσπαση άνευ απωλειών συνένωσης της R αν έχουμε **πάντα:**

$$r_1 \triangleright \triangleleft r_2 = r$$

Υπάρχει πιο εύκολο κριτήριο;

Αποσυνθέσεις – άνευ απωλειών

Θεώρημα:

Αποσύνθεση άνευ απωλειών συνένωσης
αν το γνώρισμα συνένωσης είναι
υπερκλειδί σε έναν τουλάχιστον από
τους νέους πίνακες

Τυπικά:

$$R1 \cap R2 \rightarrow R1 \text{ ή}$$

$$R1 \cap R2 \rightarrow R2$$

Αποσύνθεση – άνευ απωλειών

Παράδειγμα:

R1

AM	Κωδ	βαθμός
123	cs331	A
123	cs351	B
234	cs211	A

AM, Κωδ → βαθμός

R2

AM	όνομα	διεύθυνση
123	Φωτίου	Ιωνίας
234	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

AM	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
234	cs211	A	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

AM, κωδ → βαθμός



Επισκόπηση - αναλυτικά

- Σχεδιασμός ΒΔ και κανονικοποίηση
 - Παγίδες κακού σχεδιασμού
 - Αποσύνθεση
 - Αποσύνθεση άνευ απωλειών συνένωσης
 - **Διατήρηση Εξαρτήσεων**
 - Κανονικές μορφές

Αποσύνθεση – Διατηρ. Εξαρτ.

Άτυπα δεν θέλουμε οι αρχικές Σ.Ε. να εκτείνονται σε 2 πίνακες - αντιπαράδειγμα:

Κωδ	Πόλη	ήπειρος
123	Λονδίνο	E
125	Παρίσι	E
234	Πεκίνο	A

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

Κωδ	ήπειρος
123	E
125	E
234	A

Κωδ → πόλη, ήπειρος

πόλη → ήπειρος

κωδ → πόλη

κωδ → ήπειρος

Αποσύνθεση – Διατηρ. Εξαρτ.

Αποσύνθεση που διατηρεί τις εξαρτήσεις

Κωδ	Πόλη	ήπειρος
123	Λονδίνο	E
125	Παρίσι	E
234	Πεκίνο	A

Κωδ → πόλη, ήπειρος

Πόλη → ήπειρος

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

Πόλη	ήπειρος
Λονδίνο	E
Παρίσι	E
Πεκίνο	A

Κωδ → πόλη

πόλη → ήπειρος

(όμως: Κωδ → ήπειρος ?)

Αποσύνθεση – Διατηρ. Εξαρτ.

Άτυπα:

δεν θέλουμε οι αρχικές Σ.Ε. να εκτείνονται σε ' δύο πίνακες

Συγκεκριμένα: ... οι Σ.Ε. **κανονικού καλύμματος**

Έστω F_j το σύνολο των Σ.Ε. F^+ που περιέχουν μόνο γνωρίσματα στην R_j .

- Ιδανικά η αποσύνθεση θα έπρεπε να διατηρεί τις εξαρτήσεις, δηλ.,
 $(F_1 \cup F_2 \cup \dots \cup F_n)^+ = F^+$
- *Διαφορετικά*, ο έλεγχος των ενημερώσεων για παραβίαση των Σ.Ε. θα απαιτούσε υπολογιστικές συνενώσεις \rightarrow μεγάλο κόστος

Αποσύνθεση – Διατηρ. Εξαρτ.

Γιατί είναι καλό να διατηρήσουμε τις Σ.Ε.;

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

Κωδ	ήπειρος
123	E
125	E
234	A

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

Πόλη	ήπειρος
Λονδίνο	E
Παρίσι	E
Πεκίνο	A

Κωδ → πόλη Κωδ → ήπειρος

Κωδ → πόλη Πόλη → ήπειρος

(πόλη → ήπειρος: 'χάνεται')

Αποσύνθεση – Διατηρ. Εξαρτ.

A: πχ., καταγραφή ότι το Πεκίνο είναι στην ήπειρο 'Α'

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

Κωδ	ήπειρος
123	E
125	E
234	A

Κωδ	Πόλη
123	Λονδίνο
125	Παρίσι
234	Πεκίνο

Πόλη	ήπειρος
Λονδίνο	E
Παρίσι	E
Πεκίνο	A

Κωδ → πόλη Κωδ → ήπειρος

Κωδ → πόλη Πόλη → ήπειρος

(πόλη → ήπειρος: 'χάνεται')

Αποσύνθεση – Διατηρ. Εξαρτ.

Για να ελέγξουμε αν μια εξάρτηση $\alpha \rightarrow \beta$ διατηρείται σε μια αποσύνθεση της σχέσης R σε R_1, R_2, \dots, R_n κάνουμε το ακόλουθο test (εφαρμόζοντας κλειστότητα γνωρίσματος στο F)

αποτέλεσμα = α

while (υπάρχουν αλλαγές στο αποτέλεσμα) do

for each R_i στην αποσύνθεση

$t = (\text{αποτέλεσμα} \cap R_i)^+ \cap R_i$

αποτέλεσμα = *αποτέλεσμα* \cup t

Αν το αποτέλεσμα περιέχει όλα τα γνωρίσματα στο β , τότε η Σ.Ε $\alpha \rightarrow \beta$ διατηρείται

Αποσύνθεση – Διατηρ. Εξαρτ.

Κάνουμε το test σε όλες τις εξαρτήσεις στο F για να ελέγξουμε αν η αποσύνθεση διατηρεί τις εξαρτήσεις

Το test απαιτεί **πολυωνυμικό χρόνο**

Ο υπολογισμός του F^+ και $(F_1 \cup F_2 \cup \dots \cup F_n)^+$ απαιτεί **εκθετικό χρόνο**



Αποσύνθεση - Συμπεράσματα

- Η αποσύνθεση πρέπει να είναι άνευ απωλειών
 - **γνώρισμα συνένωσης → υπερκλειδί**

- Όποτε είναι δυνατό, πρέπει να διατηρούνται οι εξαρτήσεις (σε κάποιες περιπτώσεις αυτό ίσως να είναι ανέφικτο)

Κανονικοποίηση με χρήση Σ.Ε.

- Κατά την αποσύνθεση ενός σχεσιακού σχήματος R με ένα σύνολο Σ.Ε. F σε R_1, R_2, \dots, R_n θέλουμε:
 - Αποσύνθεση άνευ απωλειών συνένωσης: διαφορετικά ... απώλεια πληροφορίας
 - Έλλειψη πλεονασμού: οι σχέσεις R_i ιδανικά θα πρέπει να είναι είτε σε Boyce-Codd Normal Form (BCNF) ή σε 3^η Κανονική Μορφή (3NF)
 - Διατήρηση Εξαρτήσεων: Έστω F_i το σύνολο των εξαρτήσεων στην F^+ που περιέχει μόνο τα γνωρίματα της R_i .
 - Ιδανικά η αποσύνθεση θα πρέπει να διατηρεί τις εξαρτήσεις, πχ. $(F_1 \cup F_2 \cup \dots \cup F_n)^+ = F^+$
 - Διαφορετικά, ο έλεγχος των ενημερώσεων για παραβίαση των Σ.Ε. θα απαιτούσε υπολογιστικές συνενώσεις \rightarrow μεγάλο κόστος

Κανονικοποίηση με χρήση Σ.Ε. - Παράδειγμα

- $R = (A, B, C)$
 $F = \{A \rightarrow B, B \rightarrow C\}$
- $R_1 = (A, B), R_2 = (B, C)$
 - Αποσύνθεση άνευ απωλειών συνένωσης:
 $R_1 \cap R_2 = \{B\}$ και $B \rightarrow BC$
 - Διατήρηση Εξαρτήσεων
- $R_1 = (A, B), R_2 = (A, C)$
 - Αποσύνθεση άνευ απωλειών συνένωσης:
 $R_1 \cap R_2 = \{A\}$ και $A \rightarrow AB$
 - Μη διατήρηση εξαρτήσεων
(δεν μπορεί να ελεγχθεί το $B \rightarrow C$ χωρίς να υπολογιστεί το $R_1 \bowtie R_2$)



Επισκόπηση - αναλυτικά

- Σχεδιασμός ΒΔ και Κανονικοποίηση
 - Παγίδες κακού σχεδιασμού
 - Αποσύνθεση (\rightarrow πώς θα λύσουμε το πρόβλημα)
 - **Κανονικές Μορφές** (\rightarrow πώς θα εντοπίσουμε το πρόβλημα)
 - BCNF,
 - 3NF,
 - (1NF, 2NF)



Κανονικές Μορφές - BCNF

Είδαμε πώς θα διορθώσουμε 'κακά' σχήματα -
'Όμως τι είναι 'καλό' σχήμα;

Απάντηση: 'καλό', αν υπακούει μια κανονική
μορφή, δηλ., ένα σύνολο κανόνων

Τυπικά: Boyce-Codd Normal Form (BCNF)

Κανονικές Μορφές - BCNF

Ορισμός: το σχεσιακό σχήμα R είναι σε BCNF ως προς το F , αν

- άτυπα: τα πάντα εξαρτώνται από το πλήρες κλειδί, και τίποτα άλλο εκτός από αυτό
- Σχεδόν-τυπικά: κάθε προσδιοριστής (determinant) του καλύμματος είναι ένα υποψήφιο κλειδί

Κανονικές Μορφές - BCNF

Παράδειγμα και αντι-παράδειγμα:

AM	Όνομα	διεύθυνση
123	Φωτίου	Ιωνίας
123	Φωτίου	Ιωνίας
234	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

AM	Κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
234	cs211	A	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

AM, κωδ → βαθμός

Κανονικές Μορφές - BCNF

Τυπικά: για κάθε Σ.Ε. $A \rightarrow B$ στο F^+

- $A \rightarrow B$ είναι τετριμμένη (το A είναι ένα υπερσύνολο του B) ή
- Το A είναι υπερκλειδί
- (ή και τα δύο)



Κανονικές Μορφές - BCNF

Θεώρημα: δεδομένου ενός σχήματος R και ενός συνόλου Σ.Ε. F , μπορούμε πάντα να το αποσυνθέσουμε σε σχήματα R_1, \dots, R_n , έτσι ώστε

- R_1, \dots, R_n να είναι σε BCNF και
- Η αποσύνθεση να είναι άνευ απωλειών

(...όμως, κάποιες αποσυνθέσεις μπορεί να απωλέσουν εξαρτήσεις)



BCNF Αποσύνθεση

Πώς?

...ουσιαστικά, διαχωρίζοντας τις Σ.Ε. του καλύμματος

Πχ. ΠΑΙΡΝΕΙ1(ΑΜ, κωδ, βαθμός, όνομα, διεύθυνση)

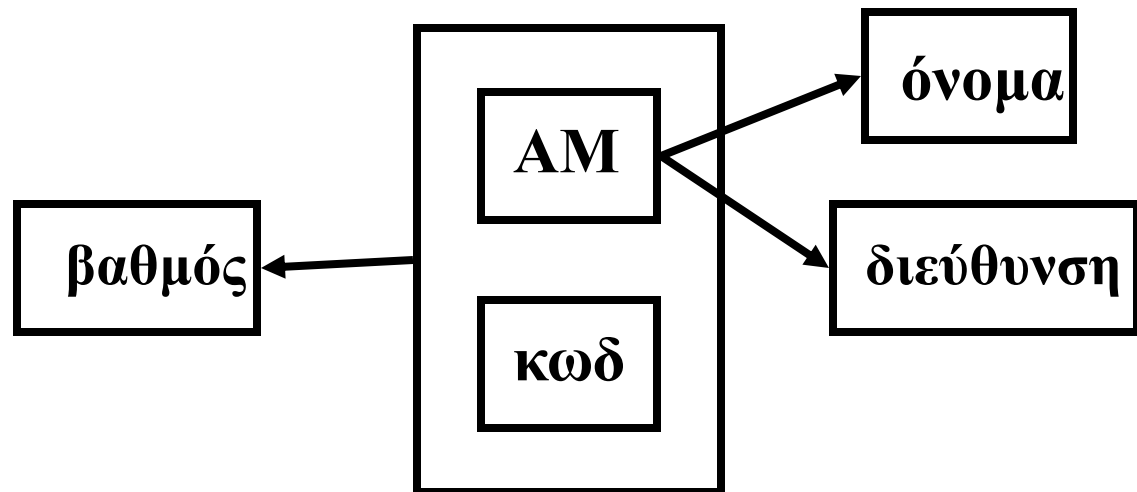
ΑΜ → όνομα, διεύθυνση

ΑΜ, κωδ → βαθμός

Κανονικές Μορφές - BCNF

Πχ. ΠΑΙΡΝΕΙ1(ΑΜ, κωδ, βαθμός, όνομα, διεύθυνση)

ΑΜ → όνομα, διεύθυνση ΑΜ, κωδ → βαθμός



Κανονικές Μορφές - BCNF

AM	κωδ	βαθμός
123	cs331	A
123	cs351	B
234	cs211	A

AM, κωδ → βαθμός

AM	όνομα	διεύθυνση
123	Φωτίου	Ιωνίας
123	Φωτίου	Ιωνίας
234	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

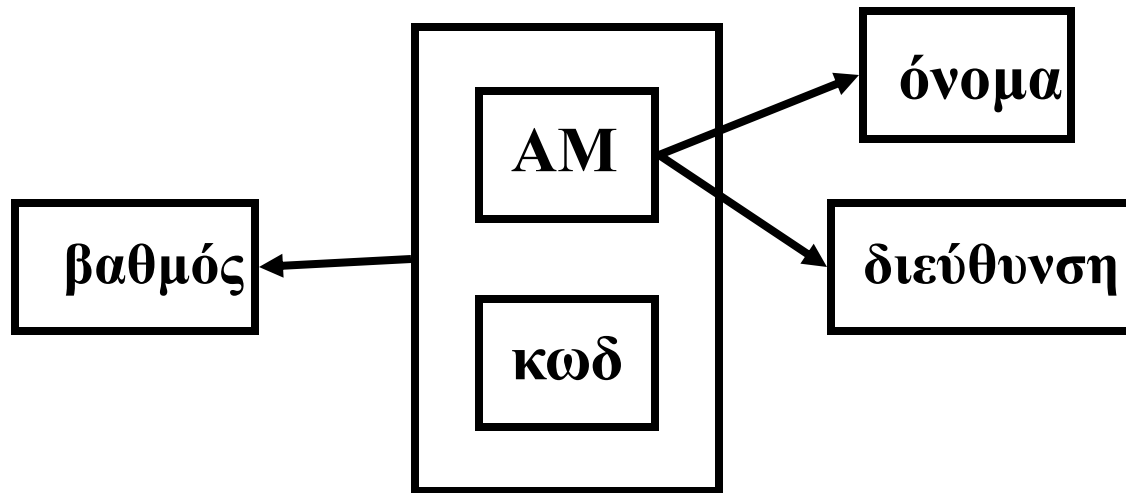
AM	κωδ	βαθμός	όνομα	διεύθυνση
123	cs331	A	Φωτίου	Ιωνίας
123	cs351	B	Φωτίου	Ιωνίας
234	cs211	A	Αλεξίου	Θράκης

AM → όνομα, διεύθυνση

AM, κωδ → βαθμός

Κανονικές Μορφές - BCNF

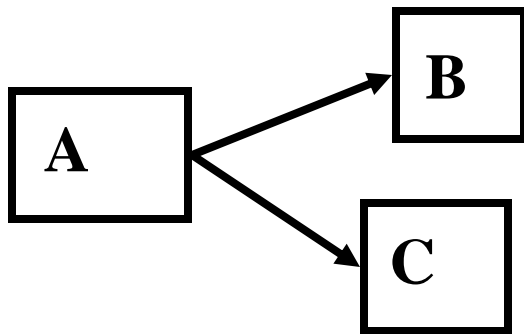
σχηματικά: θέλουμε το σχήμα ενός 'αστεριού'



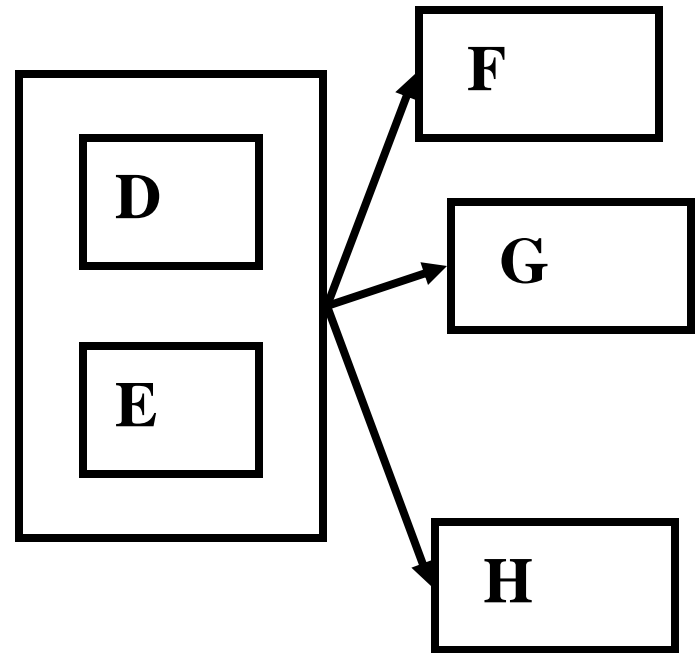
:δεν είναι σε BCNF

Κανονικές Μορφές - BCNF

σχηματικά: θέλουμε το σχήμα ενός 'αστεριού'



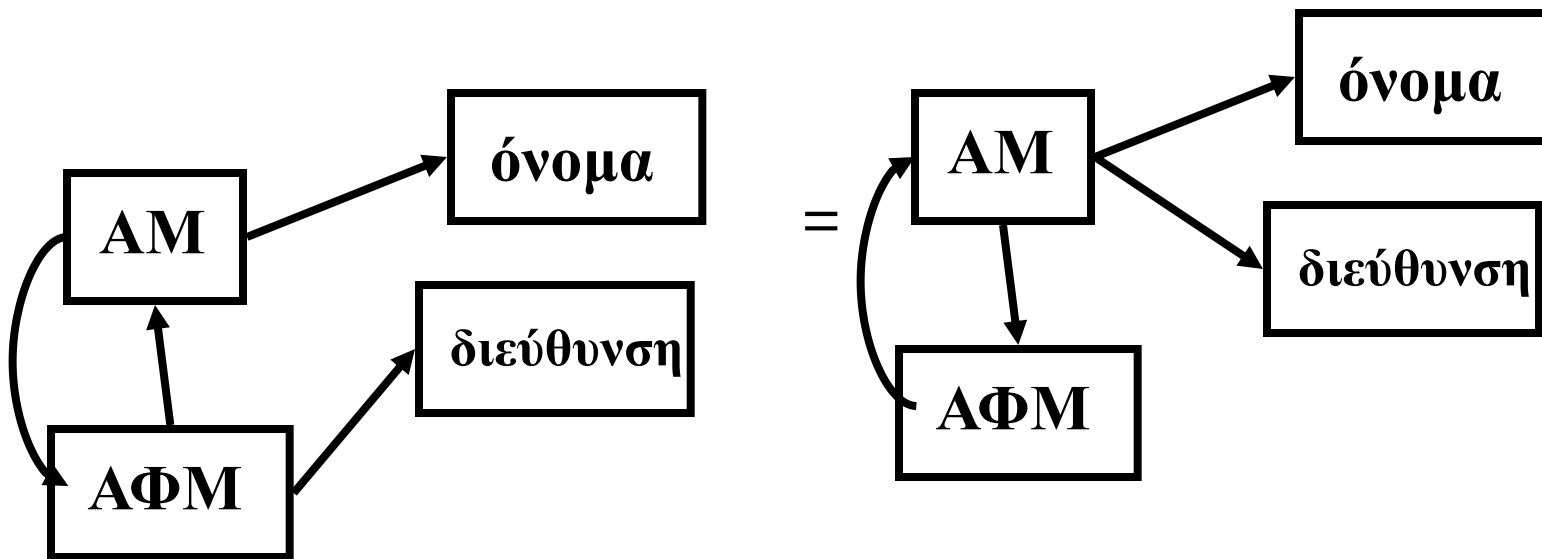
ή



Κανονικές Μορφές - BCNF

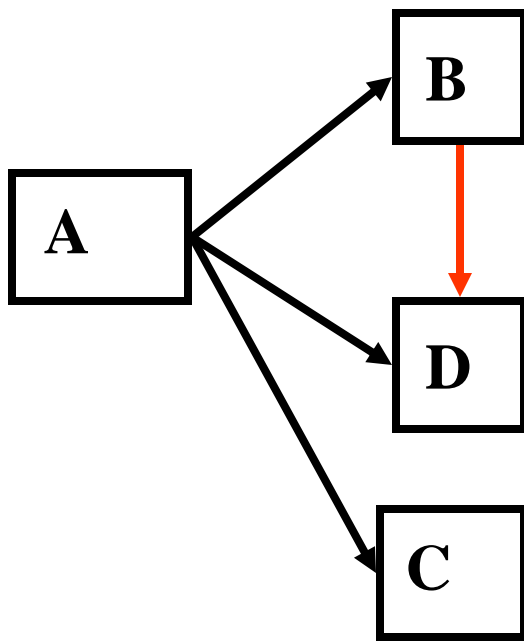
Ή να μοιάζει με αστέρι: (πχ. 2 υποψήφια κλειδιά):

ΦΟΙΤΗΤΗΣ(AM, ΑΦΜ., όνομα, διεύθυνση)

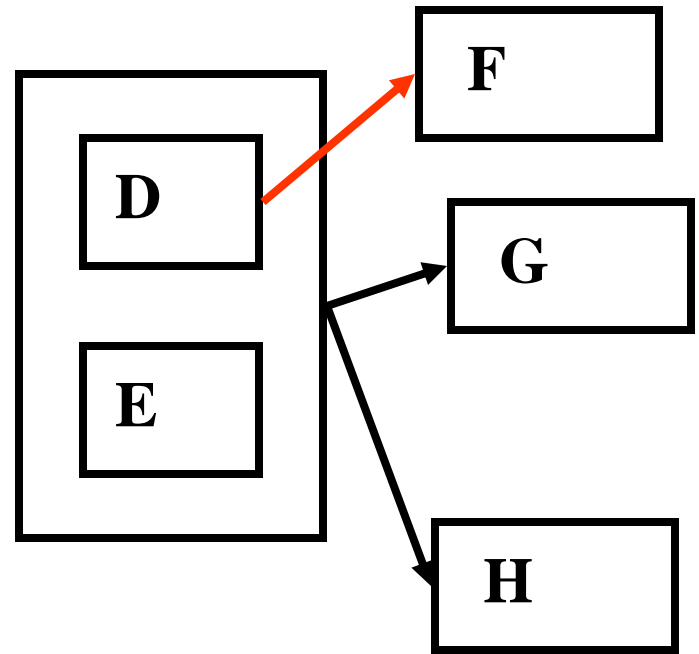


Κανονικές Μορφές - BCNF

Αλλά **όχι**:



ή



BCNF Αποσύνθεση

```
αποτέλεσμα := {R};
done := false;
υπολόγισε  $F^+$ ;
while (not done) do
    if (υπάρχει ένα σχήμα  $R_i$  στο αποτέλεσμα που δεν είναι σε BCNF)
        then begin
            έστω  $\alpha \rightarrow \beta$  μη τετριμμένη συναρτησιακή
            εξάρτηση που ισχύει στην  $R_i$ 
            έτσι ώστε  $\alpha \rightarrow R_i$  δεν είναι στο  $F^+$ ,
            και  $\alpha \cap \beta = \emptyset$ ;
            αποτέλεσμα := (αποτέλεσμα -  $R_i$ )  $\cup$  ( $R_i - \beta$ )  $\cup$  ( $\alpha, \beta$ );
        end
    else done := true;
```

Σημ: κάθε R_i είναι σε BCNF, και η αποσύνθεση είναι
άνευ απωλειών συνένωσης

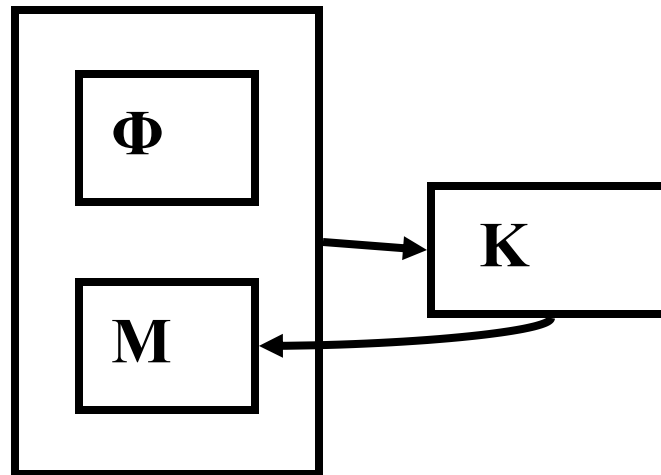
Κανονικές Μορφές - 3NF

Θεωρείστε την 'κλασική' περίπτωση:
ΦΚΜ (Φοιτητής, Καθηγητής, Μάθημα)

$K \rightarrow M$

$\Phi, M \rightarrow K$

Είναι σε BCNF?

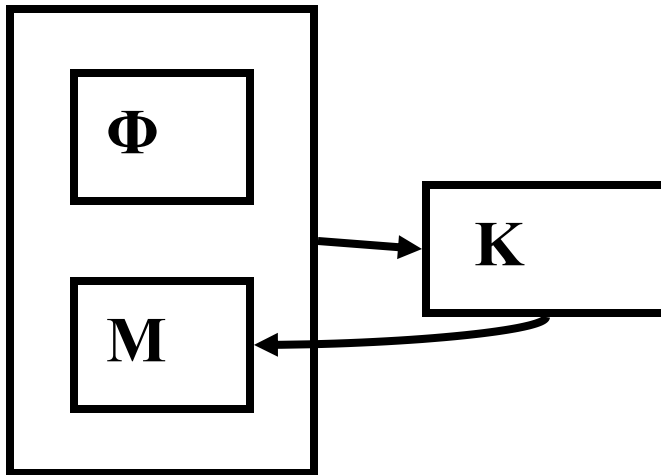


Κανονικές Μορφές - 3NF

ΦΚΜ (Φοιτητής, Καθηγητής, Μάθημα)

$K \rightarrow M$ $\Phi, M \rightarrow K$

Πώς θα το αποσυνθέσουμε σε BCNF?





Κανονικές Μορφές - 3NF

ΦΚΜ (Φοιτητής, Καθηγητής, Μάθημα)

$K \rightarrow M$ $\Phi, M \rightarrow K$

1) $R_1(K, M)$ $R_2(\Phi, M)$

(BCNF; - άνευ απωλειών; - διατ. εξαρτ.;)

2) $R_1(K, M)$ $R_2(\Phi, K)$

(BCNF; - άνευ απωλειών; - διατ. εξαρτ.;)

Κανονικές Μορφές - 3NF

ΦΚΜ (Φοιτητής, Καθηγητής, Μάθημα)

$K \rightarrow M$ $\Phi, M \rightarrow K$

1) R1(K,M) R2(Φ,M)

(BCNF; **N+N** – άνευ απωλειών; **O** – Διατ. Εξαρτ.; **O**)

2) R1(K,M) R2(Φ,K)

(BCNF; **N+N** - άνευ απωλειών; **N** - Διατ. Εξαρτ.; **O**)

Κανονικές Μορφές - 3NF

ΦΚΜ (Φοιτητής, Καθηγητής, Μάθημα)

$K \rightarrow M$ $\Phi, M \rightarrow K$

Σε αυτή την περίπτωση: αδύνατο να έχουμε και τα δύο

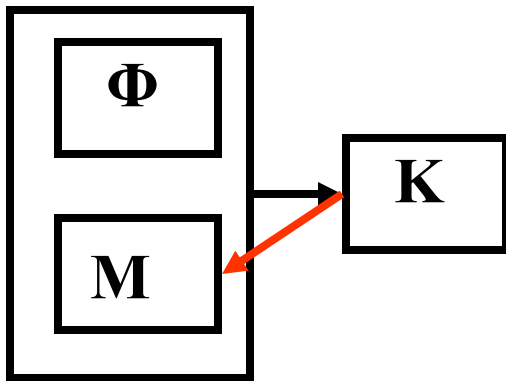
- BCNF **και**
- Διατήρηση εξαρτήσεων

Ας υποδεχτούμε την 3NF (...μια ασθενέστερη κανονική μορφή)!

Κανονικές Μορφές - 3NF

ΦΚΜ (Φοιτητής, Καθηγητής, Μάθημα)

$K \rightarrow M$ $\Phi, M \rightarrow K$

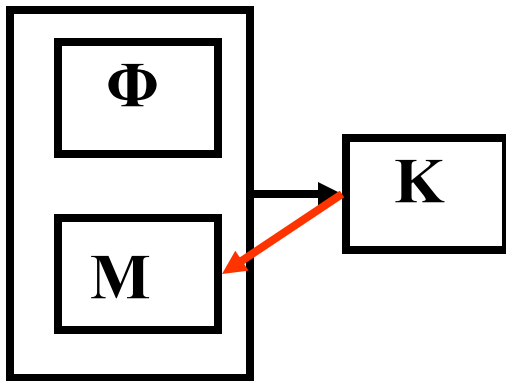


άτυπα, η 3NF

‘επιτρέπει’ το κόκκινο βέλος στο κανονικό κάλυμμα

Κανονικές Μορφές - 3NF

ΦΚΜ (Φοιτητής,
Καθηγητής, Μάθημα)
 $K \rightarrow M \quad \Phi, M \rightarrow K$



Τυπικά, ένα σχεσιακό σχήμα R με Σ.Ε. 'F' είναι σε 3NF αν:

Για κάθε $a \rightarrow b$ στο F^+ :

- Είναι τετριμμένη ή
- Το a είναι υπερκλειδί ή
- Κάθε γνώρισμα $b-a$: είναι μέρος υποψήφιου κλειδιού

Κανονικές Μορφές - 3NF

- $R = (J, K, L)$
 $F = \{JK \rightarrow L, L \rightarrow K\}$
δύο υποψήφια κλειδιά = JK και JL
- Το R δεν είναι σε BCNF
- Οποιαδήποτε αποσύνθεση του R θα αποτύγχανε να διατηρήσει

$$JK \rightarrow L$$

- Η BCNF αποσύνθεση έχει (JL) και (LK)
 - Ο έλεγχος για το $JK \rightarrow L$ απαιτεί συνένωση
 - Το R είναι σε 3NF
 - $JK \rightarrow L$ το JK είναι υπερκλειδί
 - $L \rightarrow K$ το K **περιέχεται** σε ένα υποψήφιο κλειδί
- Υπάρχει κάποιος πλεονασμός σε αυτό το σχήμα...

Κανονικές Μορφές - 3NF

ΕΛΕΓΧΟΣ ΓΙΑ 3NF

- Βελτιστοποίηση: Χρειάζεται να ελέγξουμε μόνο τις Σ.Ε. στο F , δε χρειάζεται να ελέγξουμε όλες τις Σ.Ε. στο F^+
- Χρησιμοποιούμε την κλειστότητα γνωρίσματος για να ελέγξουμε, για κάθε εξάρτηση $\alpha \rightarrow \beta$, αν το α είναι υπερκλειδί
- Αν το α δεν είναι υπερκλειδί, πρέπει να επιβεβαιώσουμε αν κάθε γνώρισμα στο β περιέχεται σε ένα υποψήφιο κλειδί για το R
 - Ο έλεγχος αυτός είναι πιο ακριβός καθώς χρειάζεται να βρούμε τα υποψήφια κλειδιά
 - Ο έλεγχος για 3NF είναι NP-hard
 - Το ενδιαφέρον είναι πως η αποσύνθεση σε 3NF (περιγράφηκε συνοπτικά) μπορεί να συντελεστεί σε πολυωνυμικό χρόνο

Αποσύνθεση σε 3NF

έστω F_c το κανονικό κάλυμμα για το F ;

$i := 0$;

for each συναρτησιακή εξάρτηση $\alpha \rightarrow \beta$ στο F_c **do**

if κανένα από τα σχήματα R_j , $1 \leq j \leq i$ δεν περιέχει τα α β

then begin

$i := i + 1$;

$R_i := \alpha \beta$

end

if κανένα από τα σχήματα R_j , $1 \leq j \leq i$ δεν περιέχει ένα υποψήφιο κλειδί για το R

then begin

$i := i + 1$;

$R_i :=$ οποιοδήποτε υποψήφιο κλειδί για το R ;

end

return (R_1, R_2, \dots, R_i)

Οι εξαρτήσεις διατηρούνται με τη δημιουργία ακριβώς ενός σχήματος για κάθε δεδομένη εξάρτηση

Εγγυάται μια αποσύνθεση άνευ απωλειών συνένωσης έχοντας τουλάχιστον ένα σχήμα που περιέχει ένα υποψήφιο κλειδί για το σχήμα που διασπάται



Κανονικές Μορφές - 3NF

Πώς θα μετατρέψουμε ένα σχήμα σε 3NF;

Εν συντομία

....για κάθε Σ.Ε. στο κάλυμμα, τοποθέτησέ την σε έναν πίνακα

Κανονικές Μορφές - 3NF vs BCNF

- Αν το 'R' είναι σε BCNF, είναι πάντα σε 3NF (αλλά όχι το αντίστροφο)
- Πρακτικά, πρέπει να στοχεύουμε σε
 - BCNF, συνένωση άνευ απωλειών και διατήρηση εξαρτήσεων
- Αν αυτό είναι αδύνατο, μπορούμε να δεχτούμε
 - 3NF, επιμένοντας όμως σε συνένωση άνευ απωλειών και διατήρηση εξαρτήσεων
 - Η 3NF παρουσιάζει προβλήματα με τις μεταβατικές εξαρτήσεις

3NF vs BCNF (συνέχεια)

- Παράδειγμα προβλημάτων εξαιτίας του πλεονασμού της 3NF

- $R = (J, K, L)$
 $F = \{JK \rightarrow L, L \rightarrow K\}$

J	L	K
j_1	l_1	k_1
j_2	l_1	k_1
j_3	l_1	k_1
$null$	l_2	k_2

Ένα σχήμα που είναι σε 3NF αλλά όχι σε BCNF έχει τα εξής προβλήματα:

- Επανάληψη πληροφορίας (πχ. η συσχέτιση l_1, k_1)
- Απαιτεί τη χρήση null τιμών (πχ. για την αναπαράσταση της συσχέτισης l_2, k_2 όπου δεν υπάρχει αντίστοιχη τιμή για το J).

Κανονικές Μορφές - περισσότερα

- Γιατί '3'NF; Τι είναι η 2NF; Η 1NF;
- 1NF: τα γνωρίσματα είναι ατομικά (δηλ. δεν υπάρχουν γνωρίσματα με τιμές συνόλου δηλ. 'επαναλαμβανόμενες ομάδες')
- 1NF: όχι 'σχέσεις μέσα σε σχέσεις' ή 'σχέσεις ως γνωρίσματα πλειάδων'

ΑΦΜ	όνομα	εξαρτώμενοι
123	Ράσης	Πέτρος Μαρία Γιάννης
234	Δήμου	Άννα Μιχάλης

Όχι 1NF

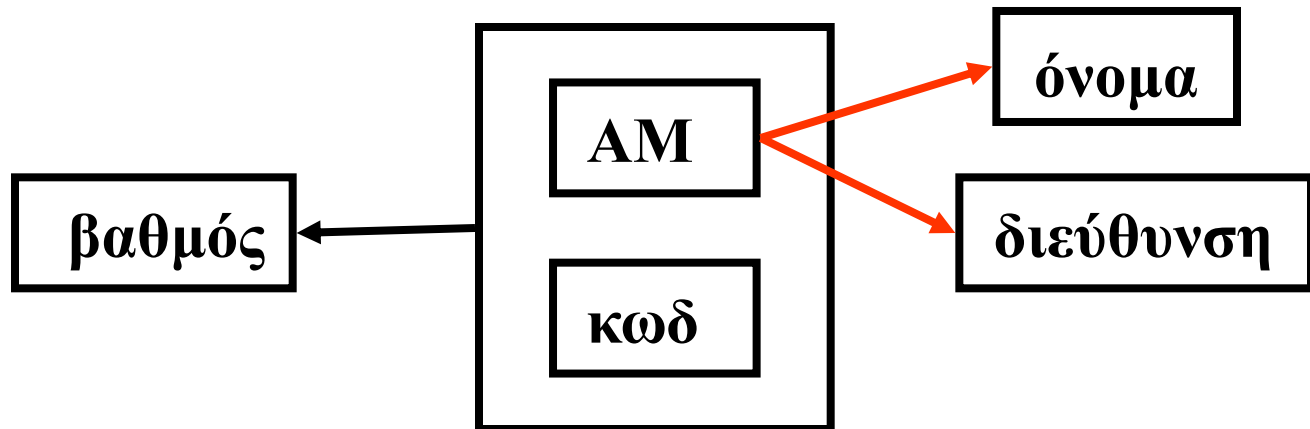
Κανονικές Μορφές - περισσότερα

2NF: 1NF και γνώρισμα που δεν είναι κλειδιά πλήρως εξαρτώμενα από το κλειδί

: αν κάθε μη πρωτεύον γνώρισμα είναι πλήρως συναρτησιακά εξαρτώμενο από το πρωτεύον κλειδί

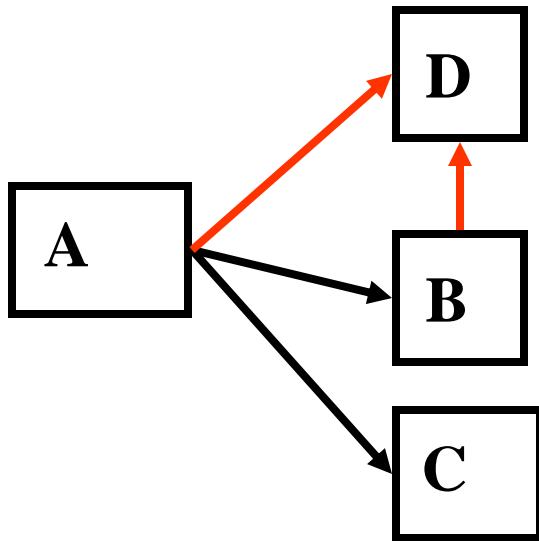
Αντι-παράδειγμα: ΠΑΙΡΝΕΙ1(ΑΜ, κωδ, βαθμός, όνομα, διεύθυνση)

ΑΜ → όνομα, διεύθυνση ΑΜ, κωδ → βαθμός



Κανονικές Μορφές - περισσότερα

- 3NF: 2NF χωρίς μεταβατικές εξαρτήσεις
: 2NF και κανένα μη πρωτεύον γνώρισμα δεν εξαρτάται μεταβατικά από το πρωτεύον κλειδί
- Αντι-παράδειγμα:



σε 2NF, αλλά **όχι** σε 3NF



Κανονικές Μορφές - περισσότερα

- 4NF, πλειότειμες εξαρτήσεις:
αργότερα...
- Πρακτικά, τα διαγράμματα E-R συνήθως οδηγούν σε πίνακες που είναι σε BCNF



Σύνοψη – συμπεράσματα

Σχεδιασμός ΒΔ και Κανονικοποίηση

- Παγίδες κακού σχεδιασμού
- Αποσυνθέσεις (άνευ απωλειών, διατήρηση εξαρτήσεων)
- Κανονικές μορφές (BCNF ή 3NF)

“τα πάντα πρέπει να εξαρτώνται από το κλειδί, **ολόκληρο** το κλειδί, και **τίποτα άλλο** εκτός από το κλειδί”