



Βάσεις Δεδομένων I

Ενότητα 3: Μετατροπή σχήματος Ο/Σ σε σχεσιακό

Δρ. Τιμπίρης Αλκιβιάδης

Τμήμα Μηχανικών Πληροφορικής ΤΕ



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο ΤΕΙ Κεντρικής Μακεδονίας» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ενότητα 3

Σχεσιακό Μοντέλο

Δρ. Τιμπίρης Αλκιβιάδης

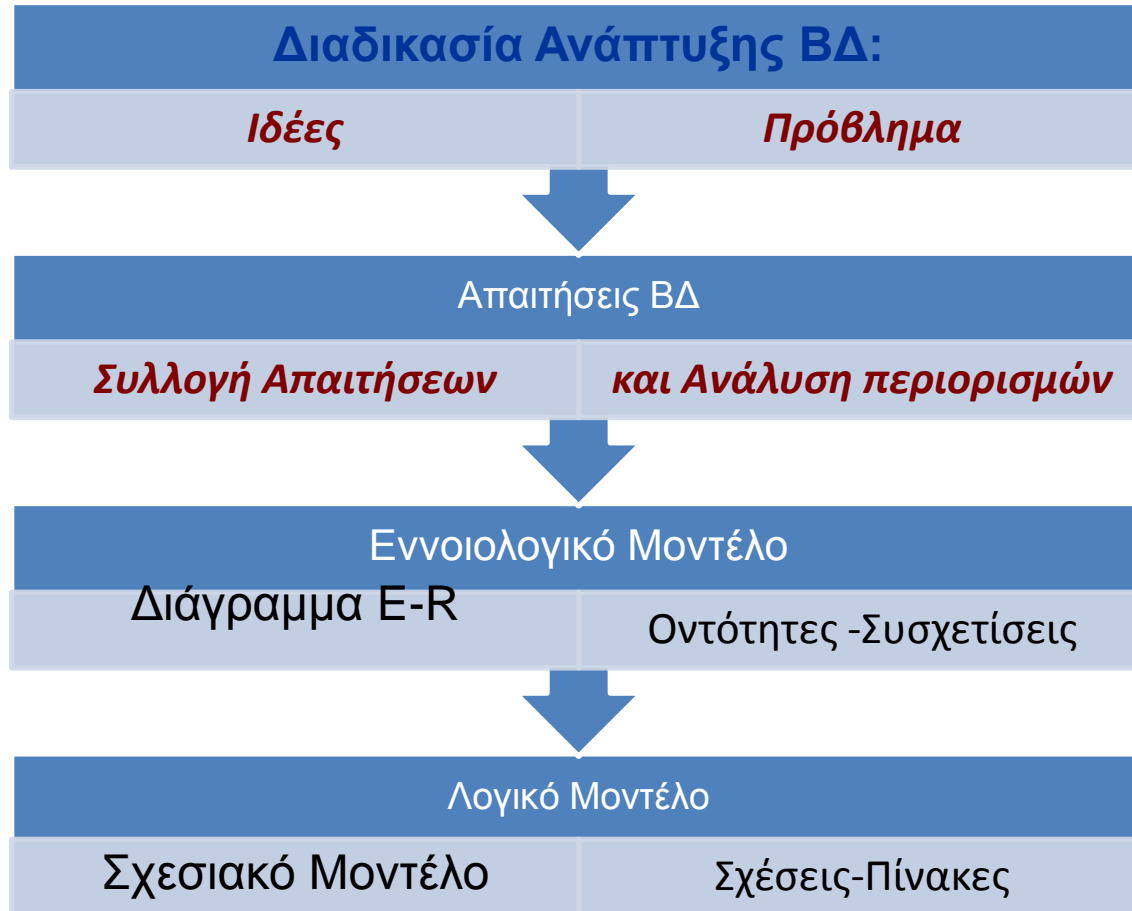
Περιεχόμενα ενότητας

- Λογικός Σχεδιασμός (Σχεσιακό Μοντέλο)
- Εισαγωγικά
- Τυπικός ορισμός Σχεσιακού Μοντέλου
- Κριτήρια Σχεδιασμού

Σκοποί ενότητας

Ο σκοπός της άσκησης αυτής είναι η παρουσίαση και η κατανόηση του σχεσιακού μοντέλου, το μοντέλο που έχει επικρατήσει για το σχεδιασμό των βάσεων δεδομένων. Παρουσιάζονται τα δομικά στοιχεία, οι δομικοί περιορισμοί του μοντέλου αυτού

Διαδικασία Ανάπτυξης ΒΔ



Δομές - Πράξεις

- ΔΟΜΕΣ: **RELATIONS** (ΣΧΕΣΕΙΣ) – μία μόνο Δομή
- ΠΡΑΞΕΙΣ: ***Επεξεργασία των Σχέσεων***
- Αντιστοιχίες με το E-R Μοντέλο

Τύπος Οντοτήτων	<->	Τύπος Σχέσης (Relation)
Οντότητα	<->	n-πλειάδα (Tuple)
Γνώρισμα	<->	Γνώρισμα (Attribute)
Συσχέτιση	<->	Τύπος Σχέσης (Relation)

Πίνακας=Σχέση?

Ένα Σχεσιακό Σχήμα είναι ένα Σύνολο Σχέσεων ή Πινάκων

Τα ονόματα των στηλών είναι τα ονόματα των Γνωρισμάτων

ΑΑΑ	345	78	ΦΦΦΦΦ
123	ΚΚΚΚΚ	ΒΒΒΒ	ΤΤΤΤΤ
1-12-1990	1234	ΥΕΣ	2-3-2000

Είναι αυτός ο πίνακας σχέση?

Στιγμιότυπο ΒΔ

ΚΩΔΙΚΟΣ_M	ΜΑΘΗΜΑ	ΕΞΑΜΗΝΟ	ΔΙΔΑΚΤΙΚΕΣ	ΚΩΔΙΚΟΣ_Κ	Κ
3	ΦΥΣΙΚΗ		3	6 ΠΑΠΑΔΟΠΟΥΛ	
4	ΜΑΘΗΜΑΤΙΚΑ		5	5 ΠΑΠΑΣ	
5	ΑΓΓΛΙΚΑ		1	3 ΔΗΜΗΤΡΙΟΥ	
6	ΒΑΣΕΙΣ_ΔΕΔΟΜΕΝΩΝ		4	7 ΔΗΜΗΤΡΙΟΥ	
7	ΣΤΑΤΙΣΤΙΚΗ		5	7 ΠΑΠΑΔΟΠΟΥΛ	
8	ΒΙΟΜΗΧΑΝΙΚΗ		5	4 ΔΗΜΗΤΡΙΟΥ	

ΔΕΜ	ΕΞΕΤΑΣΤΙΚΗ	ΚΩΔΙΚΟΣ_ΜΑΘΗΜΑΤΟΣ	ΒΑΘΜΟΣ
ΠΑΠΑΔΟΠΟΥΛ	ΕΑΡ2008-2009	ΦΥΣΙΚΗ	5
ΠΑΠΑΔΟΠΟΥΛΟΥ	ΧΕΙΜ2009-2010	ΜΑΘΗΜΑΤΙΚΑ	5
ΠΑΠΑΔΟΠΟΥΛΟΥ	ΧΕΙΜ2010-2011	ΑΓΓΛΙΚΑ	1
ΠΑΠΑΔΟΠΟΥΛΟΥ	ΧΕΙΜ2010-2011	ΒΑΣΕΙΣ_ΔΕΔΟΜΕΝΩΝ	4
ΠΑΠΑΔΟΠΟΥΛΟΥ	ΕΑΡ2008-2009	ΣΤΑΤΙΣΤΙΚΗ	5
ΣΤΑΜΟΣ	ΧΕΙΜ2010-2011	ΦΥΣΙΚΗ	7
ΣΤΑΜΟΣ	ΧΕΙΜ2009-2010	ΜΑΘΗΜΑΤΙΚΑ	3
ΣΤΑΜΟΣ	ΧΕΙΜ2011-2012	ΑΓΓΛΙΚΑ	8
ΣΤΑΜΟΣ	ΧΕΙΜ2010-2011	ΒΑΣΕΙΣ_ΔΕΔΟΜΕΝΩΝ	6
ΣΤΑΜΟΣ	ΕΑΡ2008-2009	ΣΤΑΤΙΣΤΙΚΗ	10
ΓΕΩΡΓΙΟΥ	ΕΑΡ2008-2009	ΦΥΣΙΚΗ	6
ΓΕΩΡΓΙΟΥ	ΕΑΡ2009-2010	ΜΑΘΗΜΑΤΙΚΑ	7
ΓΕΩΡΓΙΟΥ	ΕΑΡ2008-2010	ΑΓΓΛΙΚΑ	6
ΓΕΩΡΓΙΟΥ	ΕΑΡ2009-2010	ΒΑΣΕΙΣ_ΔΕΔΟΜΕΝΩΝ	5
ΓΕΩΡΓΙΟΥ	ΧΕΙΜ2009-2010	ΣΤΑΤΙΣΤΙΚΗ	3

ΕΠΙΘΕΤΟ	ΟΝΟΜΑ	ΔΙΕΥΘΥΝΣΗ	ΕΙΔΙΚΟΤΗΤΑ	ΠΟΛΗ	ΤΚ	ΚΙΝΗΤΟ	ΣΤΑΘΕΡΟ	E_MAIL	ΦΩΤΟΓΡΑΦ	ΣΧΟΛΙΑ
ΠΑΠΑΣ	ΗΛΙΑΣ	Δ.ΘΡΑΚΗΣ 1	ΜΑΘΗΜΑΤΙΚΟΣ	ΣΕΡΡΕΣ	65789	687266268	232107788	kastanhs@gmail.com	Bitmap Image	
ΠΑΠΑΔΟΠΟΥΛΟΣ	ΠΑΝΝΗΣ	ΗΡΩΔΟΤΟΥ 3	ΦΥΣΙΚΟΣ	ΑΘΗΝΑ	78973	678976278	758373938	kapoukranhs@gmail.co	Bitmap Image	
ΔΗΜΗΤΡΙΟΥ	ΕΥΑΓΓΕΛΙΑ	ΚΑΡΑΪΣΧΑΚΗ 1	ΧΗΜΙΚΟΣ	ΘΕΣΣΑΛΟΝΙΚΗ	98987	698025780	838388389	soyltanh@gmail.com	Bitmap Image	

ΔΕΜ	ΟΝΟΜΑ	ΕΠΙΘΕΤΟ	ΟΝΟΜΑ_ΠΑΤΡΟΣ	ΗΜΕΡΟΜΗΝΙΑ_ΕΓΓΡΑΦΗΣ	ΗΜΕΡΟΜΗΝΙΑ_ΓΕΝΝΗΣΗΣ	ΠΟΛΗ	ΔΙΕΥΘΥΝΣΗ
1	ΜΑΡΙΑ	ΠΑΠΑΔΟΠΟΥΛΟΥ	ΧΡΗΣΤΟΣ	1/9/2009	1/1/1985	ΣΕΡΡΕΣ	ΣΤΑΜΟΥΛΗ 19
2	ΚΩΣΤΑΣ	ΣΤΑΜΟΣ	ΠΑΝΝΗΣ	2/9/2009	10/9/1989	ΘΕΣΣΑΛΟΝΙΚΗ	ΙΟΥΣΤΙΝΙΑΝΟΥ 3
3	ΕΛΕΝΗ	ΓΕΩΡΓΙΟΥ	ΑΛΕΞΙΟΣ	1/2/2009	9/3/1989	ΑΘΗΝΑ	ΡΑΜΕΣΤΟΥ 16
*	(Νέο)						

Βασικές Πράξεις

3 ΒΑΣΙΚΕΣ ΠΡΑΞΕΙΣ

- *ΕΠΙΛΟΓΗ (SELECT)*
- *ΠΡΟΒΟΛΗ (PROJECT)*
- *ΚΑΡΤΕΣΙΑΝΟ ΓΙΝΟΜΕΝΟ (PRODUCT)*

ΜΗ ΒΑΣΙΚΗ ΑΛΛΑ ΧΡΗΣΙΜΗ ΠΡΑΞΗ

- *ΣΥΝΔΕΣΗ (JOIN). Είναι συνδυασμός καρτεσιανού γινομένου και επιλογής*

Ενημερώσεις/τροποποιήσεις

INSERT:	<i>Εισαγωγή γραμμών</i>	-
DELETE:	<i>Διαγραφή γραμμών</i>	
UPDATE:	<i>Τροποποίηση</i>	

Η συντριπτική πλειοψηφία των DBMS είναι Σχεσιακά και διατίθενται σε ΟΛΕΣ τις Υπολογιστικές Πλατφόρμες.

E.F. Codd (“A relational model for large shared data banks”, CACM), σαν Θεωρία για Μοντέλα Δεδομένων

Ορισμός

- Σχεσιακή Βάση Δεδομένων είναι ένα σύνολο από ΣΧΕΣΕΙΣ
- **ΣΧΕΣΗ (RELATION)**: Ένας Πίνακας τιμών.
 - Κάθε **στήλη** στον πίνακα έχει όνομα, ονομάζεται **γνώρισμα, πεδίο (attribute –field-column)**.
 - Κάθε **γραμμή** ονομάζεται **πλειάδα, εγγραφή (Tuple, record)** και παριστά τα χαρακτηριστικά μιας οντότητας στο μοντέλο.

Τυπικός Ορισμός

- **ΔΟΜΕΣ**
 - Μόνο μία δομή: *relations (σχέσεις)* (ΜΕ ΜΟΝΑΔΙΚΟ ΟΝΟΜΑ)
- Ένα **Σχισιακό Σχήμα** R είναι το όνομα και τα γνωρίσματα σε μια Σχέση, μαζί με τα αντίστοιχα πεδία τιμών για τα γνωρίσματα. Όταν είναι προφανές, τα πεδία τιμών παραλείπονται.
Συμβολισμός: $R(A_1, A_2, \dots, A_n)$
π.χ., ΦΟΙΤΗΤΗΣ (ΑΕΜ, ΕΠΙΘΕΤΟ, ΟΝΟΜΑ, ΤΗΛ)
- Ένα **Σχήμα Βάσης Δεδομένων** S είναι ένα σύνολο Σχέσεων.
Συμβολισμός: $S = \{R_1, R_2, \dots, R_m\}$
π.χ., ΤΕΙ = { ΦΟΙΤΗΤΕΣ, ΚΑΘΗΓΗΤΕΣ, ... }

Τυπικός Ορισμός

Μια **πλειάδα (Tuple)** t μιας Σχέσης $R(A_1, A_2, \dots, A_n)$ είναι μια (διατεταγμένη) λίστα τιμών $t = \langle v_1, v_2, \dots, v_n \rangle$, όπου κάθε τιμή v_i είναι ένα στοιχείο του πεδίου $D(A_i)$.

Ένα **στιγμιότυπο σχέσης (relation instance)** $r(R)$, πιο απλά, **σχέση (relation)**, είναι ένα σύνολο πλειάδων

$$r(R) = \{ t_1, t_2, \dots, t_k \}$$

Η **πληθικότητα (cardinality)** της R είναι ο αριθμός των πλειάδων στην $r(R)$, και συμβολίζεται με **CARD_R**

Χαρακτηριστικά των Σχέσεων

- Η ΔΙΑΤΑΞΗ των στηλών σε μια σχέση δεν είναι σημαντική
- Η ΔΙΑΤΑΞΗ των γραμμών σε μια σχέση δεν είναι σημαντική
- Κάθε εγγραφή αποθηκεύεται ΜΙΑ ΦΟΡΑ σε μια σχέση
- Μια τιμή μπορεί να παρουσιάζεται ΠΟΛΛΕΣ ΦΟΡΕΣ σε μια στήλη και είναι ΑΤΟΜΙΚΗ
- Τα ονόματα των στηλών είναι μοναδικά
- Σε κάθε στήλη οι τιμές είναι του ίδιου τύπου

- Μια ειδική τιμή, ονομαζόμενη ΚΕΝΟ (NULL), χρησιμοποιείται για την να παραστήσει στη βάση μιας τιμή που είναι **μη εφαρμόσιμος** (non-applicable) ή **άγνωστη** (unknown).
Π.χ, τηλέφωνο, Διεύθυνση δεν τα γνωρίζουμε
- Συμβολισμός: τιμή γνωρίσματος για μια πλειάδα t , $t[A_i] = v_i$

Δομικοί Περιορισμοί

- Υπάρχουν **τρία** είδη περιορισμών *που είναι έμφυτοι στο μοντέλο*:
 - ΚΛΕΙΔΙ (KEY),
 - ΑΚΕΡΑΙΟΤΗΤΑ ΟΝΤΟΤΗΤΑΣ (ENTITY INTEGRITY), και
 - ΑΝΑΦΟΡΙΚΗ ΑΚΕΡΑΙΟΤΗΤΑ (REFERENTIAL INTEGRITY.)
- Υπάρχουν **τρία** είδη *ρητών περιορισμών*:
 - ΠΕΔΙΟ ΤΙΜΩΝ (DOMAIN),
 - ΣΤΗΛΩΝ (COLUMN) και
 - ΟΡΙΖΟΜΕΝΩΝ από τον ΧΡΗΣΤΗ (USER-DEFINED)

Έμφυτοι Δομικοί Περιορισμοί

- **Περιορισμοί Κλειδιών:** Τα διαφορετικά κλειδιά, όπως ορίστηκαν στο μοντέλο E-R, ισχύουν και στο Σχεσιακό Μοντέλο.
Ένα σύνολο γνωρισμάτων **A B** σχήματος μιας σχέσης **R** για το οποίο κάθε πλειάδα σε στιγμιότυπο $r(R)$ πρέπει να έχει *μοναδική τιμή* (είναι ένα
 - υπέρ-κλειδί (superkey). Δηλαδή, για διαφορετικά t_1 και t_2 , ισχύει $t_1[AB] \neq t_2[AB]$
 - Ένα υποψήφιο κλειδί (candidate key) **K** είναι ένα *ελάχιστο* υπέρ-κλειδί (δηλαδή, δεν υπάρχει υποσύνολο του **K** που να είναι και αυτό υπέρ-κλειδί). Το **K** ονομάζεται συνήθως ΚΛΕΙΔΙ (key).
 - Ένα πρωτεύον κλειδί (primary key) **PK** είναι ένα από τα υποψήφια κλειδιά που συμφωνείται να παίξει το ρόλο του του προσδιοριστή για τις πλειάδες της σχέσης (τα πρωτεύοντα κλειδιά υπογραμμίζονται)

Έμφυτοι Δομικοί Περιορισμοί

- **ΑΚΕΡΑΙΟΤΗΤΑ ΟΝΤΟΤΗΤΑΣ (ENTITY INTEGRITY):**
Το κύριο κλειδί PK στο σχήμα της σχέσης R δεν μπορεί να έχει ΚΕΝΕΣ (NULL) τιμές σε πλειάδες μιας σχέσης $r(R)$.

$$t[PK] \neq \text{NULL}, \text{ για κάθε } t \text{ στην } r(R)$$

- Ο λόγος πίσω από αυτό τον περιορισμό είναι ότι το κύριο κλειδί χρησιμοποιείται για τον προσδιορισμό μιας εγγραφής σε μια Σχέση.
- Σημειώνεται ότι και άλλα γνωρίσματα στην R μπορεί να περιορίζονται στο να μην έχουν NULL από ΡΗΤΟΥΣ περιορισμούς.

Έμφυτοι Δομικοί Περιορισμοί

- **ΑΝΑΦΟΡΙΚΗ ΑΚΕΡΑΙΟΤΗΤΑ (REFERENTIAL INTEGRITY):**
Αυτός ο δομικός περιορισμός εμπλέκει ΔΥΟ σχέσεις και χρησιμοποιείται για να καταγράψει τη συνέπεια σε μια συσχέτιση μεταξύ εγγραφών των δυο σχέσεων.
- Η συνήθης μορφή είναι αυτή των **ξένων κλειδιών**.
 - Ένα **ξένο κλειδί (foreign key) FK** είναι ένα σύνολο γνωρισμάτων σε μια σχέση R1 που αποτελεί **κύριο κλειδί** σε μια άλλη σχέση R2.
Μια πλειάδα t_1 στην $r(R_1)$ λέγεται ότι **αναφέρεται** σε μια άλλη πλειάδα t_2 στην $r(R_2)$, εάν: $t_1[FK] = t_2[PK]$

Ρητοί Περιορισμοί

- **Περιορισμοί Πεδίου Τιμών:** Είναι οι κανόνες που ορίζονται για το πεδίο τιμών και κληρονομούνται από τις στήλες (γνωρίσματα) **που παίρνουν τιμές από το πεδίο.**
 - Το πεδίο μπορεί να οριστεί μαζί με κανόνες ακεραιότητας (π.χ., το πεδίο των *integers* με όλους τους κανόνες για ακέραιους). Αυτοί είναι (κυρίως) οι βασικοί **τύποι δεδομένων (*data types*.)**
- **Περιορισμοί Στηλών:** Είναι **επιπρόσθετοι** των περιορισμών πεδίου τιμών και αναφέρονται στις τιμές για τα γνωρίσματα.
 - Για παράδειγμα., η στήλη των *small integers* ή ***integers between 1 and 10, κλπ.*** είναι επιπλέον περιορισμοί των ακεραίων

USER-DEFINED

- **ΠΕΡΙΟΡΙΣΜΟΙ ΟΡΙΖΟΜΕΝΟΙ από τον ΧΡΗΣΤΗ (USER-DEFINED)**
 - επίσης stored procedures, triggers, methods (για object-oriented systems)*
- Γενικά, τα DBMS είναι **αδύνατα σε υποστήριξη περιορισμών**

Πράξεις

- Διαχωρίζονται σε (α) ΕΝΗΜΕΡΩΣΕΙΣ, (β) ΑΝΑΚΤΗΣΕΙΣ
- Το σύνολο των πράξεων στο Σχεσιακό Μοντέλο είναι **ΚΛΕΙΣΤΟ** δηλαδή οι πράξεις ορίζονται σε Σχέσεις και έχουν αποτέλεσμα νέες Σχέσεις
- **Ενημερώσεις (UPDATE) σε Σχέσεις**
 - Εισαγωγή (INSERT) πλειάδας
 - Διαγραφή (DELETE) πλειάδας
 - Τροποποίηση (MODIFY) πλειάδας
- Οι περιορισμοί ακεραιότητας δεν πρέπει να παραβιάζονται με την εκτέλεση μιας πράξης ενημέρωσης. Για αυτό, ενημερώσεις μπορεί είτε να **απορρίπτονται** ή να **διορθώνουν** με την εκτέλεση νέων **(επιπλέον) ενημερώσεων**.
 - Π.χ., όταν μια πλειάδα του EMPLOYEE διαγράφεται, όλες οι πλειάδες στην WORKING_ON που έχουν την ίδια τιμή στο SSN διαγράφονται **(ανύπαρκτοι υπάλληλοι δεν δουλεύουν σε έργα!)**

Άτυπα Κριτήρια Σχεδιασμού

- **Σχεδιασμός Σχέσεων:** Ομαδοποίηση Γνωρισμάτων έτσι ώστε να επιτευχθούν “καλά” σχήματα Σχέσεων (Σχέσεις Βάσης)
- **ΑΤΥΠΑ ΚΡΙΤΗΡΙΑ:**
 - Προσπάθεια για εννοιολογική καθαρότητα
 - Προσπάθεια για αποδοτικότητα χώρου (αποφυγή πλεονασμού)
 - Προσπάθεια για ακεραιότητα (αποφυγή ανωμαλιών ενημέρωσης)
 - Προσπάθεια για πληρότητα (αποφυγή NULL τιμών σε πλειάδες)
 - Προσπάθεια για γλωσσολογική αποδοτικότητα
 - Προσπάθεια για καλές Επιδόσεις (*performance*)

Σχεδιασμός Σχεσιακού Μοντέλου

- Α τρόπος: από εννοιολογικό σχεδιασμό (π.χ. με διάγραμμα E-R) στο λογικό σχεδιασμό (π.χ. Σχεσιακό μοντέλο)
 - Βήματα μετάβασης από το E-R σε πίνακες του σχεσιακού μοντέλου
- Β τρόπος: από το «μηδέν»
 - Ορισμός «καθολικής» (universal) σχέσης που τα περιέχει ΟΛΑ
 - Αποσύνθεση σε περισσότερες σχέσεις με βάση τη **Θεωρία Κανονικοποίησης**