

Το Σχεσιακό Μοντέλο

Μοντελοποίηση

Σχήμα (database schema): η περιγραφή της δομής της πληροφορίας που είναι αποθηκευμένη στη βδ με τη χρήση ενός μοντέλου δεδομένων

Μοντέλο Δεδομένων: ένα σύνολο από έννοιες (δομικά στοιχεία) που μπορούν να χρησιμοποιηθούν για την περιγραφή της δομής της πληροφορίας

Βήματα Σχεδιασμού

1. Συλλογή και Ανάλυση Απαιτήσεων (requirement analysis)

Τι δεδομένα θα αποθηκευτούν, ποιες εφαρμογές θα κτιστούν πάνω στα δεδομένα

Περιγραφή σε φυσική γλώσσα

2. Εννοιολογικός Σχεδιασμός/Μοντελοποίηση (conceptual design)

Υψηλού-επιπέδου περιγραφή:

(α) Δεδομένα (οντότητες και συσχετίσεις) που θα αποθηκευτούν στη βδ

(β) Τι είδους πληροφορία για αυτά θα αποθηκεύσουμε

(γ) Περιορισμοί ακεραιότητας (integrity constraints)

Χρήση Μοντέλου Οντοτήτων/Συσχετίσεων

Βήματα Σχεδιασμού

3. Λογικός Σχεδιασμός (ή Απεικόνιση των Μοντέλων Δεδομένων) (logical design)

- Επιλογή ενός ΣΔΒΔ για την υλοποίηση του σχεδιασμού
- Μετατροπή του εννοιολογικού σχεδιασμού σε ένα σχήμα στο μοντέλο δεδομένων του *επιλεγμένου ΣΔΒΔ*

χρήση Σχεσιακού Μοντέλου (πίνακες)

επίσης Κανονικοποίηση, π.χ., έλεγχοι πλεονασμού

Βελτίωση Σχήματος (Schema Refinement)

Σχήμα και Στιγμιότυπο

Σχήμα της Βάσης

Πρόθεση (intension)

Μοντέλο : (1) δομικά στοιχεία

(2) περιορισμοί ακεραιότητας

Ανάπτυξη (extension)

Στιγμιότυπο της Βάσης (κατάσταση ή σύνολο εμφανίσεων ή σύνολο στιγμιοτύπων)

(αρχική κατάσταση, έγκυρη κατάσταση)

Τι θα δούμε σήμερα

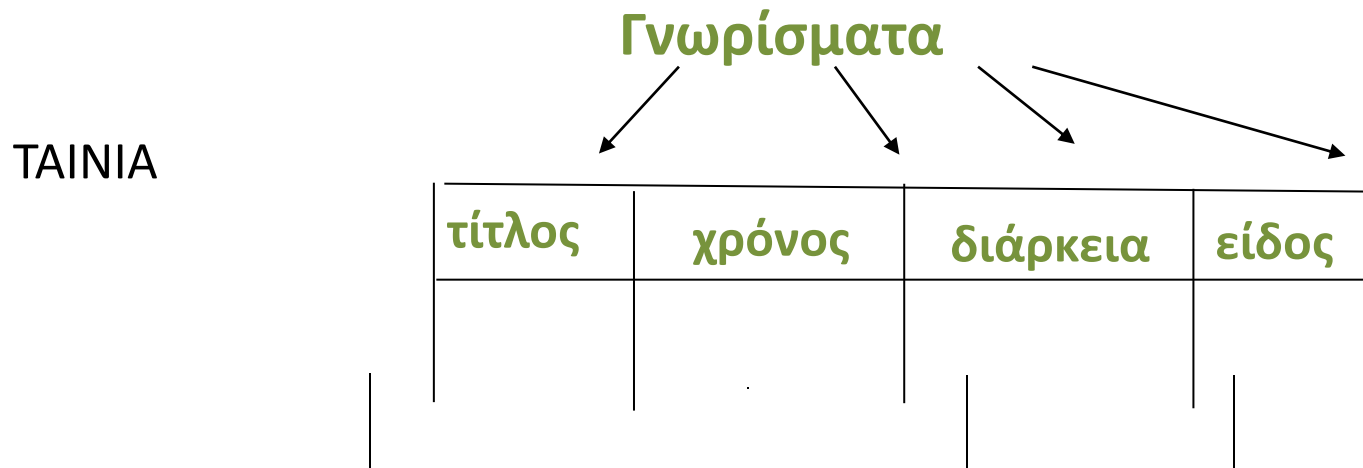
- I. Το Σχεσιακό Μοντέλο
- II. Μετατροπή/αντιστοιχία Σχήματος
O/Σ σε Σχεσιακό Σχήμα

Σχήμα Σχέσης

Βασικό δομικό στοιχείο είναι οι «πίνακες» ή «σχέσεις»

Σχήμα σχέσης R δηλώνεται $R(A_1, A_2, \dots, A_n)$ αποτελείται από ένα όνομα σχέσης και μια λίστα από γνωρίσματα.

Παράδειγμα - ΤΑΙΝΙΑ(τίτλος, χρόνος, διάρκεια, είδος)



Βαθμός: το πλήθος των γνωρισμάτων

Στιγμιότυπο Σχέσης

απλός τρόπος αναπαράστασης δεδομένων: ένας δισδιάστατος πίνακας που λέγεται **σχέση**

ΤΑΙΝΙΑ

Γνωρίσματα

τίτλος	χρόνος	διάρκεια	είδος
Star Wars	1997	124	έγχρωμη
Mighty Ducks	1991	104	έγχρωμη
Wayne's World	1992	95	έγχρωμη

Οι γραμμές της σχέσης (εκτός της επικεφαλίδας) ονομάζονται **πλειάδες**.

Στιγμιότυπο: Σύνολο από Πλειάδες

Παράδειγμα: {(Star Wars, 1997, 124, έγχρωμη), (Mighty Ducks, 1991, 104, έγχρωμη), (Wayne's World, 1992, 95, έγχρωμη)}

Σχήμα - Στιγμιότυπο

Σχήμα σχέσης R που δηλώνεται $R(A_1, A_2, \dots, A_n)$ αποτελείται από ένα όνομα σχέσης και μια λίστα από γνωρίσματα.

Μία σχέση r ή $r(R)$ (ή ένα στιγμιότυπο r του σχήματος σχέσης R) είναι ένα σύνολο από πλειάδες.

Πεδίο Ορισμού

Κάθε γνώρισμα A_i παίρνει τιμές από κάποιο σύνολο D που ονομάζεται **πεδίο ορισμού** του A_i και συμβολίζεται με $dom(A_i)$

Το γνώρισμα είναι το όνομα ενός ρόλου που παίζει κάποιο πεδίο ορισμού D στο σχήμα σχέσης R

Το πεδίο ορισμού D είναι ένα σύνολο από ατομικές τιμές

(παράδειγμα: ακέραιοι, συμβολοσειρές - όχι εγγραφές, πίνακες, λίστες)

Κάθε τιμή γνωρίσματος μιας πλειάδας ατομική.

Πλειάδες

Κάθε **πλειάδα** είναι μια διατεταγμένη λίστα από τιμές $\langle v_1, v_2, \dots, v_n \rangle$ όπου κάθε τιμή v_i είναι ένα στοιχείο του $\text{dom}(A_i)$ ή η ειδική τιμή null

Κάθε **σχέση** είναι ένα υποσύνολο του καρτεσιανού γινομένου:

$$r(R) \subseteq \text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n)$$

Παρατηρήσεις

- ✓ Διάταξη των πλειάδων σε μια σχέση
- ✓ Διάταξη των γνωρισμάτων στο σχήμα σχέσης

Συμβολισμός

Σχήμα σχέσης βαθμού n $R(A_1, A_2, \dots, A_n)$

Πλειάδα t της σχέσης $r(R) \langle v_1, v_2, \dots, v_n \rangle$

- ✓ αναφορά στις συνιστώσες τιμές $t[A_i]$
- ✓ $t[A_u, A_w, \dots, A_z]$
- ✓ όνομα γνωρίσματος $t.A_i$

Σχήμα Σχεσιακής Βάσης Δεδομένων

Σχήμα μιας σχεσιακής βάσης δεδομένων είναι ένα σύνολο από σχήματα σχέσεων

Παράδειγμα - ΤΑΙΝΙΑ(τίτλος, χρόνος, διάρκεια, είδος)

ΗΘΟΠΟΙΟΣ(όνομα, διεύθυνση, έτος-γέννησης)

ΠΑΙΖΕΙ(όνομα_ηθοποιού, τίτλος, χρόνος)

Παράδειγμα

Θεωρείστε μια βάση δεδομένων για ταινίες και ηθοποιούς όπου κρατάμε

- το έτος γέννησης, διεύθυνση, και όνομα (που είναι μοναδικό) για τους *Ηθοποιούς*
- το τίτλο, έτος, διάρκεια και είδος για τις *Ταινίες*. Ο τίτλος μιας ταινίας δεν είναι μοναδικός, αλλά υπάρχει μόνο μια ταινία με τον ίδιο τίτλο κάθε έτος.
- ποιος ηθοποιός έπαιξε σε ποια ταινία και ποιο ρόλο

Σχεσιακό Σχήμα

ΤΑΙΝΙΑ

Τίτλος	Έτος	Διάρκεια	Είδος
--------	------	----------	-------

ΗΘΟΠΟΙΟΣ

Όνομα	Διεύθυνση	Έτος-Γέννησης
-------	-----------	---------------

ΠΑΙΖΕΙ

Όνομα-Ηθοποιού	Τίτλος	Έτος
----------------	--------	------

Παράδειγμα

Θεωρείστε μια βάση δεδομένων για το φοιτητολόγιο για ένα πανεπιστήμιο που να περιέχει τις παρακάτω πληροφορίες:

- το όνομα, διεύθυνση, αριθμό μητρώου (που είναι μοναδικός) για τους *Φοιτητές*
- το όνομα, κωδικό (που είναι μοναδικός), μονάδες, εξάμηνο για τα *Μαθήματα*
- το *βαθμό* που πήρε ένας φοιτητής σε κάποιο μάθημα

Υποθέστε ότι καταγράφεται μόνο ένας (ο τελικός βαθμός) του φοιτητή στο μάθημα

Περιορισμός Κλειδιού

Μια σχέση ορίζεται ως ένα σύνολο πλειάδων, άρα όλες οι πλειάδες πρέπει να είναι διαφορετικές.

(Υπερ)-κλειδί είναι ένα υποσύνολο γνωρισμάτων του σχήματος σχέσης R τέτοια ώστε σε κάθε στιγμιότυπο $r(R)$, κανένα ζευγάρι πλειάδων δε μπορεί να έχει τον ίδιο συνδυασμό τιμών για τα γνωρίσματα αυτά, δηλαδή, K είναι κλειδί, αν

δεν μπορούν να υπάρχουν σε οποιοδήποτε στιγμιότυπο της σχέσης δύο διαφορετικές πλειάδες t_1 και t_2 , για τις οποίες $t_1[K] = t_2[K]$

Περιορισμός Κλειδιού

(υπερ)κλειδί - υποψήφιο κλειδί - πρωτεύον κλειδί

υποψήφιο κλειδί K : κλειδί με την ιδιότητα ότι αν αφαιρεθεί ένα οποιοδήποτε γνώρισμα A από το K , το K' που προκύπτει δεν είναι κλειδί

Συμβολισμός: υπογραμμίζουμε τα γνώρισμα του πρωτεύοντος κλειδιού

✓ Κάθε σχέση έχει τουλάχιστον ένα υπερ-κλειδί, ποιο;

Από τον ορισμό, κάθε (σχήμα) σχέσης έχει τουλάχιστον ένα (πρωτεύον) κλειδί – δεν υπάρχουν «ασθενείς» σχέσεις

Περιορισμός Κλειδιού

Υποθέσεις:

(1) Το όνομα του ηθοποιού είναι μοναδικό

(2) Ο τίτλος μιας ταινίας δεν είναι μοναδικός, αλλά μόνο μια ταινία με τον ίδιο τίτλο κάθε έτος

(3) Σε μια ταινία μπορεί να παίζουν πολλοί ηθοποιοί και ένα ηθοποιός μπορεί να παίζει σε πολλές ταινίες

ΤΑΙΝΙΑ

Τίτλος	Έτος	Διάρκεια	Είδος
--------	------	----------	-------

ΗΘΟΠΟΙΟΣ

Όνομα	Διεύθυνση	Έτος-Γέννησης
-------	-----------	---------------

ΠΑΙΖΕΙ

Όνομα-Ηθοποιού	Τίτλος	Έτος
----------------	--------	------

Περιορισμός Κλειδιού

ΤΑΙΝΙΑ

<u>Τίτλος</u>	<u>Έτος</u>	Διάρκεια	Είδος
---------------	-------------	----------	-------

ΗΘΟΠΙΟΣ

<u>Όνομα</u>	Διεύθυνση	Έτος-Γέννησης
--------------	-----------	---------------

ΠΑΙΖΕΙ

<u>Όνομα-Ηθοποιού</u>	<u>Τίτλος</u>	<u>Έτος</u>
-----------------------	---------------	-------------

Περιορισμός Κλειδιού

Έστω το παρακάτω στιγμιότυπο ενός σχήματος σχέσης $R(A, B, C, D)$

A	B	C	D
6	7	1	1
1	7	7	2
3	7	7	1
1	5	9	2

Τι μπορείτε να πείτε για τα κλειδιά της R;

*Ο περιορισμός του κλειδιού αφορά το σχήμα, από ένα στιγμιότυπο, μπορούμε να πούμε ποια σύνολα γνωρισμάτων **δεν** έχουν την ιδιότητα του κλειδιού, αλλά δεν μπορούμε να πούμε ποια την έχουν*

Περιορισμός Ακεραιότητας Οντοτήτων

Δε μπορεί η τιμή του πρωτεύοντος κλειδιού (οποιοδήποτε γνωρίσματος που ανήκει στο κλειδί) να είναι null.

Περιορισμός Αναφορικής Ακεραιότητας

Έστω δύο σχήματα σχέσεων $R_1(X)$ και $R_2(Y)$, ένα σύνολο γνωρισμάτων F της R_1 είναι *ξένο κλειδί* που *αναφέρεται* στην R_2 αν

(1) το σύνολο F αποτελείται από το ίδιο πλήθος και με το ίδιο πεδίο ορισμού γνωρισμάτων όπως και το *πρωτεύον κλειδί* K της R_2 και

(2) σε οποιοδήποτε στιγμιότυπο, για μια πλειάδα t_1 της R_1 ισχύει ότι είτε

(α) όλα τα γνωρίσματα F της t_1 έχουν την τιμή null είτε

(β) στο ίδιο στιγμιότυπο, υπάρχει μια πλειάδα t_2 της R_2 , τέτοια ώστε $t_1[F] = t_2[K]$.

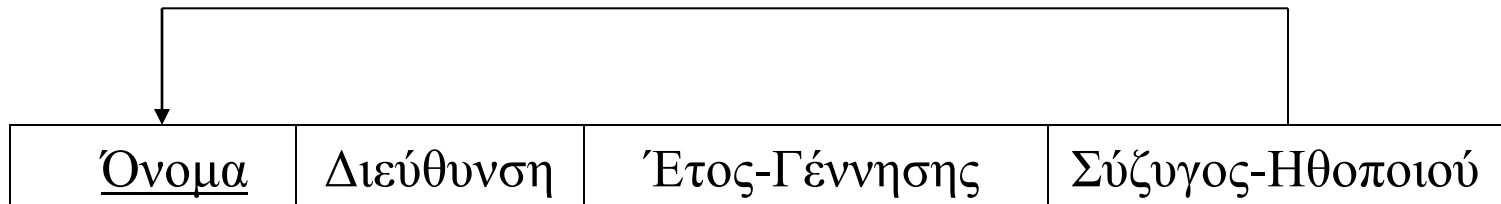
Λέμε ότι η πλειάδα t_1 της R_1 αναφέρεται στην πλειάδα t_2 της R_2 .

Η R_2 καλείται *αναφερόμενη* σχέση και η R_1 *αναφέρουσα* σχέση.

Περιορισμός Αναφορικής Ακεραιότητας

- ✓ Συνήθως προκύπτουν από συσχετίσεις μεταξύ οντοτήτων
- ✓ Το ξένο κλειδί μπορεί να αναφέρεται στη δική του σχέση (συνήθως, προκύπτει από αναδρομική συσχέτιση)

ΗΘΟΠΟΙΟΣ



Περιορισμός Σημασιολογικής Ακεραιότητας

Παραδείγματα:

- ο μισθός ενός εργαζομένου δεν μπορεί να υπερβαίνει το μισθό του προϊσταμένου του
- ο μέγιστος αριθμός ωρών που ένας εργαζόμενος μπορεί να απασχοληθεί σε όλα τα έργα ανά εβδομάδα είναι 56.

Περιορισμοί Ακεραιότητας (integrity constraints)

- Περιορισμός **Πεδίου Ορισμού**: Η τιμή κάθε γνωρίσματος A πρέπει να είναι μία *ατομική* τιμή από το πεδίο ορισμού αυτού του γνωρίσματος $\text{dom}(A)$
- Περιορισμός **Κλειδιού**
- Περιορισμός **Ακεραιότητας Οντοτήτων**: Δε μπορεί η τιμή του πρωτεύοντος κλειδιού να είναι null
- Περιορισμός **Αναφορικής Ακεραιότητας**
- Περιορισμός **Σημασιολογικής Ακεραιότητας**

Σχεσιακό Σχήμα

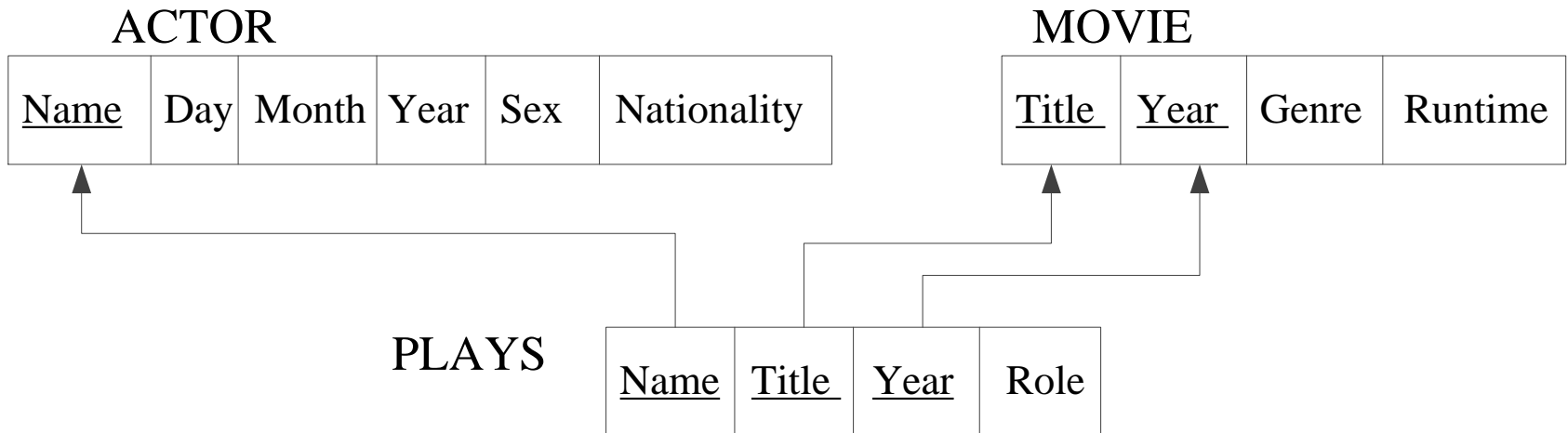
Ένα **σχεσιακό σχήμα βάσης** δεδομένων είναι ένα σύνολο από σχήματα σχέσεων $\Sigma = \{R_1, R_2, \dots, R_n\}$ και ένα σύνολο από περιορισμούς ακεραιότητας.

Ένα **στιγμιότυπο μιας σχεσιακής βάσης δεδομένων** ΒΔ του Σ είναι ένα σύνολο από στιγμιότυπα σχέσεων (σχέσεις) $B\Delta = \{r_1, r_2, \dots, r_n\}$ τέτοια ώστε κάθε r_i είναι ένα στιγμιότυπο του R_i που ικανοποιεί τους περιορισμούς ορθότητας (πεδίου ορισμού, κλειδιού, ακεραιότητας οντοτήτων, και αναφορικής ακεραιότητας)

Προσοχή: οι περιορισμοί ακεραιότητας πρέπει να ισχύουν σε κάθε στιγμιότυπο.

Σχήμα μιας Βάσης Δεδομένων

ΠΡΟΣΟΧΗ - το παρακάτω σχήμα για ταινίες είναι διαφορετικό από αυτό στις προηγούμενες διαφάνειες



Στιγμιότυπο του σχήματος

ACTOR

Name	Day	Month	Year	Sex	Nationality
George Clooney	6	May	1961	Male	American
Emmanuelle Riva	24	Feb	1927	Female	French
Sandra Bullock	26	Jul	1964	Female	American
Brad Pitt	18	Dec	1963	Male	American
Frank Sinatra	12	Dec	1915	Male	American

MOVIE

Title	Year	Genre	Runtime
Amour	2012	drama	117
Gravity	2013	science-fiction	91
Ocean's Eleven	1960	crime	127
Frozen	2013	comedy	102
Ocean's Eleven	2001	crime	116

PLAYS

Name	Title	Year	Role
George Clooney	Gravity	2013	Mark Kowalski
George Clooney	Ocean's Eleven	2001	Danny Ocean
Sandra Bullock	Gravity	2013	Ryan Stone
Emmanuelle Riva	Amour	2012	Anne
Frank Sinatra	Ocean's Eleven	1960	Danny Ocean
Brad Pitt	Ocean's Eleven	2001	Rusty Ryan

Ερωτήσεις;