

Σχεδίαση Σχεσιακών ΒΔ



- Εννοιολογικός σχεδιασμός
 - το Μοντέλο Οντοτήτων-Συσχετίσεων (ER)
 - Από τα διαγράμματα ER στο Σχεσιακό Μοντέλο
- Θεωρία Κανονικοποίησης (Κανονικές Μορφές – ΚΜ)
 - 1^η ΚΜ(1NF)
 - Αποσύνθεση - Συναρτησιακές Εξαρτήσεις
 - 2^η (2NF), 3^η (3NF), Boyce-Codd ΚΜ (BCNF)
 - Καθολική Διαδικασία Σχεδίασης ΒΔ

InfoLab | Τμήμα Πληροφορικής | Πανεπιστήμιο Πειραιώς (<http://infolab.cs.unipi.gr/>)

έκδοση: Ιαν. 2012

Βασική πηγή διαφανειών: Silberschatz et al., "Database System Concepts", 4/e



- Εννοιολογικός σχεδιασμός
 - το Μοντέλο Οντοτήτων-Συσχετίσεων (ER)
 - Από τα διαγράμματα ER στο Σχεσιακό Μοντέλο
- Θεωρία Κανονικοποίησης (Κανονικές Μορφές – ΚΜ)
 - 1^η ΚΜ(1NF)
 - Αποσύνθεση - Συναρτησιακές Εξαρτήσεις
 - 2^η (2NF), 3^η (3NF), Boyce-Codd ΚΜ (BCNF)
 - Καθολική Διαδικασία Σχεδίασης ΒΔ

Διαδικασία σχεδίασης ΒΔ



- **Σχεδιασμός ΒΔ** είναι η διαδικασία δημιουργίας του **σχεσιακού σχήματος** (relational schema) της ΒΔ
- Υπάρχουν 2 «σχολές» σχεδίασης σχεσιακών ΒΔ
 - Μέσω του Εννοιολογικού Σχεδιασμού. Βήματα:
 - Διατύπωση ενός **εννοιολογικού μοντέλου δεδομένων** (Conceptual Data Model) σε μορφή **Διαγράμματος Οντοτήτων – Συσχετίσεων** (Entity-Relationship diagram)
 - Μετατροπή του διαγράμματος ER σε Σχεσιακό Σχήμα
 - Με χρήση **Κανονικών Μορφών** (Normal Forms) και εφαρμογή της **Θεωρίας Κανονικοποίησης** (Normalization theory)

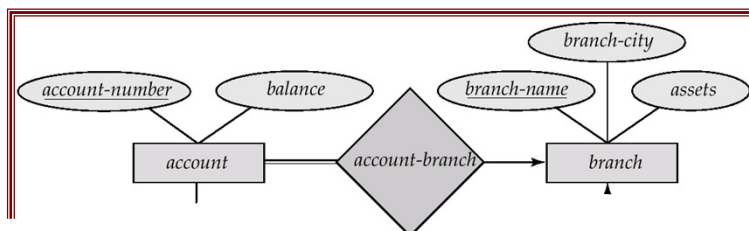
Εννοιολογικός Σχεδιασμός



(conceptual design)

- Προσπάθεια για ξεκαθάρισμα των εννοιών
- Εύρεση και καταγραφή των Οντοτήτων της ΒΔ, των μεταξύ τους Συσχετίσεων καθώς και της σημασιολογίας που τα συνοδεύει
- ΣΤΟΧΟΣ:
 - Μια **αφαιρετική**, αλλά **πλήρης** περιγραφή του τμήματος του μικρόκοσμου που θα αναπαρασταθεί στην βάση δεδομένων.
 - Αυτή η περιγραφή γίνεται με τη χρήση μιας ημι-τυπικής σημειογραφίας / συμβολισμού.

Παράδειγμα εννοιολογικού σχεδιασμού



- Μοντέλο οντοτήτων-συσχετίσεων (Entity-Relationship model – ER)
 - Τα υποκαταστήματα (branches) ανοίγουν λογαριασμούς καταθέσεων (accounts)
 - για τους καταθετικούς λογαριασμούς, καταγράφουμε κωδικό, υπόλοιπο
 - για τα υποκαταστήματα, καταγράφουμε όνομα, έδρα, αποθεματικό

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

5

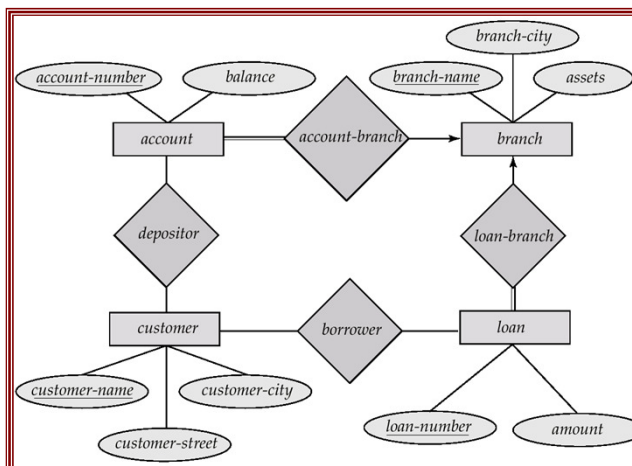
ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο Οντοτήτων-Συσχετίσεων



E-R (Entity-Relationship) model

- Ένα απλό και σαφές μοντέλο εννοιολογικού σχεδιασμού που στηρίζεται στο γραφικό συμβολισμό
- **διαισθητικό** μοντέλο που προσδιορίζει **αφαιρετικά** τις πληροφορίες που μια ΒΔ αποθηκεύει και οργανώνει



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

6

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Συστατικά του E-R Μοντέλου (1)



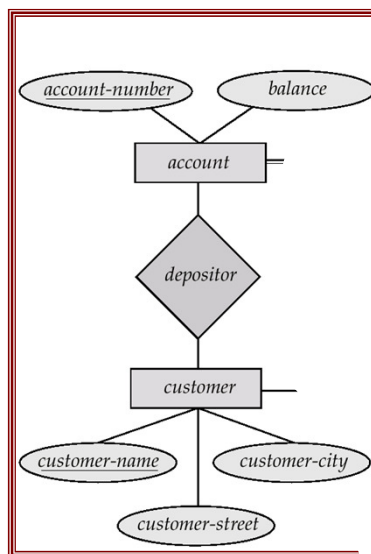
- Υπάρχουν δυο βασικές εννοιολογικές έννοιες:
- **Οντότητες (entities)**
 - Συγκεκριμένα αντικείμενα που υπάρχουν (ή πιστεύεται ότι υπάρχουν) και μπορούν να αναπαρασταθούν στην ΒΔ
 - π.χ., ο ΠΕΛΑΤΗΣ “Νίκος Νικολάου”, το ΥΠΟΚΑΤΑΣΤΗΜΑ “Ομόνοιας”, το ΔΑΝΕΙΟ “L-123”, ...
 - Για κάθε οντότητα καταγράφουμε ορισμένα **χαρακτηριστικά (attributes)**
- **Συσχετίσεις (relationships)**
 - Είναι επίσης (ειδικά) αντικείμενα που αντιστοιχούν δύο ή περισσότερες ξεχωριστές οντότητες με ένα συγκεκριμένο νόημα (τυπικά, μια Συσχέτιση είναι ένα διατεταγμένο σύνολο οντοτήτων)
 - π.χ., το δάνειο “L-123” ΔΟΘΗΚΕ από το υποκατάστημα “Ομόνοιας” κλπ.

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

7

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Συστατικά του E-R Μοντέλου (2)



- Οι Οντότητες μπορούν να έχουν **χαρακτηριστικά (attributes)** που είναι ιδιότητες που τα χαρακτηρίζουν
 - π.χ. μια οντότητα ΠΕΛΑΤΗΣ έχει χαρακτηριστικά: όνομα, διεύθυνση, ...
- ... αλλά και οι Συσχετίσεις μπορούν να έχουν χαρακτηριστικά
 - π.χ. αν υπήρχε κύριος / δευτερεύων κάτοχος του λογαριασμού καταθέσεων, αυτή η πληροφορία θα αποτελούσε χαρακτηριστικό της συσχέτισης depositor

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

8

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Συστατικά του E-R Μοντέλου (3)



- Είδη χαρακτηριστικών
 - **Απλά (simple)**: μια οντότητα έχει ατομική τιμή για αυτό
 - π.χ., υπόλοιπο λογαριασμού
 - **Σύνθετα (composite)**: το χαρακτηριστικό αποτελείται από 2 ή περισσότερα τμήματα
 - π.χ., διεύθυνση = {Δρόμος, Αριθμός, ΤΚ, Πόλη, Χώρα}
 - **Πλειότιμα (multi-valued)**: το χαρακτηριστικό έχει πολλαπλές τιμές
 - π.χ., τηλέφωνο(-α) ενός πελάτη

Οντότητες customer και loan



- 2 οντότητες με τα χαρακτηριστικά τους

321-12-3123	Jones	Main	Harrison
019-28-3746	Smith	North	Rye
677-89-9011	Hayes	Main	Harrison
555-55-5555	Jackson	Dupont	Woodside
244-66-8800	Curry	North	Rye
963-96-3963	Williams	Nassau	Princeton
335-57-7991	Adams	Spring	Pittsfield

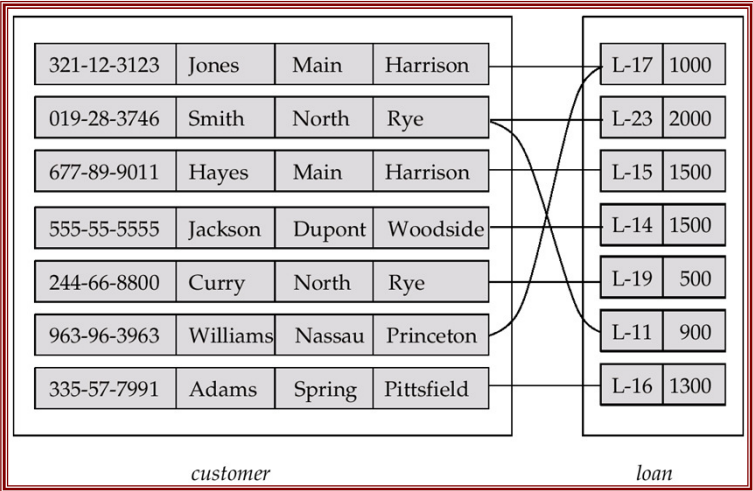
customer

L-17	1000
L-23	2000
L-15	1500
L-14	1500
L-19	500
L-11	900
L-16	1300

loan

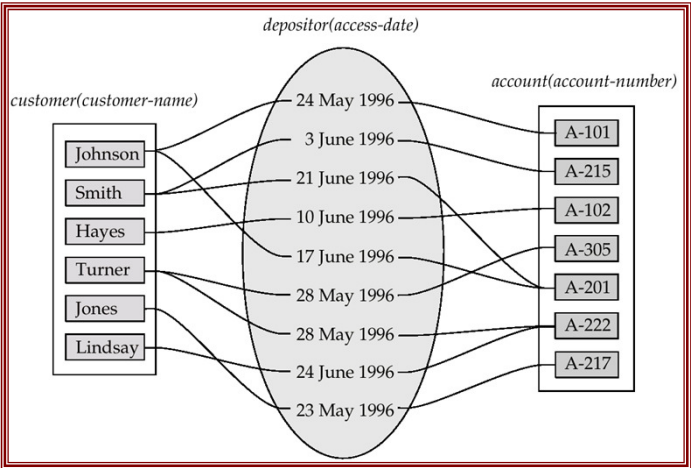
Συσχέτιση borrower

- Συσχέτιση μεταξύ των οντοτήτων customer και loan

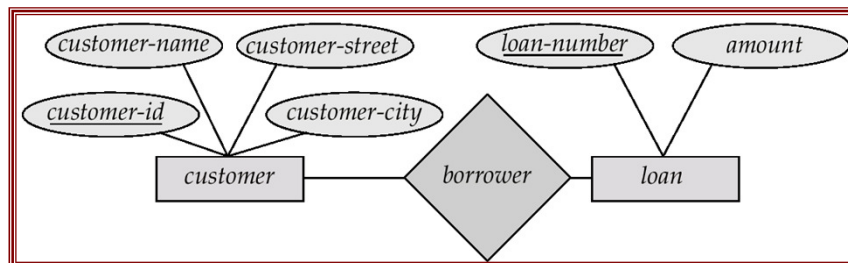


Χαρακτηριστικό πάνω σε συσχέτιση

- π.χ. ημερομηνία τελευταίας ενημέρωσης του λογαριασμού από τον κάτοχό του



Διάγραμμα Οντοτήτων-Συσχετίσεων (E-R)



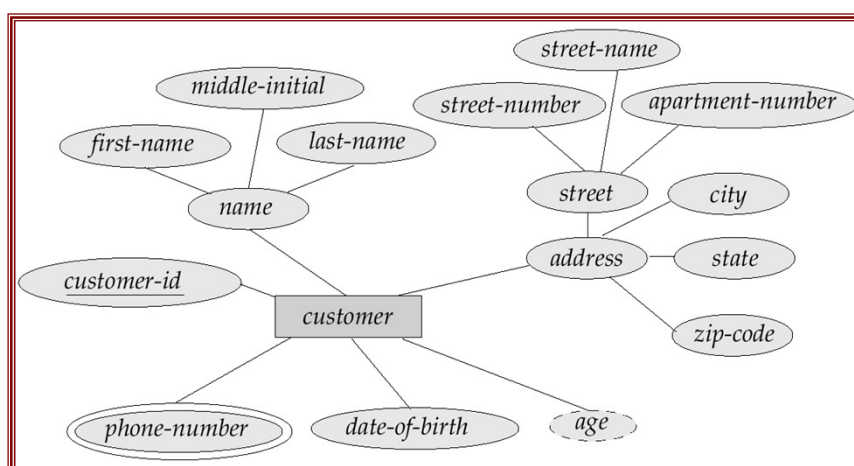
- **Ορθογώνια:** οντότητες
- **Ρόμβοι:** συσχετίσεις
- **Γραμμές:** συνδέουν χαρακτηριστικά με οντότητες, οντότητες με συσχετίσεις
- **Ελλείψεις:** χαρακτηριστικά
 - **Διπλές ελλείψεις:** πλειότιμα χαρακτηριστικά
 - **Διακεκομμένες ελλείψεις:** εξαρτημένα χαρακτηριστικά
- **Υπογραμμίσεις:** πρωτεύοντα κλειδιά

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

13

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σύνθετα, πλειότιμα, εξαρτημένα χαρακτηριστικά

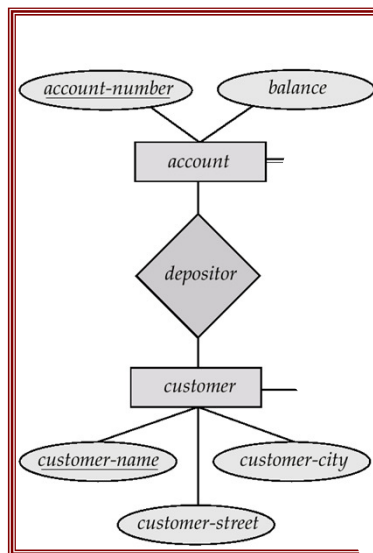


ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

14

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Ιδιότητες τύπων συσχετίσεων (1)



- Ένας τύπος συσχέτισης έχει **βαθμό (degree)**
 - Συσχέτιση μεταξύ 2 συνόλων οντοτήτων → **διμερής** (binary relationship)
 - Συσχέτιση μεταξύ 3 συνόλων οντοτήτων → **τριμερής** (ternary relationship)
 - Συσχέτιση μεταξύ N συνόλων οντοτήτων → **N-μερής** (N-ary relationship)

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

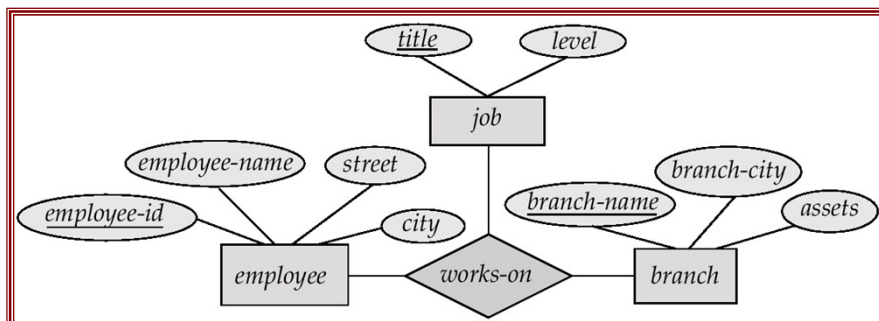
15

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα τριμερούς συσχέτισης



- Ένας υπάλληλος (employee) έχει διαφορετική απασχόληση (job) σε διαφορετικά υποκαταστήματα (branch)
 - π.χ. ο υπάλληλος Γιάννης απασχολείται ως πωλητής στο υποκατάστημα Α και ως ταμίας στο υποκατάστημα Β



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

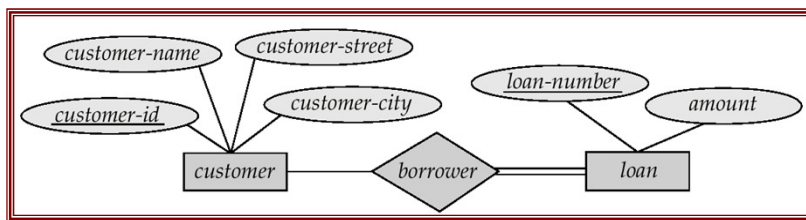
16

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Ιδιότητες τύπων συσχετίσεων (2)



- Ο περιορισμός ύπαρξης ορίζει αν η **συμμετοχή** μιας οντότητας στον τύπο συσχέτισης είναι **ολική (total)** ή **μερική (partial)**
 - Ολική συμμετοχή (συμβολισμός στο E-R: διπλή γραμμή): κάθε οντότητα συμμετέχει σε τουλάχιστον μια συσχέτιση
 - η συμμετοχή του loan στο borrower είναι ολική
 - Μερική συμμετοχή (συμβολισμός στο E-R: απλή γραμμή): μερικές οντότητες μπορεί να μη συμμετέχουν σε καμία συσχέτιση
 - η συμμετοχή του customer στο borrower είναι μερική

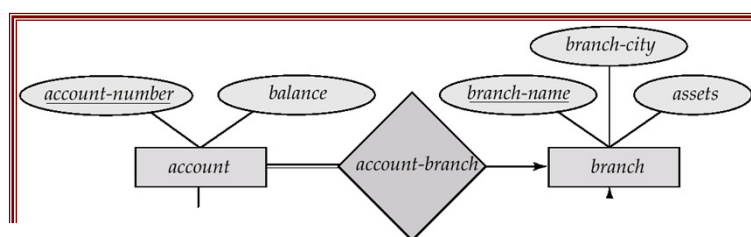


ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

17

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Ιδιότητες τύπων συσχετίσεων (3)



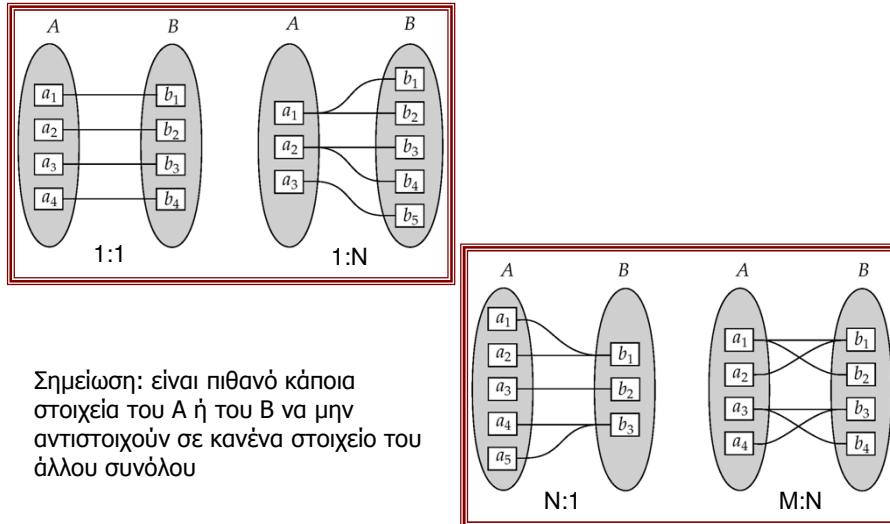
- Η **πληθικότητα (cardinality)** ενός τύπου συσχέτισης ορίζει το πόσες οντότητες από το πρώτο σύνολο οντοτήτων στην συσχέτιση μπορούν να συνδεθούν με πόσες οντότητες από το δεύτερο σύνολο οντοτήτων
- π.χ. ο τύπος συσχέτισης account-branch είναι πολλά-προς-ένα
 - περισσότεροι του ενός λογαριασμοί (μπορούν να) αντιστοιχούν σε κάποιο υποκατάστημα
 - ένα μόνο υποκατάστημα αντιστοιχεί σε κάποιο λογαριασμό

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

18

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Πληθικότητα συσχέτισης (4 περιπτώσεις)



Σημείωση: είναι πιθανό κάποια στοιχεία του A ή του B να μην αντιστοιχούν σε κανένα στοιχείο του άλλου συνόλου

Η έννοια του κλειδιού

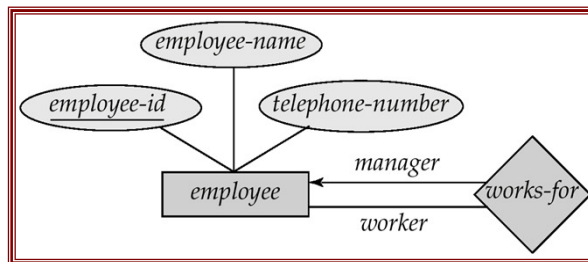


- Ένα χαρακτηριστικό (ή σύνολο χαρακτηριστικών) ενός τύπου οντοτήτων / συσχετίσεων για τον οποίο κάθε οντότητα / συσχέτιση στο σύνολο πρέπει να έχει μοναδική τιμή ονομάζεται **κλειδί (key)**.
 - π.χ. ΑΦΜ του ΥΠΑΛΛΗΛΟΥ, {ΑΦΜ, ΟΝΟΜΑ} του ΥΠΑΛΛΗΛΟΥ, κλπ.
- υποψήφιο κλειδί (candidate key)** ονομάζεται ένα **ελάχιστο (minimal) κλειδί** (δηλαδή, κανένα υποσύνολο των χαρακτηριστικών του δεν είναι κλειδί)
 - π.χ. ΑΦΜ είναι Υποψήφιο Κλειδί για τον ΥΠΑΛΛΗΛΟ, αλλά ο συνδυασμός {ΑΦΜ, ΟΝΟΜΑ} δεν είναι.
- Το **κύριο κλειδί (primary key)** είναι ένα από τα υποψήφια κλειδιά που ορίζεται σαν **αναγνωριστής (identifier)** για τον τύπο οντοτήτων / συσχετίσεων
 - π.χ. ΑΦΜ είναι μια καλή επιλογή Κυρίου Κλειδιού στον τύπο οντοτήτων ΥΠΑΛΛΗΛΟΣ

Η έννοια του ρόλου



- Οι οντότητες που συσχετίζονται μέσω μιας συσχέτισης δε χρειάζεται να είναι διακριτές
- Οι ετικέτες “manager” and “worker” ονομάζονται **ρόλοι (roles)**. Δείχνουν πώς οι υπάλληλοι αλληλεπιδρούν μέσω της συσχέτισης works-for
- Οι ρόλοι εμφανίζονται στα διαγράμματα E-R με ετικέτες πάνω στις γραμμές που συνδέουν ρόμβους με ορθογώνια.
- Οι ετικέτες ρόλων είναι προαιρετικές (ξεκαθαρίζουν τη σημασιολογία της συσχέτισης)



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

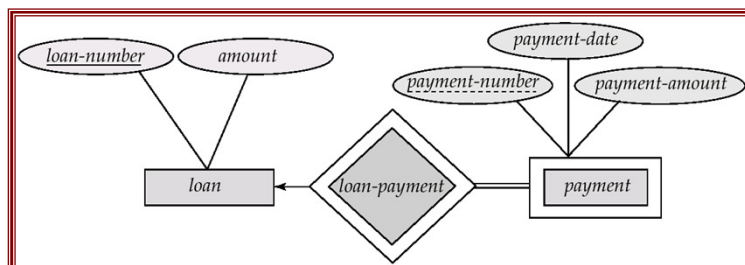
21

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Αδύναμα σύνολα οντοτήτων



- Αδύναμο ονομάζεται ένα σύνολο οντοτήτων όταν η ύπαρξή του εξαρτάται από την ύπαρξη ενός άλλου (ισχυρού) συνόλου οντοτήτων
 - συμβολισμός στο E-R: διπλό ορθογώνιο για το σύνολο οντοτήτων, διακεκομμένη υπογράμμιση για το πρωτεύον κλειδί του
 - π.χ. payment: δεν υπάρχει αν δεν υπάρχει το αντίστοιχο loan



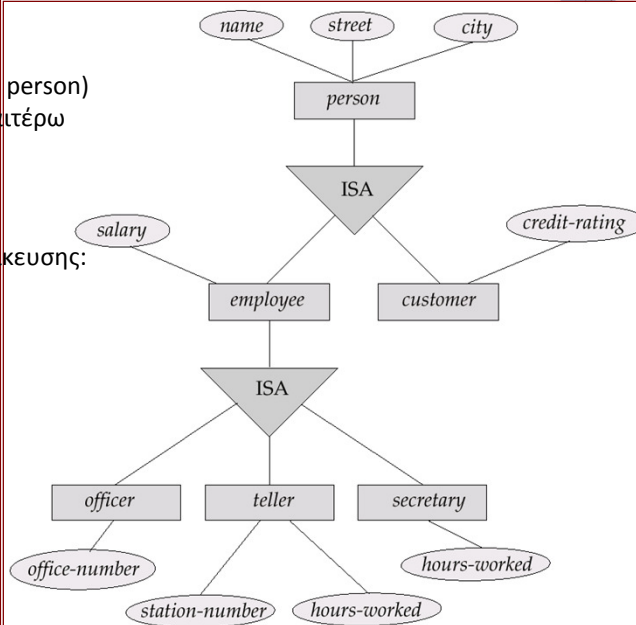
ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

22

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Εξειδίκευση (Specialization)

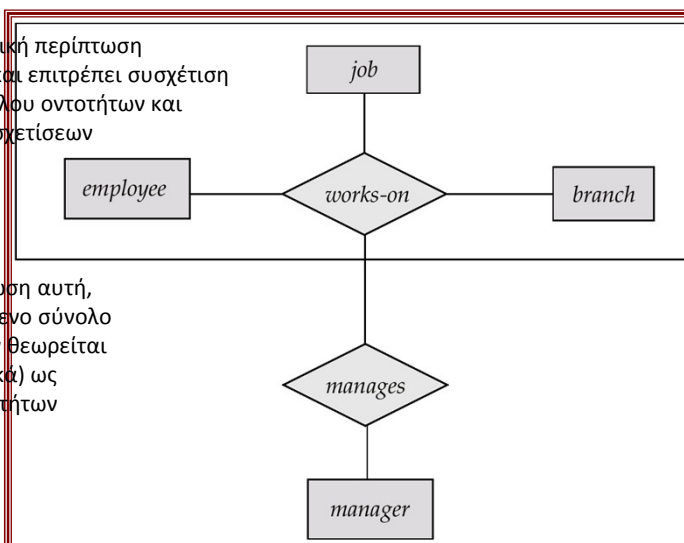
- Μια οντότητα (π.χ. person) εξειδικεύεται περαιτέρω σε άλλες οντότητες (π.χ. employee ή customer)
- Περιπτώσεις εξειδίκευσης:
 - Πλήρης ή μη
 - Διακριτή ή επικαλυπτόμενη



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

Συσώρευση (Aggregation)

- Αποτελεί ειδική περίπτωση συσχέτισης και επιτρέπει συσχέτιση μεταξύ συνόλου οντοτήτων και συνόλου συσχετίσεων
- Στην περίπτωση αυτή, το εμπλεκόμενο σύνολο συσχετίσεων θεωρείται (καταχρηστικά) ως σύνολο οντοτήτων

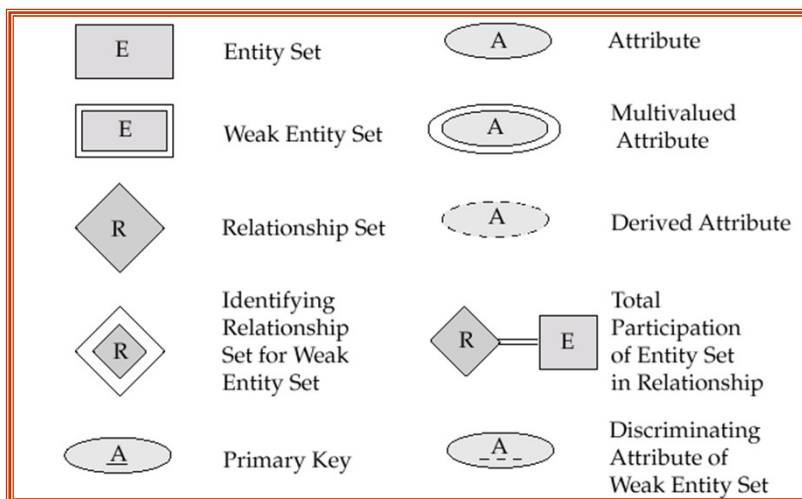


ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

24

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σύμβολα που χρησιμοποιούνται στα διαγράμματα E-R (1)

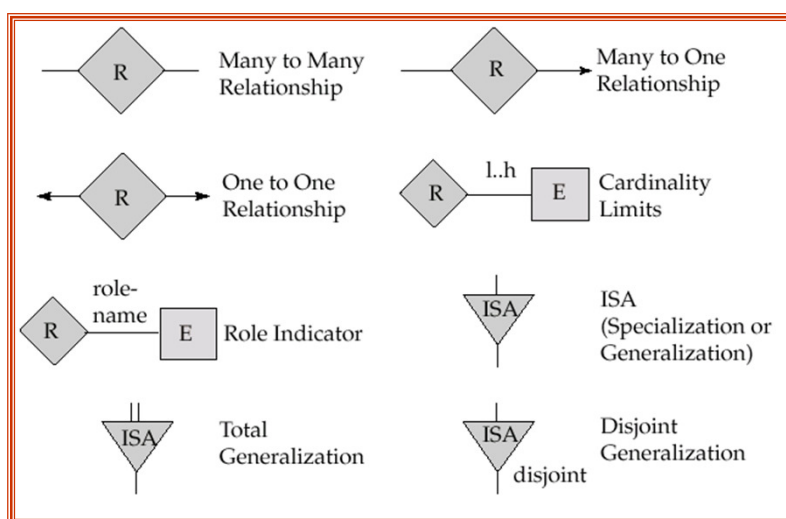


ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

25

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σύμβολα που χρησιμοποιούνται στα διαγράμματα E-R (2)

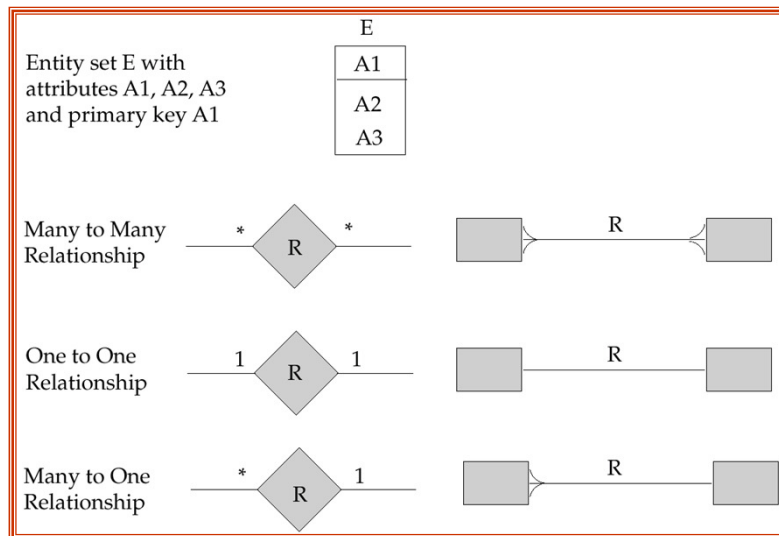


ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

26

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σύμβολα που χρησιμοποιούνται στα διαγράμματα E-R (3)



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

27

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μετατροπή σχήματος E-R σε σχεσιακό σχήμα



- Τα πρωτεύοντα κλειδιά επιτρέπουν να εκφραστούν τα σύνολα οντοτήτων και τα σύνολα συσχετίσεων ως πίνακες που αναπαριστούν τα περιεχόμενα μιας βάσης δεδομένων
- Μια βάση δεδομένων συμβατή με ένα διάγραμμα E-R μπορεί να αναπαρασταθεί με μια συλλογή πινάκων
- Η μετατροπή ενός διαγράμματος E-R σε συλλογή πινάκων αποτελεί την αφετηρία για να προκύψει μια σχεσιακή βάση δεδομένων από μια εννοιολογική σχεδίαση στο μοντέλο E-R
- Βασικοί κανόνες μετατροπής E-R σε σχεσιακό ...

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

28

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Αναπαράσταση συνόλων οντοτήτων



- Ένα ισχυρό σύνολο οντοτήτων μετατρέπεται σε πίνακα (με τα ίδια χαρακτηριστικά)

<i>customer-id</i>	<i>customer-name</i>	<i>customer-street</i>	<i>customer-city</i>
019-28-3746	Smith	North	Rye
182-73-6091	Turner	Putnam	Stamford
192-83-7465	Johnson	Alma	Palo Alto
244-66-8800	Curry	North	Rye
321-12-3123	Jones	Main	Harrison
335-57-7991	Adams	Spring	Pittsfield
336-66-9999	Lindsay	Park	Pittsfield
677-89-9011	Hayes	Main	Harrison
963-96-3963	Williams	Nassau	Princeton

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

29

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Αναπαράσταση αδύναμων συνόλων οντοτήτων



- Ένα αδύναμο σύνολο οντοτήτων γίνεται πίνακας που περιλαμβάνει μια στήλη για το πρωτεύον κλειδί του ισχυρού συνόλου οντοτήτων που το ταυτοποιεί

<i>loan-number</i>	<i>payment-number</i>	<i>payment-date</i>	<i>payment-amount</i>
L-11	53	7 June 2001	125
L-14	69	28 May 2001	500
L-15	22	23 May 2001	300
L-16	58	18 June 2001	135
L-17	5	10 May 2001	50
L-17	6	7 June 2001	50
L-17	7	17 June 2001	100
L-23	11	17 May 2001	75
L-93	103	3 June 2001	900
L-93	104	13 June 2001	200

ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

30

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Αναπαράσταση συνόλων συσχετίσεων



- Ένα σύνολο συσχετίσεων M:N αναπαρίσταται ως πίνακας με στήλες για τα πρωτεύοντα κλειδιά των οντοτήτων που συμμετέχουν, και επιπλέον όλα τα χαρακτηριστικά του συνόλου συσχετίσεων
- Π.χ.: πίνακας για το σύνολο συσχετίσεων borrower

<i>customer-id</i>	<i>loan-number</i>
019-28-3746	L-11
019-28-3746	L-23
244-66-8800	L-93
321-12-3123	L-17
335-57-7991	L-16
555-55-5555	L-14
677-89-9011	L-15
963-96-3963	L-17

Αναπαράσταση συνόλων συσχετίσεων (συν.)

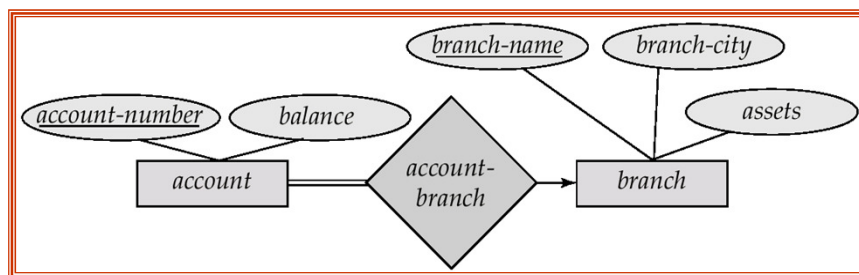


- Οι συσχετίσεις N:1 και 1:N μπορούν να αναπαρασταθούν απλά με προσθήκη ενός επιπλέον χαρακτηριστικού στην πλευρά 'N', με το πρωτεύον κλειδί της πλευράς '1'
 - Εάν η συμμετοχή στην πλευρά 'N' είναι μερική, μπορεί προκύψει η περίπτωση μια στήλη του πίνακα να έχει πολλές κενές τιμές. Οπότε ίσως θα συνέφερε η δημιουργία νέου πίνακα (όπως στην περίπτωση M:N)
- Ειδική περίπτωση: συσχέτιση αδύναμου με ισχυρό σύνολο οντοτήτων
 - Δεν απαιτείται σε καμία περίπτωση δημιουργία νέου πίνακα αφού ο πίνακας που προέρχεται από το αδύναμο σύνολο οντοτήτων έχει ήδη σαν στήλη το πρωτεύον κλειδί του ισχυρού συνόλου οντοτήτων.

Αναπαράσταση συνόλων συσχετίσεων (συν.)



- Παράδειγμα: αντί να δημιουργήσουμε έναν πίνακα για τη συσχέτιση *account-branch*, προσθέτουμε ένα χαρακτηριστικό *branch* στην οντότητα *account*



Αναπαράσταση συνόλων συσχετίσεων (συν.)



- Για τις συσχετίσεις 1:1 έχουμε 3 εναλλακτικές προσεγγίσεις:
 - είτε να προσθέσουμε το πρωτεύον κλειδί της μιας πλευράς ως επιπλέον χαρακτηριστικό στον πίνακα της άλλης πλευράς
 - και αυτό να γίνει για τον ένα ή και τους δύο πίνακες (στη δεύτερη περίπτωση, υπάρχει πλεονασμός πληροφορίας)
 - είτε να δημιουργηθεί νέος πίνακας (όπως στην περίπτωση M:N)

Αναπαράσταση σύνθετων / πλειότιμων χαρακτηριστικών



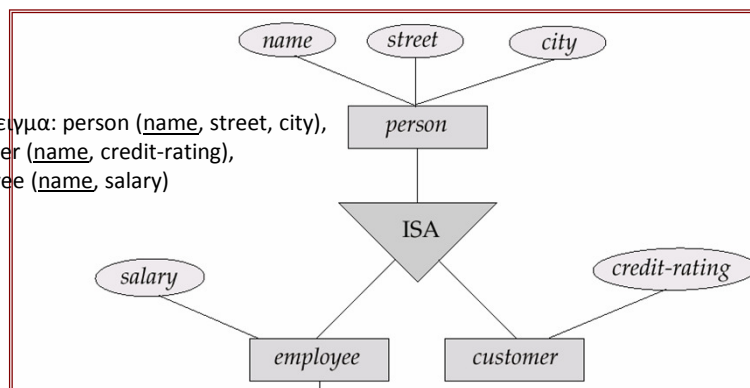
- Τα σύνθετα χαρακτηριστικά μετατρέπονται σε ένα σύνολο απλών
 - Παράδειγμα: έστω το σύνθετο χαρακτηριστικό name με συστατικά first-name και last-name. Ο πίνακας που προκύπτει θα έχει, μεταξύ άλλων, δύο χαρακτηριστικά name.first-name και name.last-name
- Από ένα πλειότιμο χαρακτηριστικό ενός συνόλου οντοτήτων προκύπτει νέος πίνακας !
 - Ο πίνακας έχει ως στήλες το πρωτεύον κλειδί του συνόλου οντοτήτων και μια ακόμη που αντιστοιχεί στο πλειότιμο χαρακτηριστικό
 - Παράδειγμα: έστω το πλειότιμο χαρακτηριστικό phone του συνόλου οντοτήτων customer (με πρωτεύον κλειδί customer-name). Προκύπτει ένας νέος πίνακας customer-phones (customer-name, phone)

Αναπαράσταση εξειδίκευσης



- 1^η μέθοδος: Προκύπτει ένας πίνακας για κάθε εμπλεκόμενο σύνολο οντοτήτων, όπου καθένας από τους πίνακες εξειδίκευσης συμπεριλαμβάνει ως στήλη το πρωτεύον κλειδί του πίνακα γενίκευσης

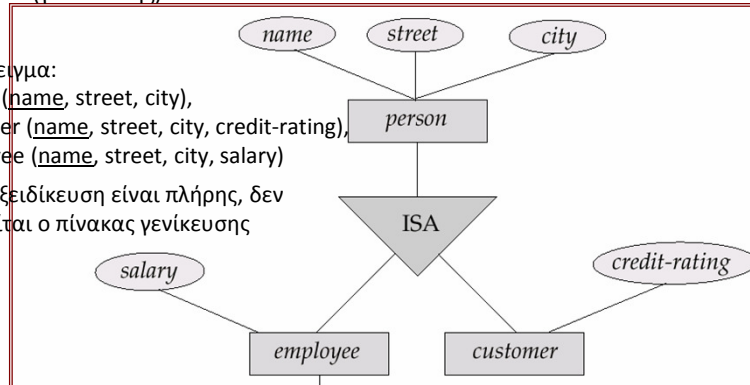
- Παράδειγμα: person (name, street, city),
customer (name, credit-rating),
employee (name, salary)



Αναπαράσταση εξειδίκευσης (συν.)



- 2^η μέθοδος: Προκύπτει ένας πίνακας για κάθε εμπλεκόμενο σύνολο οντοτήτων, όπου καθένας από τους πίνακες εξειδίκευσης συμπεριλαμβάνει ως στήλες όλα τα χαρακτηριστικά του συνόλου οντοτήτων ανώτερου επιπέδου (γενίκευσης)
- Παράδειγμα:
 person (name, street, city),
 customer (name, street, city, credit-rating),
 employee (name, street, city, salary)
- Εάν η εξειδίκευση είναι πλήρης, δεν απαιτείται ο πίνακας γενίκευσης



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

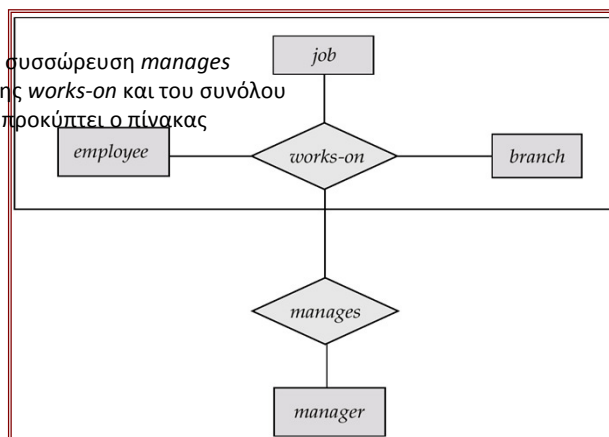
37

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Αναπαράσταση συσώρευσης



- Προκύπτει πίνακας με στήλες το πρωτεύον κλειδί της εμπλεκόμενης συσχέτισης, το πρωτεύον κλειδί του εμπλεκόμενου συνόλου οντοτήτων και τυχόν επιπλέον χαρακτηριστικά της συσώρευσης
- Παράδειγμα: από τη συσώρευση *manages* μεταξύ της συσχέτισης *works-on* και του συνόλου οντοτήτων *manager* προκύπτει ο πίνακας *manages* (employee-id, branch-name, title, manager-name)



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

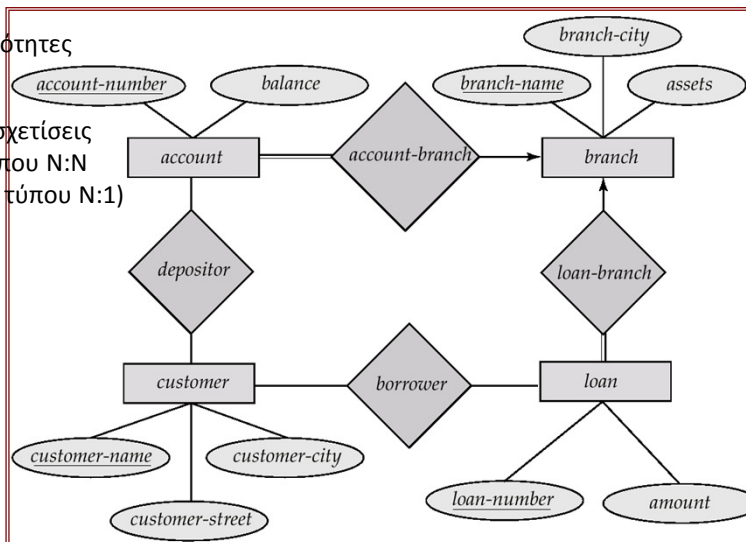
38

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα διαγράμματος E-R (ΒΔ τράπεζας)



- 4 οντότητες
- 4 συσχετίσεις
(2 τύπου N:N
και 2 τύπου N:1)



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

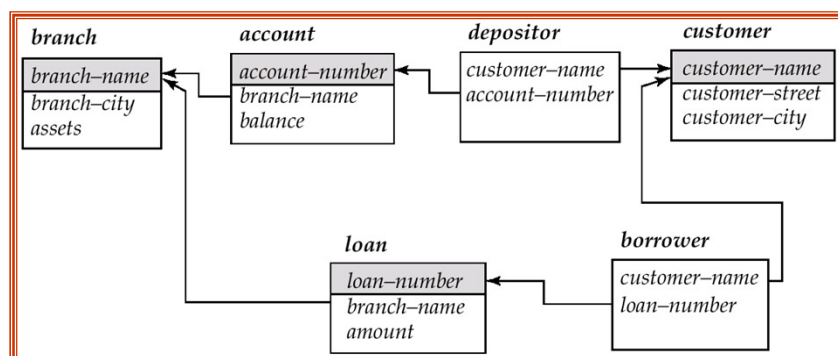
39

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα σχεσιακού σχήματος (ΒΔ τράπεζας)



- 6 (4+2) πίνακες ...



ΒΔ: [4] Σχεδίαση Σχεσιακών ΒΔ

40

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

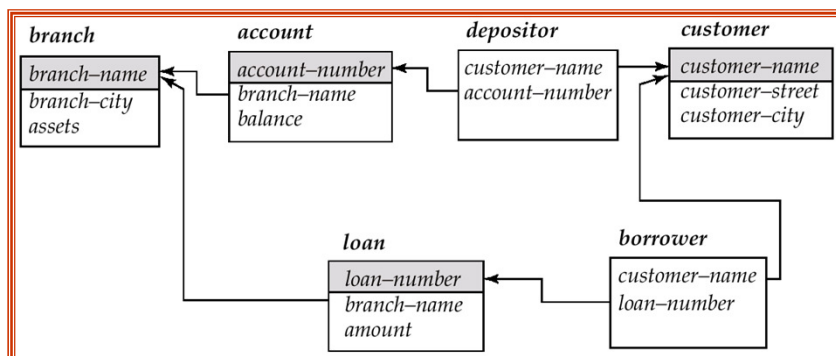


- Εννοιολογικός σχεδιασμός
 - το Μοντέλο Οντοτήτων-Συσχετίσεων (ER)
 - Από τα διαγράμματα ER στο Σχεσιακό Μοντέλο
- Θεωρία Κανονικοποίησης (Κανονικές Μορφές – ΚΜ)
 - 1^η ΚΜ(1NF)
 - Αποσύνθεση - Συναρτησιακές Εξαρτήσεις
 - 2^η (2NF), 3^η (3NF), Boyce-Codd ΚΜ (BCNF)
 - Καθολική Διαδικασία Σχεδίασης ΒΔ

Το πρόβλημα ...



- Γιατί αυτοί οι πίνακες και όχι κάποιοι άλλοι (περισσότεροι ή λιγότεροι, με διαφορετικό μοίρασμα χαρακτηριστικών);



Στόχοι στη σχεδίαση σχεσιακών ΒΔ

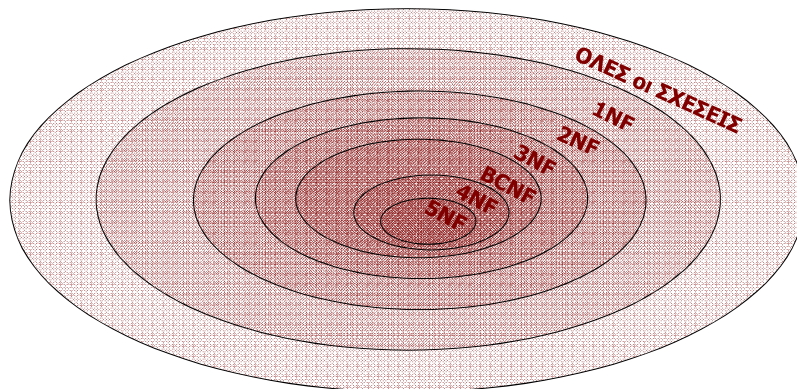


- Στόχος της σχεδίασης σχεσιακών ΒΔ είναι να βρούμε μια "καλή" συλλογή τύπων σχέσεων. Κακή σχεδίαση μπορεί να σημαίνει:
 - Επανάληψη / πλεονασμό πληροφορίας
 - Αδυναμία αναπαράστασης κάποιας πληροφορίας
- Σχεδιαστικοί στόχοι:
 - Αποφυγή πλεονασμού στα δεδομένα
 - Εξασφάλιση ότι αναπαριστώνται όλες οι συσχετίσεις μεταξύ των οντοτήτων
 - Διευκόλυνση των ελέγχων για τυχόν παραβίαση των περιορισμών ορθότητας
- Το «εργαλείο»: **κανονικοποίηση ΒΔ**
 - Υπακοή σε κάποια κανονική μορφή (1^η, 2^η, 3^η, ...)

Κανονικές μορφές



- Έχουν οριστεί 5 **κανονικές μορφές** (normal forms - NF) που η επόμενη είναι πιο αυστηρή από την προηγούμενη.
 - Μεταξύ 3^{ης} και 4^{ης} κανονικής μορφής, έχει οριστεί η κανονική μορφή Boyce-Codd (BCNF) που πρακτικά αποτελεί το 'στόχο' της διαδικασίας σχεδίασης ΒΔ



Πρώτη Κανονική Μορφή



- Οι τιμές του πεδίου ορισμού ενός χαρακτηριστικού είναι **ατομικές** εάν θεωρούνται ως αδιαίρετες μονάδες.
- Ένας τύπος σχέσης R βρίσκεται στην **1^η κανονική μορφή (first normal form – 1NF)** εάν τα πεδία τιμών όλων των χαρακτηριστικών του R είναι **ατομικά**
- Αντιπαραδείγματα:
 - Πλειότιμα χαρακτηριστικά (τηλέφωνα)
 - Σύνθετα χαρακτηριστικά (ονοματεπώνυμο)
 - Συμβολοσειρές (π.χ. κωδικοί) που σπάζουν σε τμήματα
 - Αν ο ΑΜ του φοιτητή είναι μια συμβολοσειρά της μορφής CS101 και οι δύο πρώτοι χαρακτήρες χρησιμοποιούνται για να βρούμε το Τμήμα του φοιτητή, τότε καταστρατηγούμε την ατομικότητα

Παράδειγμα «προβληματικού» σχήματος



- Έστω το σχήμα:
Lending-schema = (branch-name, branch-city, assets, customer-name, loan-number, amount)

<i>branch-name</i>	<i>branch-city</i>	<i>assets</i>	<i>customer-name</i>	<i>loan-number</i>	<i>amount</i>
Downtown	Brooklyn	9000000	Jones	L-17	1000
Redwood	Palo Alto	2100000	Smith	L-23	2000
Perryridge	Horseneck	1700000	Hayes	L-15	1500
Downtown	Brooklyn	9000000	Jackson	L-14	1500

- Πλεονασμός:
 - Επανάληψη των δεδομένων *branch-name, branch-city, assets* για κάθε δάνειο που δίνει το συγκεκριμένο υποκατάστημα → σπατάλη χώρου
 - Δυσκολία στην τροποποίηση δεδομένων, πιθανότητα ασυνεπών τιμών
- Τιμές Null
 - Δεν μπορούμε να αποθηκεύσουμε τα στοιχεία ενός υποκαταστήματος αν δεν έχει δοθεί κανένα δάνειο από αυτό
 - Μπορούμε να χρησιμοποιήσουμε τιμές null αλλά είναι δύσκολος ο χειρισμός τους.

Αποσύνθεση (decomposition)



- Αποσύνθεση του σχήματος

<i>branch-name</i>	<i>branch-city</i>	<i>assets</i>	<i>customer-name</i>	<i>loan-number</i>	<i>amount</i>
Downtown	Brooklyn	9000000	Jones	L-17	1000
Redwood	Palo Alto	2100000	Smith	L-23	2000
Perryridge	Horseneck	1700000	Hayes	L-15	1500
Downtown	Brooklyn	9000000	Jackson	L-14	1500

σε

Branch-schema = (*branch-name*, *branch-city*, *assets*)

Loan-info-schema = (*customer-name*, *loan-number*, *branch-name*, *amount*)

- Όλα τα χαρακτηριστικά του αρχικού τύπου σχέσης (*R*) πρέπει να διατηρούνται στην αποσύνθεση (*R*₁, *R*₂): $R = R_1 \cup R_2$
- Στόχος της σχεδίασης σχεσιακών βάσεων δεδομένων: να αποφασίσουμε αν ένας τύπος σχέσης *R* είναι σε "καλή" μορφή.
- Το ζητούμενο είναι: **αποσύνθεση χωρίς απώλεια πληροφορίας (lossless decomposition)**

- Φορμαλιστικά:

$$\forall \text{ σχέση } r \text{ τύπου } R, r = \Pi_{R_1}(r) \bowtie \Pi_{R_2}(r)$$

Συναρτησιακές Εξαρτήσεις



- Περιορισμοί πάνω στο σύνολο των έγκυρων σχέσεων
- Απαιτούν η τιμή ορισμένων χαρακτηριστικών να προσδιορίζει μοναδικά την τιμή άλλων χαρακτηριστικών
- Γενίκευση της έννοιας του κλειδιού
- Ορισμός:** Έστω *R* ένας τύπος σχέσης, με $\alpha \subseteq R$ and $\beta \subseteq R$. Η **συναρτησιακή εξάρτηση (functional dependence - FD)** $\alpha \rightarrow \beta$ ισχύει πάνω στον *R* εάν και μόνο εάν

$$t_1[\alpha] = t_2[\alpha] \Rightarrow t_1[\beta] = t_2[\beta]$$

- Παραδείγματα:

branch-name \rightarrow *assets*, *branch-city*

loan-number \rightarrow *amount*, *branch-name*

Συναρτησιακές Εξαρτήσεις (συν.)



- Το σύνολο χαρακτηριστικών K αποτελεί **υπερκλειδί (superkey)** για τον τύπο σχέσης R εάν και μόνο εάν $K \rightarrow R$
- Το K αποτελεί **υποψήφιο κλειδί (candidate key)** για τον R εάν και μόνο εάν
 - $K \rightarrow R$, και
 - δεν υπάρχει $\alpha \subset K$, $\alpha \rightarrow R$
- Οι συναρτησιακές εξαρτήσεις επιτρέπουν να εκφράσουμε περιορισμούς που δεν μπορούν να εκφραστούν μέσω των υπερκλειδιών.
 - Για παράδειγμα, ας θεωρήσουμε τον τύπο σχέσης:

$$\text{loan-info-schema} = (\text{customer-name}, \text{loan-number}, \text{branch-name}, \text{amount})$$
 περιμένουμε να ισχύουν αυτές οι συναρτησιακές εξαρτήσεις:

$$\text{loan-number} \rightarrow \text{amount} \text{ και } \text{loan-number} \rightarrow \text{branch-name}$$
 αλλά όχι αυτή: $\text{loan-number} \rightarrow \text{customer-name}$

Συναρτησιακές Εξαρτήσεις (συν.)



- Χρησιμοποιούμε συναρτησιακές εξαρτήσεις για να:
 - Ελέγχουμε την εγκυρότητα μιας σχέσης r σύμφωνα με ένα δοθέν σύνολο συναρτησιακών εξαρτήσεων F (οπότε λέμε ότι η r **ικανοποιεί** την F).
 - Προδιαγράψουμε περιορισμούς F πάνω σε έναν τύπο σχέσης R (οπότε λέμε ότι η F **ισχύει για** τον R)
- Μια συναρτησιακή εξάρτηση καλείται **τετριμμένη** εάν ικανοποιείται οπωσδήποτε σε κάθε στιγμιότυπο ενός τύπου σχέσης. π.χ.
 - $\text{customer-name}, \text{loan-number} \rightarrow \text{customer-name}$
 - Γενικά, $\alpha \rightarrow \beta$ είναι τετριμμένη εάν $\beta \subseteq \alpha$

Κανονικοποίηση με χρήση συναρτησιακών εξαρτήσεων



- Όταν αποσυνθέτουμε ένα τύπο σχέσης R με ένα σύνολο συναρτησιακών εξαρτήσεων F στους τύπους σχέσεων R_1, R_2, \dots, R_n έχουμε τριπλό στόχο:
 - Στόχος 1: **όχι απώλεια πληροφορίας**
 - Η Αποσύνθεση πρέπει να είναι χωρίς απώλεια πληροφορίας (lossless decomposition)
 - Στόχος 2: **όχι πλεονασμός πληροφορίας**
 - Οι τύποι σχέσεων R_i που προκύπτουν από την αποσύνθεση, κατά προτίμηση να είναι σε Boyce-Codd Κανονική Μορφή ή 3^η Κανονική Μορφή (θα τις ορίσουμε σε λίγο ...)
 - Στόχος 3: **διατήρηση εξαρτήσεων**
 - Έστω F_i το σύνολο των εξαρτήσεων που περιέχουν μόνο χαρακτηριστικά του R_i . Η αποσύνθεση επιθυμούμε να διατηρεί τις εξαρτήσεις.

Δεύτερη Κανονική Μορφή



- Ένας τύπος σχέσης R είναι σε **2^η Κανονική Μορφή (2nd Normal Form - 2NF)** εάν είναι σε 1NF και επιπλέον:
 - Κανένα χαρακτηριστικό (από αυτά που δεν συμμετέχουν σε υποψήφιο κλειδί) δεν **προσδιορίζεται** από **τμήμα** ενός υποψήφιου κλειδιού K
- Φορμαλιστικά: $\nexists (a \rightarrow b) \text{ in FD: } a \subset K \wedge b \notin K$
- Παράδειγμα:
 - $R = (\text{customer-name}, \text{loan-number}, \text{amount})$
 - Καταρχήν η R είναι σε 1NF
 - Όμως, ισχύει η εξάρτηση $\text{loan-number} \rightarrow \text{amount}$ και το χαρακτηριστικό loan-number είναι μέρος του υποψήφιου κλειδιού $\{\text{customer-name}, \text{loan-number}\}$.
 - Άρα η R δεν είναι σε 2NF

Δεύτερη Κανονική Μορφή (συν.)



- Πώς γίνεται αποσύνθεση σε 2NF:
 - Απαλλάσσουμε την R από την 'προβληματική' εξάρτηση, η οποία με τη σειρά της σχηματίζει ένα νέο τύπο σχέσης.
- Παράδειγμα:
 - $R = (\underline{customer-name}, \underline{loan-number}, amount)$
 - Η 'προβληματική' εξάρτηση είναι η $loan-number \rightarrow amount$
 - Άρα: $R_1 = (\underline{customer-name}, \underline{loan-number})$, $R_2 = (\underline{loan-number}, amount)$
 - Τόσο η R_1 όσο και η R_2 είναι σε 2NF

Τρίτη Κανονική Μορφή



- Ένας τύπος σχέσης R είναι σε **3^η Κανονική Μορφή (3rd Normal Form - 3NF)** εάν είναι σε 2NF και επιπλέον:
 - Κανένα χαρακτηριστικό (από αυτά που δεν συμμετέχουν σε υποψήφιο κλειδί) δεν **προσδιορίζεται μεταβατικά** (δηλ. μέσω άλλου χαρακτηριστικού) από ένα υποψήφιο κλειδί K
 - Η μεταβατική εξάρτηση ισχύει όταν το χαρακτηριστικό που παρεμβάλλεται **δεν** συμμετέχει σε υποψήφιο κλειδί
- Φορμαλιστικά:

$\nexists (a \rightarrow b), (b \rightarrow c) \text{ in FD: } a = K \wedge b \notin K \wedge c \notin K$
- Παράδειγμα:
 - $R = (\underline{loan-number}, amount, branch-name, branch-city)$
 - Καταρχήν η R είναι σε 2NF. Όμως:
 - $loan-number \rightarrow branch-name$ και $branch-name \rightarrow branch-city$, οπότε το $branch-city$ προσδιορίζεται μεταβατικά από το $loan-number$ μέσω του $branch-name$.
 - Άρα η R δεν είναι σε 3NF

Τρίτη Κανονική Μορφή (συν.)



- Πώς γίνεται αποσύνθεση σε 3NF:
 - Απαλλάσσουμε την R από την 'προβληματική' εξάρτηση, η οποία με τη σειρά της σχηματίζει ένα νέο τύπο σχέσης.
- Παράδειγμα:
 - $R = (\underline{\text{loan-number}}, \text{amount}, \text{branch-name}, \text{branch-city})$
 - Η 'προβληματική' εξάρτηση είναι η $\text{branch-name} \rightarrow \text{branch-city}$
 - Άρα: $R_1 = (\underline{\text{loan-number}}, \text{amount}, \text{branch-name})$, $R_2 = (\underline{\text{branch-name}}, \text{branch-city})$
 - Τόσο η R_1 όσο και η R_2 είναι σε 3NF
- Αποδεικνύεται ότι πάντοτε μπορεί να βρεθεί μια αποσύνθεση σε 3NF που να είναι αποσύνθεση χωρίς απώλεια πληροφορίας και να διατηρεί τις εξαρτήσεις.

Boyce-Codd Κανονική Μορφή



- Ένας τύπος σχέσης R είναι σε **Boyce-Codd Κανονική Μορφή (Boyce-Codd Normal Form - BCNF)** εάν για όλες τις συναρτησιακές εξαρτήσεις $\alpha \rightarrow \beta$ που ισχύουν στην R , το α αποτελεί υποψήφιο κλειδί της R
 - Προσοχή: δεν ξεκινάμε τον έλεγχο με την παραδοχή ότι η R είναι σε 3NF
 - Απλά, μπορεί να αποδειχθεί ότι αν η R είναι σε BCNF τότε είναι και σε 3NF.
- Φορμαλιστικά: $\nexists (a \rightarrow b) \text{ in FD: } a \neq K$
- Παράδειγμα:
 - Έστω $R = (A, B, C)$, ισχύουν οι συναρτησιακές εξαρτήσεις $F = \{A \rightarrow B, B \rightarrow C\}$ και $\{A\}$ είναι το κλειδί
 - \Rightarrow Η R δεν είναι σε BCNF (μάλιστα, δεν είναι ούτε σε 3NF !!)
 - Έστω ότι γίνεται η αποσύνθεση της R σε $R_1 = (A, B)$, $R_2 = (B, C)$
 - \Rightarrow οι R_1 και R_2 είναι σε BCNF

Boyce-Codd Κανονική Μορφή (συν.)



- Πώς γίνεται αποσύνθεση σε 3NF:
 - Κάθε συναρτησιακή εξάρτηση σχηματίζει ένα νέο τύπο σχέσης.
- Παράδειγμα:
 - $R = (branch-name, branch-city, assets, \underline{customer-name}, \underline{loan-number}, amount)$

<i>branch-name</i>	<i>branch-city</i>	<i>assets</i>	<i>customer-name</i>	<i>loan-number</i>	<i>amount</i>
Downtown	Brooklyn	9000000	Jones	L-17	1000
Redwood	Palo Alto	2100000	Smith	L-23	2000
Perryridge	Horseneck	1700000	Hayes	L-15	1500
Downtown	Brooklyn	9000000	Jackson	L-14	1500

- συναρτησιακές εξαρτήσεις:
 - $branch-name \rightarrow assets, branch-city$
 - $loan-number \rightarrow amount, branch-name$
- Άρα: $R_1 = (branch-name, branch-city, assets)$, $R_2 = (branch-name, \underline{loan-number}, amount)$, $R_3 = (\underline{customer-name}, \underline{loan-number})$

BCNF και Διατήρηση Εξαρτήσεων



- Ιδιότητες της BCNF:
 - Αποσύνθεση χωρίς απώλεια πληροφορίας \rightarrow ΝΑΙ
 - Διατήρηση εξαρτήσεων \rightarrow ΙΣΩΣ (!)
 - Με άλλα λόγια, δεν είναι πάντοτε δυνατό να πάρουμε μια αποσύνθεση BCNF που να διατηρεί τις εξαρτήσεις !!
- Παράδειγμα:
 - $R = (J, K, L)$, $F = \{JK \rightarrow L, L \rightarrow K\}$, δύο υποψήφια κλειδιά JK και JL
 $\Rightarrow R$ είναι σε 3NF αλλά δεν είναι σε BCNF
 - Οποιαδήποτε αποσύνθεση της R αδυνατεί να διατηρήσει την εξάρτηση $JK \rightarrow L$
- Τότε καταφεύγουμε στη λύση της 3NF (μπορεί να αποδειχθεί ότι αν ένας τύπος σχέσης R είναι σε BCNF τότε είναι και σε 3NF)
 - Αυτή η λύση επιτρέπει κάποιας μορφής πλεονασμό (με συνεπαγόμενα προβλήματα όμως)

BCNF και Διατήρηση Εξαρτήσεων (συν.)



- Προβλήματα που μπορεί να προκύψουν λόγω πλεονασμού στην 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

- Επανάληψη πληροφορίας (π.χ., η συσχέτιση l_1, k_1)
- Ανάγκη για χρήση τιμών null (π.χ., για να αναπαρασταθεί η συσχέτιση l_2, k_2 όταν δεν υπάρχει αντίστοιχη τιμή για το J).

Τέταρτη (4NF) και Πέμπτη (5NF) Κανονική Μορφή



- Με τη χρήση των συναρτησιακών εξαρτήσεων (functional dependencies - FD) μπορούμε να ορίσουμε 2NF, 3NF και BCNF
- Έχουν μελετηθεί και άλλες εξαρτήσεις:
 - πλειότιμες εξαρτήσεις (Multi-Valued Dependencies - MVD),
 - εξαρτήσεις συνένωσης (Join Dependencies),
 - εξαρτήσεις εγκλεισμού (Inclusion Dependencies), ...
- Αυτές οι εξαρτήσεις μας πάνε σε περαιτέρω κανονικές μορφές, 4NF και 5NF
 - Έχουν μελετηθεί θεωρητικά αλλά ...
 - ... δεν είναι στους σκοπούς του μαθήματος !!

Καθολική διαδικασία σχεδίασης ΒΔ



- Οι στόχοι στη σχεδίαση σχεσιακών ΒΔ είναι:
 - BCNF
 - Σύνδεση χωρίς απώλεια πληροφορίας
 - Διατήρηση εξαρτήσεων
- Αν δεν μπορούμε να τα πετύχουμε, κάνουμε υποχωρήσεις:
 - Μη διατήρηση κάποιων εξαρτήσεων
 - 3NF αντί BCNF (πλεονασμός ...)
- Ας σημειωθεί ότι η SQL δεν παρέχει κάποιον άμεσο τρόπο να προδιαγραφούν οι συναρτησιακές εξαρτήσεις, πέρα από τους περιορισμούς του κλειδιού.
 - Μπορούν να προδιαγραφούν με assertions, αλλά ο έλεγχός τους κοστίζει

Καθολική διαδικασία σχεδίασης ΒΔ (συν.)



- Κάνουμε την παραδοχή ότι ο τύπος σχέσης R :
 - Θα μπορούσε να προκύψει όταν μετατρέπουμε ένα διάγραμμα E-R σε ένα σύνολο πινάκων.
 - Θα μπορούσε να θεωρηθεί ως μία και μόνη σχέση που περιέχει **όλα** τα χαρακτηριστικά που μας ενδιαφέρουν στη ΒΔ (ονομάζεται **καθολική σχέση - universal relation**).
 - Θα μπορούσε να είναι το αποτέλεσμα κάποιων σχεδιαστικών αποφάσεων στην πορεία ελέγχου ή μετατροπής μιας σχέσης σε κάποια κανονική μορφή

Η προσέγγιση της καθολικής σχέσης



- Η σχέση $r_1 \bowtie r_2 \bowtie \dots \bowtie r_n$ ονομάζεται **καθολική σχέση (universal relation)** καθώς εμπλέκει όλα τα χαρακτηριστικά της ΒΔ όπως ορίζονται από το $R_1 \cup R_2 \cup \dots \cup R_n$
- «άκυρες» πλειάδες** – πλειάδες που “εξαφανίζονται” όταν υπολογίζουμε μια σύνδεση.
 - Έστω $r_1(R_1), r_2(R_2), \dots, r_n(R_n)$ ένα σύνολο από σχέσεις
 - Μια πλειάδα r της σχέσης r_i είναι «άκυρη» πλειάδα εάν η r δεν βρίσκεται μέσα στη σχέση:

$$\Pi_{R_i}(r_1 \bowtie r_2 \bowtie \dots \bowtie r_n)$$

Παράδειγμα αποσύνθεσης καθολικής σχέσης



- Έστω σε εννοιολογικό επίπεδο έχουν καταγραφεί τα εξής:
 - Οντότητες: customer (customer-name), loan (loan-number, amount), branch (branch-name, branch-city, assets)
 - Συσχετίσεις: customer-loan (τύπου M:N), loan-branch (τύπου N:1)
- Το σχήμα καθολικής σχέσης που προκύπτει είναι:
 - $U = (\text{branch-name, branch-city, assets, customer-name, loan-number, amount})$
- Ένα στιγμότυπο της σχέσης:

branch-name	branch-city	assets	customer-name	loan-number	amount
Downtown	Brooklyn	9000000	Jones	L-17	1000
Redwood	Palo Alto	2100000	Smith	L-23	2000
Perryridge	Horseneck	1700000	Hayes	L-15	1500
Downtown	Brooklyn	9000000	Jackson	L-14	1500

Έλεγχος για 2NF → ??

για 3NF → ??

για BCNF → ??

Μοντέλο E-R και Κανονικοποίηση



- Όταν το διάγραμμα E-R σχεδιάζεται προσεκτικά, με επιτυχή εντοπισμό των οντοτήτων, οι πίνακες που προκύπτουν δεν χρειάζονται περαιτέρω κανονικοποίηση
- Όμως, σε πραγματικές καταστάσεις, μπορούν να υπάρχουν συναρτησιακές εξαρτήσεις από χαρακτηριστικά που δεν είναι κλειδιά προς άλλα χαρακτηριστικά μιας οντότητας
 - π.χ. οντότητα *employee* με χαρακτηριστικά *department-number* και *department-address* και με τη συναρτησιακή εξάρτηση *department-number* → *department-address*
 - Ένας καλός σχεδιασμός θα είχε αναγνωρίσει το *department* ως οντότητα

Απο-κανονικοποίηση (denormalization) για λόγους απόδοσης



- Μερικές φορές για λόγους απόδοσης προτιμούμε μη κανονικοποιημένα σχήματα
 - π.χ. αν θέλουμε τα *customer-name*, *account-number* και *balance* να είναι μαζί στο αποτέλεσμα μιας ερώτησης απαιτείται *account* ⋈ *depositor*
- 1^ο εναλλακτικό: απο-κανονικοποίηση με τα χαρακτηριστικά των σχέσεων να είναι μαζί σε μία (μη κανονικοποιημένη) σχέση
 - + μεγαλύτερη ταχύτητα στις ερωτήσεις
 - μεγαλύτερες απαιτήσεις χώρου και χρόνου εκτέλεσης των αλλαγών στη ΒΔ, κίνδυνος σφαλμάτων
- 2^ο εναλλακτικό : ορισμός μιας υλοποιημένης θεώρησης (materialized view) *account* ⋈ *depositor*
 - τα ίδια υπέρ και κατά όπως πριν, εκτός του κινδύνου σφαλμάτων

Άλλα σχεδιαστικά ζητήματα



- Υπάρχουν προβλήματα που δεν οφείλονται στη (μη) κανονικοποίηση αλλά σε κακή σύλληψη. Παράδειγμα: αντί του πίνακα *earnings(company-id, year, amount)*, να έχουμε
 - 1^η λύση: πολλούς πίνακες: *earnings-2000, earnings-2001, earnings-2002*, κλπ., που όλοι θα είναι του τύπου (*company-id, earnings*).
 - Είναι μεν σε BCNF, αλλά καθιστούν δύσκολη την εκτέλεση ερωτήσεων όταν εμπλέκονται πολλά έτη και επιπλέον απαιτείται ένας νέος πίνακας κάθε έτος!
 - 2^η λύση: *company-year(company-id, earnings-2000, earnings-2001, earnings-2002)*
 - Επίσης σε BCNF, αλλά καθιστούν δύσκολη την εκτέλεση ερωτήσεων όταν εμπλέκονται πολλά έτη και επιπλέον απαιτείται ένα νέο χαρακτηριστικό κάθε έτος!
 - Αποτελεί κλασικό παράδειγμα λειτουργίας **crosstab**, όπου οι τιμές ενός χαρακτηριστικού γίνονται στήλες
 - Χρησιμοποιείται σε spreadsheets και σε εργαλεία ανάλυσης δεδομένων (βλ. μάθημα επόμενου εξαμήνου)



ΤΕΛΟΣ ΕΝΟΤΗΤΑΣ 4