



Εισαγωγή

Βάσεις Δεδομένων



Αντικείμενο: Θεμελιώδες πρόβλημα της επιστήμης μας
Διαχείριση Δεδομένων

- Μοντελοποίηση
- Αποθήκευση
- Επεξεργασία (εύρεση πληροφορίας σχετικής με μια συγκεκριμένη ερώτηση)
- Σωστή Λειτουργία (αποτυχίες συστήματος, συνέπεια)

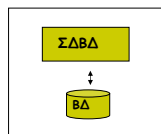
Βασικές Έννοιες



Τι είναι μια βάση δεδομένων;

Βάση Δεδομένων: συλλογή από σχετιζόμενα δεδομένα

Σύστημα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ): λογισμικό (σύνολο από προγράμματα) για δημιουργία και χρήση μιας βάσης δεδομένων



Σύστημα Βάσεων Δεδομένων

Βασικές Έννοιες



Κάποιες λειτουργίες ενός ΣΔΒΔ

- **Ορισμός** μιας βάσης δεδομένων: προδιαγραφή των τύπων, των δομών και των περιορισμών των δεδομένων που θα αποθηκευτούν στη ΒΔ
- **Κατασκευή** μια βάσης δεδομένων: αποθήκευση των ίδιων των δεδομένων
- **Χειρισμός** (manipulation) μιας βάσης δεδομένων: υποβολή ερωτήσεων για την ανάκτηση δεδομένων, ενημέρωση
- **Άλλες λειτουργίες:** Διαμοιρασμός, προστασία από αστοχίες υλικού και λογισμικού, ασφάλεια

Βασικές Έννοιες



Γιατί ένα ΣΔΒΔ:

- Κοινή λειτουργικότητα ήδη υλοποιημένη
- Σωστή υλοποίηση
- Ανεξαρτησία δεδομένων (θα δούμε περισσότερα σε λίγο)
- Ενδιαφέρουσες ιδέες/μοντελοποιήσεις
- Μόνιμη αποθήκευση, έλεγχος της επανάλψης πληροφορίας, ορθότητα, έλεγχος συνδρομικότητας, έλεγχος προσπέλασης, ασφάλεια, και άλλα πολλά

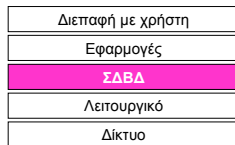
Βασικές Έννοιες



Γιατί όχι:

- Επένδυση σε λογισμικό και υλικό, καθώς και για εκπαίδευση
- Η γενικότητα που παρέχει δε χρειάζεται - προκαλεί χρονική επιβάρυνση (overhead)
- Δε σας χρειάζονται όσα προσφέρει

Η θέση των ΣΔΒΔ στη στοιβα του λογισμικού συστημάτων



Σκοπός του μαθήματος

- Σχεδιασμός και Προγραμματισμός μια βάσης δεδομένων χρησιμοποιώντας ένα ΣΔΒΔ
- Κάποια θέματα υλοποίησης ενός ΣΔΒΔ (το εσωτερικό του)
- Γενικές τεχνικές/αρχές/αλγόριθμους που διέπουν τη διαχείρισης δεδομένων

Μερικά «διαχειριστικά» θέματα

- web σελίδα <http://www.cs.uoi.gr/~pitoura>
 - Βιβλίο
«Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων», 4η Έκδοση, Τόμος Α
έρχεται ...
- Βαθμός
- Ασκήσεις (3 - 4 σύνολα) + Μια προγραμματιστική άσκηση (ε 4.0)
 - Τελικό διαγώνισμα (ε 5.0)
 - Τελικός Βαθμός (ε 5.0)
 - 40% * (Βαθμός Ασκήσεων+Προγραμματιστικής) +
 - 60% * Βαθμός Τελικού Διαγωνίσματος

Συμβουλές προς ναυτιλλόμενους

- Ναι, πρέπει να **μελετήσετε**
- Καλό θα είναι να *παρακολουθείτε το μάθημα* (τις διαλέξεις, αλλά και το ρυθμό του)
- Η ύλη/σειρά στο βιβλίο μπορεί να διαφέρει από το μάθημα -αλλά ό,τι πούμε στο μάθημα και ό,τι υπάρχει στα σχετικά κεφάλαια του βιβλίου αρκεί για να «περάσετε» το μάθημα
- Και όμως, ναι θα πρέπει να σκεφτείτε και να λύσετε προβλήματα «από το μυαλό σας»

Τι θα δούμε σήμερα

- Ιστορική Αναδρομή ...
- Η ύλη του μαθήματος σε λιγότερο από 30'

Ιστορία

Δεκαετία του 1950

Κάρτες και ταινίες (σειριακή επεξεργασία) - Batch processing

Αρχή του 1960

πρώτο γενικού-σκοπού ΣΔΒΔ: Integrated Data Store (GE)

Charles Bachman (Recipient of the 1st Turing Award, 1973)

network data model (δίκτυωτό)

Τέλη του 1960

Information Management System (IMS) IBM

hierarchical data model (Ιεραρχικό)

SABRE Airline Reservation System (AA+IBM, travelocity!!)



1970

Edgar Codd (IBM, San Jose) σχεσιακό μοντέλο δεδομένων (relational data model)

(Recipient of the **Turing Award**, 1981)

Ερευνητικά Προγράμματα: System R, INGRES - Γλώσσες: SEQUEL, QBE, QUEL

Δεκαετία του 1980

SQL (μέρος του System R)

transaction management (Jim Gray, **Turing Award**, 1999)

[*Τάσεις: αντικειμενοστραφή, αρχιτεκτονική πελάτη-εξυπηρέτη, κατανεμημένες, έμπειρα*]



Δεκαετία του 1990

εμπορικά αντικειμενοστραφή συστήματα

[*Τάσεις: πολυβάσεις, χωρικές & χρονικές, πολυμέσα, συμπερασματικές, αποθήκες δεδομένων (αναλυτική επεξεργασία), προγραμματισμό πόρων της επιχείρησης (ERP - Enterprise Resource Planning) και της διαχείρισης τους (MRP - Management Resource Planning), Internet*]



Δεκαετία του 2000

Σύστημα Διαχείρισης Χρωμοσωμάτων (Human Genome Project)

Σύστημα Παρατήρησης της Γης (Earth Observation System)

[*Τάσεις: Asilomar Report 1998*]

http://www.research.microsoft.com/~gray/Asilomar_DB_98.html

- *The Web Changes Everything*
- *Unifying Program Logic and Database Systems*
- *Hardware Advances: Scale up to MegaServers and Scale Down to Appliances*



Δεκαετία του 2000

[*Τάσεις: The Lowell Database Research Self-Assessment Meeting, May 2003*]

<http://research.microsoft.com/~Gray/Lowell/>

Rethink basic DBMS architecture with an eye toward supporting:

- Structured data
- Text, space, time, image, and multimedia data
- Procedural data, that is data types and the methods that encapsulate them
- Triggers
- Data streams and queues

as co-equal first-class components within the DBMS architecture



Δεκαετία του 2000

[*Τάσεις: The Lowell Database Research Self-Assessment Meeting, May 2003*]

<http://research.microsoft.com/~Gray/Lowell/>

Μερικά θέματα

- Sensor Data and Sensor Networks
- Multimedia Queries
- Reasoning about Uncertain Data
- Personalization
- Data Mining
- Privacy
- και άλλα (Ρίξτε μια ματιά στη σελίδα για μια ιδέα)



Σήμερα (τέλος του 2006) (hot topics!)

XML

web services

peer-to-peer (p2p)/ grids

streams

IR + ΒΔ

- Ιστορική Αναδρομή
- **Η ύλη του μαθήματος σε λιγότερο από 30'**

ΜΕΡΟΣ 1

Μοντελοποίηση - Ορισμός **Με χρήση ΣΔΒΔ**
Προγραμματισμός
 Δημιουργία/Κατασκευή
 Εισαγωγή Δεδομένων
 Επεξεργασία Δεδομένων

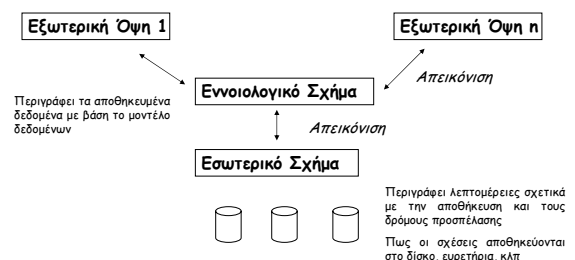
ΜΕΡΟΣ 2

Υλοποίηση ΣΔΒΔ **Το εσωτερικό ενός ΣΔΒΔ**
 (το μάθημα σε λιγότερο από 30')

Μοντέλο Δεδομένων: ένα σύνολο από έννοιες που μπορούν να χρησιμοποιηθούν για την περιγραφή της δομής της βδ

- **Υψηλού επιπέδου (εννοιολογικά) μοντέλα**
 Μοντέλο Οντοτήτων/Συσχετίσεων
- **Παραστατικά μοντέλα ή μοντέλα υλοποίησης**
 Σχισιακό Μοντέλο, Ίεραρχικό Μοντέλο, Δικτυωτό Μοντέλο
- **Χαμηλού επιπέδου ή φυσικά μοντέλα**
 Δρόμος Προσπέλασης

(το μάθημα σε λιγότερο από 30')



(το μάθημα σε λιγότερο από 30')

Ανεξαρτησία Δεδομένων: αλλαγή του σχήματος ενός επιπέδου χωρίς να αλλάξουμε το σχήμα του αμέσως υψηλότερου επιπέδου

- **Λογική Ανεξαρτησία Δεδομένων**
 αλλαγή του εννοιολογικού δεν επηρεάζει τα εξωτερικά σχήματα ή τα προγράμματα εφαρμογών
- **Φυσική Ανεξαρτησία Δεδομένων**
 αλλαγή του εσωτερικού σχήματος χωρίς να χρειάζεται αλλαγή του εννοιολογικού

αλλαγή μόνο της απεικόνισης

(το μάθημα σε λιγότερο από 30')

Σύστημα Βάσεων Δεδομένων για Κινηματογραφικές Ταινίες

ΒΗΜΑ 1: Μοντελοποίηση

- Εννοιολογικό Μοντέλο (Μοντέλο Οντοτήτων/Συσχετίσεων)
- Μοντέλο Υλοποίησης (Σχισιακό μοντέλο)

ΒΗΜΑ 2: Προγραμματισμός/Υλοποίηση

- Ορισμός Σχέσεων (πρόθεση/σχήμα)
- Εισαγωγή Στοιχείων (δημιουργία του αρχικού στιγμιότυπου)
- Διατύπωση Ερωτήσεων

(το μάθημα σε λιγότερο από 30' - ΜΕΡΟΣ 1)

Σχήματα και Στιγμιότυπα

Σχήμα της Βάσης

(δομικό στοιχείο, περιορισμοί, κατάλογος του συστήματος)

Πρόθεση (intension)

Ανάπτυξη (extension)

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

(αρχική κατάσταση, έγκυρη κατάσταση)

(το μάθημα σε λιγότερο από 30')

Βάσεις Δεδομένων 2006-2007 Ευαγγελία Πιτουρά 25

Παράδειγμα

Σύστημα Βάσεων Δεδομένων για Κινηματογραφικές Ταινίες

ΒΗΜΑ 1: Μοντελοποίηση

- Εννοιολογικό Μοντέλο (μοντέλο Οντοτήτων/Συσχετίσεων)
- Μοντέλο Υλοποίησης (σχεσιακό μοντέλο)

(το μάθημα σε λιγότερο από 30' - ΜΕΡΟΣ 1)

Βάσεις Δεδομένων 2006-2007 Ευαγγελία Πιτουρά 26

Παράδειγμα

ΒΗΜΑ 2: Προγραμματισμός/Υλοποίηση - σε σχεσιακό ΣΔΒΔ

- Ορισμός Σχέσεων (πρόθεση/σχήμα)
- Εισαγωγή Στοιχείων (δημιουργία του αρχικού στιγμιότυπου)
- Διατύπωση Ερωτήσεων

(το μάθημα σε λιγότερο από 30' - ΜΕΡΟΣ 1)

Βάσεις Δεδομένων 2006-2007 Ευαγγελία Πιτουρά 27

Γλώσσες ΣΔΒΔ

Γλώσσα Ορισμού

Γλώσσα Ορισμού Δεδομένων
Γλώσσα Αποθήκευσης Δεδομένων
Γλώσσα Ορισμού Όψεων

Γλώσσα Χειρισμού Δεδομένων (εισαγωγή, διαγραφή, τροποποίηση και ανάκτηση δεδομένων)

δυνατότητα εμφύτευσης σε μια γλώσσα υψηλού επιπέδου
μίας εγγραφής τη φορά ή συνόλου τη φορά
διαδικαστικές και μη διαδικαστικές (δηλωτικές)

(το μάθημα σε λιγότερο από 30')

Βάσεις Δεδομένων 2006-2007 Ευαγγελία Πιτουρά 28

Παράδειγμα

ΒΗΜΑ 2: Προγραμματισμός/Υλοποίηση - σε σχεσιακό ΣΔΒΔ

- Ορισμός Σχέσεων (πρόθεση/σχήμα)
- Εισαγωγή Στοιχείων (δημιουργία του αρχικού στιγμιότυπου)
- Διατύπωση Ερωτήσεων

create table R(A1 T1, A2, T2, ...)

insert/delete

select
from
where

(το μάθημα σε λιγότερο από 30' - ΜΕΡΟΣ 1)

Βάσεις Δεδομένων 2006-2007 Ευαγγελία Πιτουρά 29

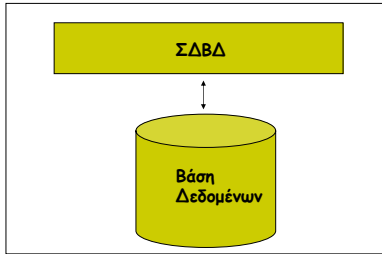
Διεπαφές ΣΔΒΔ

- Βασισζόμενες σε μενού (κατάλογο από επιλογές)
- Γραφικών
- Βασισζόμενες σε φόρμες
- Φυσικής γλώσσας
- Για παραμετρικούς χρήστες
- Για το ΔΒΔ

(το μάθημα σε λιγότερο από 30')

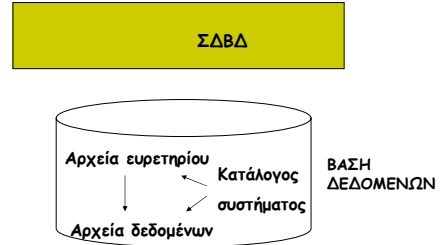
Βάσεις Δεδομένων 2006-2007 Ευαγγελία Πιτουρά 30

Η Δομή ενός ΣΔΒΔ



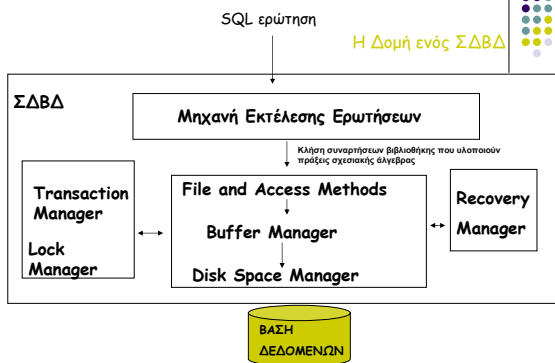
(το μάθημα σε λιγότερο από 30')

Η Δομή ενός ΣΔΒΔ



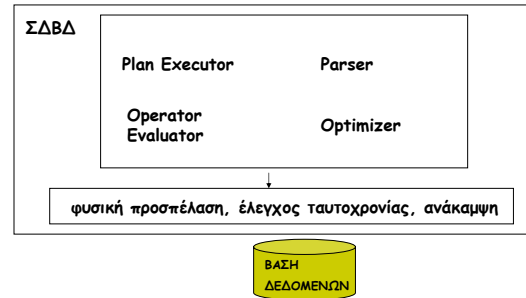
(το μάθημα σε λιγότερο από 30')

Η Δομή ενός ΣΔΒΔ

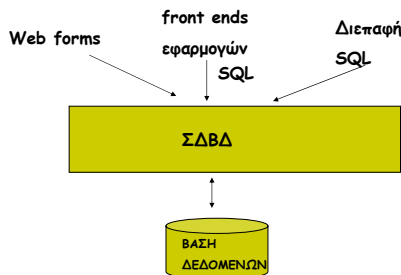


(το μάθημα σε λιγότερο από 30')

Η Δομή ενός ΣΔΒΔ



Η Δομή ενός ΣΔΒΔ



Πλεονεκτήματα ΣΔΒΔ

- Ανεξαρτησία Δεδομένων
- Αποδοτική Προσπέλαση Δεδομένων
- Ακεραιότητα Δεδομένων και Ασφάλεια
- Διαχείριση Δεδομένων
- Ταυτόχρονη προσπέλαση και ανάρρωση από σφάλματα
- Γρήγορη Ανάπτυξη Εφαρμογών

(το μάθημα σε λιγότερο από 30')



- Έλεγχος πλεονασμών
- Εξουσιοδότηση Προσπέλασης
- Παροχή μόνιμης αποθήκευσης
- Πολλαπλές Διεπαφές και Όψεις

(το μάθημα σε λιγότερο από 30')