



Διαχείριση Χωρικών Δεδομένων

Γιάννης Θεοδωρίδης

InfoLab, Τμήμα Πληροφορικής, Πανεπιστήμιο Πειραιά

<http://infolab.cs.unipi.gr>

version: Dec.2009

Πηγές:



- P. A. Longley et al. Geographic Information Systems and Science, 2/e (2005)
- S. Shekhar & S. Chawla. Spatial Databases: A Tour (2003)
- Ε. Στεφανάκης. Βάσεις Γεωγραφικών Δεδομένων και Συστήματα Γεωγραφικών Πληροφοριών (2003)

Περιεχόμενα



- Εισαγωγή
 - GIS - παραδείγματα λογισμικού και εφαρμογών, αρχιτεκτονική
- Spatial DBMS
 - Εισαγωγή – Τύποι χωρικών δεδομένων, Συστατικά ενός SDBMS
 - Σχεδίαση χωρικών ΒΔ – Εννοιολογικός, Λογικός, Φυσικός σχεδιασμός
 - Εμπορικά SDBMS – Oracle Spatial, PostgreSQL
- Σύνοψη, ερευνητικά θέματα

3

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Περιεχόμενα

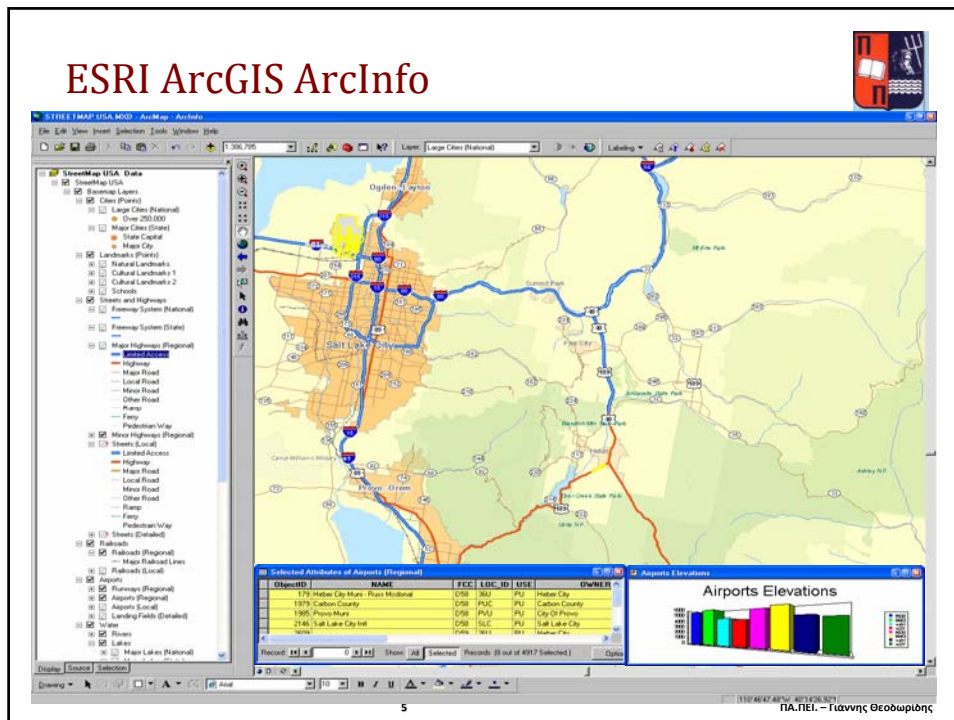


- Εισαγωγή
 - GIS - παραδείγματα λογισμικού και εφαρμογών, αρχιτεκτονική
- Spatial DBMS
 - Εισαγωγή – Τύποι χωρικών δεδομένων, Συστατικά ενός SDBMS
 - Σχεδίαση χωρικών ΒΔ – Εννοιολογικός, Λογικός, Φυσικός σχεδιασμός
 - Εμπορικά SDBMS – Oracle Spatial, PostgreSQL
- Σύνοψη, ερευνητικά θέματα

4

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

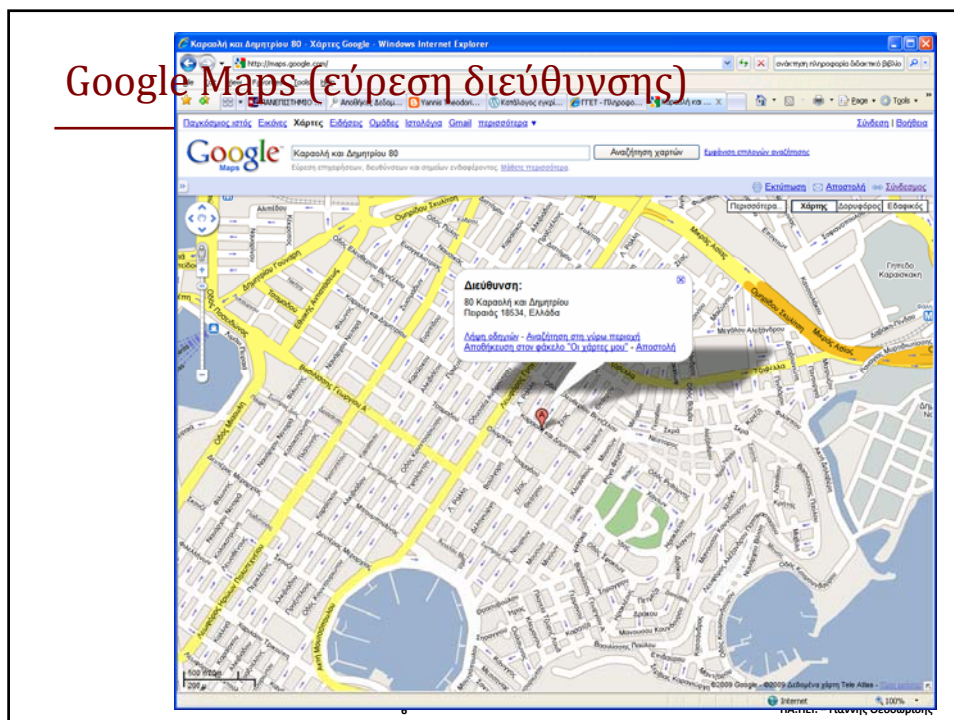
ESRI ArcGIS ArcInfo



5

ΠΑ.ΠΕΙ. - Γιάννης Θεοδοριάδης

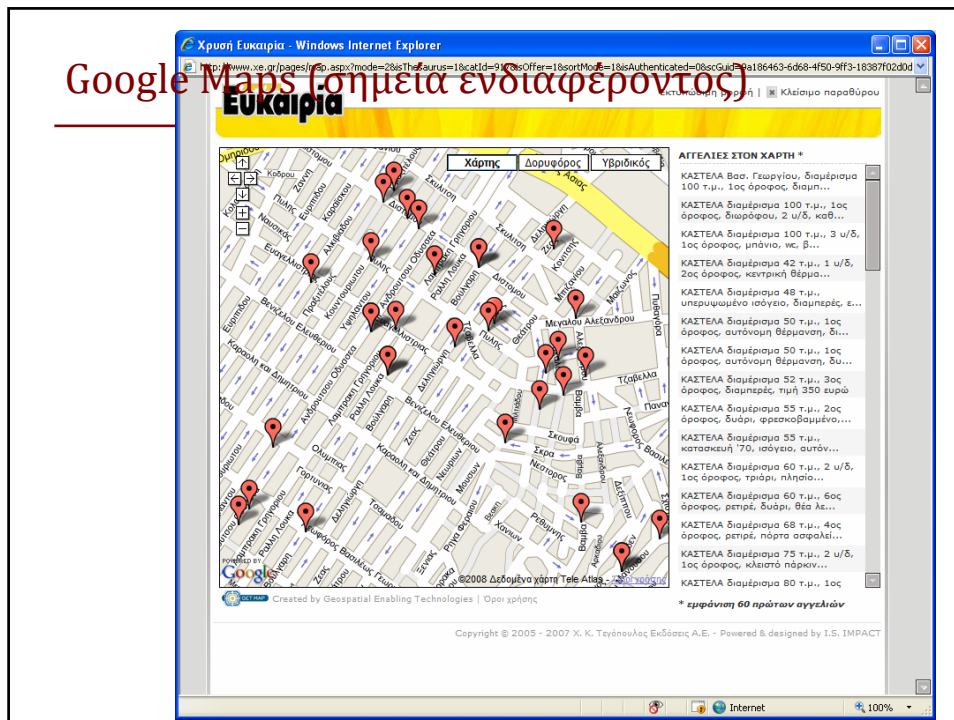
Google Maps (εύρεση διεύθυνσης)



6

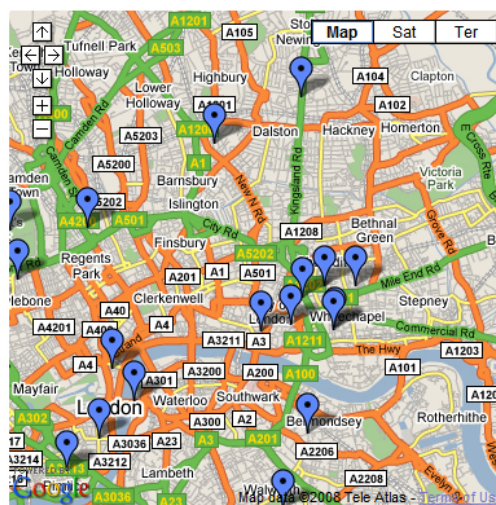
ΠΑ.ΠΕΙ. - Γιάννης Θεοδοριάδης

Google Maps (σημεία ενδιαφέροντος)

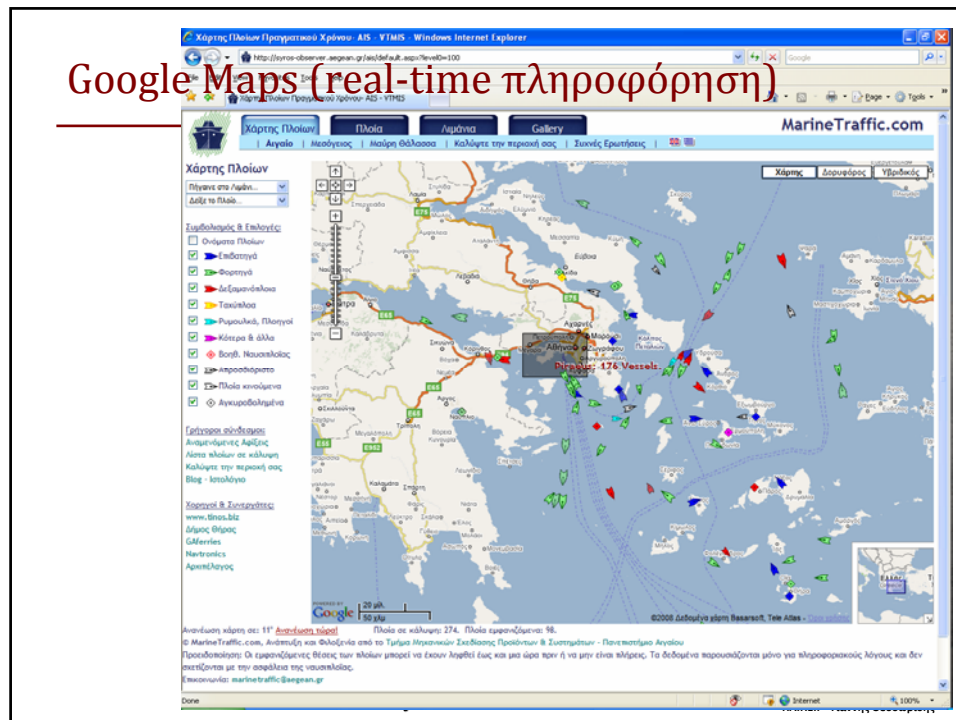


Google Maps (σημεία ενδιαφέροντος)

- Τα σημεία όπου 'έδρασε' ο "Τζακ ο Αντεροβγάλτης" (Λονδίνο, 1888)



Πηγή: archive.timesonline.co.uk



Χωρικά vs. μη-χωρικά δεδομένα



- Παραδείγματα **μη-χωρικών** δεδομένων
 - Ονόματα, τηλέφωνα, διευθύνσεις email, ...
- Παραδείγματα **χωρικών** δεδομένων
 - Κτηματολόγιο, δασολόγιο, κλπ.
 - Δημογραφικά δεδομένα
 - Περιβαλλοντικά (μετεωρολογικά κλπ.) δεδομένα
 - Ιατρικά δεδομένα (ιατρικές εικόνες)
- Άσκηση: εντοπίστε χωρικά και μη-χωρικά δεδομένα στο Facebook, το Amazon κλπ.

Αρχιτεκτονική 3 επιπέδων στα GIS



- Η αλληλεπίδραση με το χρήστη πραγματοποιείται γραφικά μέσω GUI
- Το GUI παρέχει την πρόσβαση στα εργαλεία GIS. Αυτά καθορίζουν τις λειτουργίες που διαθέτει το λογισμικό GIS για την επεξεργασία των γεωγραφικών δεδομένων
- Τα δεδομένα αποθηκεύονται σε αρχεία ή βάσεις δεδομένων από λογισμικό διαχείρισης δεδομένων

Presentation

User Interface

Business Logic

Tools

Data Server

Data Management

Data

11

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Περιεχόμενα



- Εισαγωγή
 - GIS - παραδείγματα λογισμικού και εφαρμογών, αρχιτεκτονική
- Spatial DBMS
 - Εισαγωγή – Τύποι χωρικών δεδομένων, Συστατικά ενός SDBMS
 - Σχεδίαση χωρικών ΒΔ – Εννοιολογικός, Λογικός, Φυσικός σχεδιασμός
 - Εμπορικά SDBMS – Oracle Spatial, PostgreSQL
- Σύνοψη, ερευνητικά θέματα

12

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Υποστήριξη Spatial Data Types σε DBMS



- Παραδοσιακά relational DBMS
 - Υποστηρίζουν απλούς τύπους δεδομένων, π.χ. number, varchar[], date
 - Υποστήριξη χωρικών δεδομένων μπορεί να γίνει απλοϊκά
 - Π.χ. ένα σημείο ως δύο αριθμοί, ένα πολύγωνο ως ...
- Object oriented (OO), Object relational (OR) DBMS
 - Υποστηρίζουν **abstract data types (ADT)** που ορίζονται από το χρήστη
 - Οπότε είναι εφικτή η προσθήκη χωρικών τύπων δεδομένων (π.χ. polygon)

13

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Ορισμός του SDBMS



Ορισμός: Το SDBMS είναι λογισμικό που

- **συνεργάζεται με ένα υφιστάμενο DBMS,**
- (σε λογικό επίπεδο)
 - υποστηρίζει ένα **μοντέλο χωρικών δεδομένων** και τους κατάλληλους **τύπους αναπαράστασης χωρικής πληροφορίας** (μέσω ADT),
 - και προσφέρει **κατάλληλες επεκτάσεις στη γλώσσα ερωτήσεων** (π.χ. SQL),
- (σε φυσικό επίπεδο)
 - προσφέρει υποστήριξη για **εξειδικευμένα ευρετήρια,**
 - **αλγόριθμους επεξεργασίας χωρικών ερωτημάτων,**
 - προσαρμοσμένους **κανόνες βελτιστοποίησης χωρικών ερωτημάτων**

14

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Συστατικά ενός SDBMS



- Άρα τα συστατικά ενός SDBMS είναι
 - μοντέλο χωρικών δεδομένων (spatial data model),
 - γλώσσα χωρικών ερωτήσεων (spatial query language) που υποστηρίζει χωρικούς τύπους δεδομένων (spatial data types),
 - ειδικές οργανώσεις αρχείων και χωρικών ευρετηρίων (spatial indices),
 - τεχνικές επεξεργασίας & βελτιστοποίησης χωρικών ερωτήσεων (spatial query processing & optimization)
- Παράδειγμα: **Oracle Spatial**
 - συνεργάζεται με το Oracle DBMS
 - υποστηρίζει εξειδικευμένους τύπους για χωρικά δεδομένα (π.χ. polygon) και τελεστές (π.χ. overlap) που αναγνωρίζονται από τη γλώσσα ερωτήσεων SQL3
 - διαθέτει χωρικά ευρετήρια (π.χ. R-trees)

15

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Λογικό επίπεδο: Μοντέλα χωρικών δεδομένων



- Χωρικές συσχετίσεις:
 - Τοπολογικές συσχετίσεις, π.χ. overlap
 - Συσχετίσεις απόστασης / κατεύθυνσης πάνω στον Ευκλείδειο χώρο
 - Συσχετίσεις πάνω σε γράφους, π.χ. Shortest-Path
- Μοντέλα χωρικών δεδομένων
 - Προσδιορίζουν διακριτά αντικείμενα ή/και ιδιότητες του χώρου
 - object vs. field model

16

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Λογικό επίπεδο: Γλώσσες χωρικών ερωτήσεων



- Μια γλώσσα χωρικών ερωτήσεων πρέπει να ενσωματώνει μέσα στη γλώσσα (π.χ. SQL3) του υφιστάμενου DBMS
 - τύπους χωρικών δεδομένων, π.χ. point, linestring, polygon, ...
 - χωρικούς τελεστές, π.χ. overlap, distance, nearest neighbor, ...
- Παράδειγμα:


```
SELECT  C.name
FROM    Census_blocks C
WHERE   C.boundary.Area() > 300
```

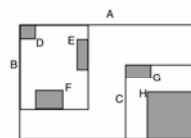
17

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

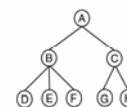
Φυσικό επίπεδο: Αρχεία και χωρικά ευρετήρια



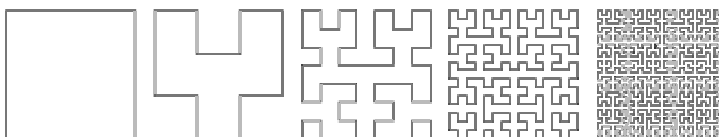
- Το πρόβλημα:
 - Μεταξύ των χωρικών δεδομένων **δεν προκύπτει μια φυσική διάταξη**
 - Δυστυχώς, πολλές μέθοδοι αποθήκευσης / δεικτοδότησης / αναζήτησης δεδομένων βασίζονται στην έννοια της διάταξης
- Πιθανές λύσεις:
 - γραμμές διάσχισης του χώρου (space filling curves)
 - χωρικά ευρετήρια (spatial indices)



a) Partitioning with MBRs



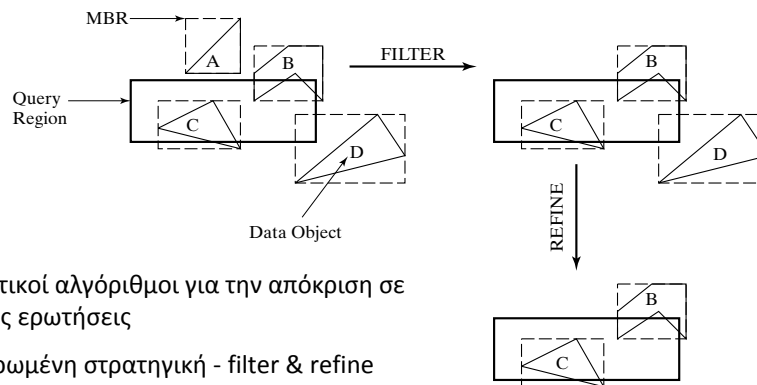
b) R-Tree



18

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Φυσικό επίπεδο: Επεξεργασία & βελτιστοποίηση χωρικών ερωτήσεων



- Αποδοτικοί αλγόριθμοι για την απόκριση σε χωρικές ερωτήσεις
- καθιερωμένη στρατηγική - filter & refine
 - Filter Step: ελέγχουμε τις MBR προσεγγίσεις των χωρικών αντικειμένων
 - Refinement Step: ελέγχουμε τα πραγματικά αντικείμενα

19

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Περιεχόμενα



- Εισαγωγή
 - GIS - παραδείγματα λογισμικού και εφαρμογών, αρχιτεκτονική
- Spatial DBMS
 - Εισαγωγή – Τύποι χωρικών δεδομένων, Συστατικά ενός SDBMS
 - Σχεδίαση χωρικών ΒΔ – Εννοιολογικός, Λογικός, Φυσικός σχεδιασμός
 - Εμπορικά SDBMS – Oracle Spatial, PostgreSQL
- Σύνοψη, ερευνητικά θέματα

20

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σχεδιασμός χωρικής ΒΔ σε τρία βήματα



- Οι εφαρμογές χωρικών ΒΔ μοντελοποιούνται βάσει μίας διαδικασίας τριών βημάτων
 - Εννοιολογικοί τύποι δεδομένων, σχέσεις και περιορισμοί (μοντέλο οντοτήτων-συσχετίσεων - ER model)
 - Λογική απεικόνιση σε ένα σχεσιακό μοντέλο και διασύνδεση με γλώσσα ερωτήσεων (π.χ. SQL)
 - Φυσική αποθήκευση, δεικτοδότηση

21

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σχεδιασμός χωρικής ΒΔ σε τρία βήματα



- Οι εφαρμογές χωρικών ΒΔ μοντελοποιούνται βάσει μίας διαδικασίας τριών βημάτων
 - Εννοιολογικοί τύποι δεδομένων, σχέσεις και περιορισμοί (μοντέλο οντοτήτων-συσχετίσεων - ER model)
 - Λογική απεικόνιση σε ένα σχεσιακό μοντέλο και διασύνδεση με γλώσσα ερωτήσεων (π.χ. SQL)
 - Φυσική αποθήκευση, δεικτοδότηση

22

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα χωρικής εφαρμογής



- Σενάριο χωρικής εφαρμογής: **Οδικό δίκτυο**
- Η μοντελοποίηση ενός οδικού δικτύου ακολουθεί τη λογική του γράφου (κόμβοι και ακμές)
 - Οι δρόμοι (streets) αποτελούνται από τμήματα του δικτύου (road segments), τα οποία με τη σειρά τους προσδιορίζονται από 2 διασταυρώσεις (road intersections)
 - Κάθε τμήμα του δικτύου συνοδεύεται από πληροφορία για την ονομασία και την αρίθμηση (δεξιά και αριστερή πλευρά) του δρόμου και τον ταχ. κώδικα
 - Πάνω στα τμήματα του δικτύου μπορούν να υπάρχουν σημεία ενδιαφέροντος (points of interest – POI)
 - Η παραπάνω ΒΔ συμπεριλαμβάνει ένα γεωγραφικό λεξικό (gazetteer)

23

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το μοντέλο οντοτήτων συσχετίσεων - ER Model

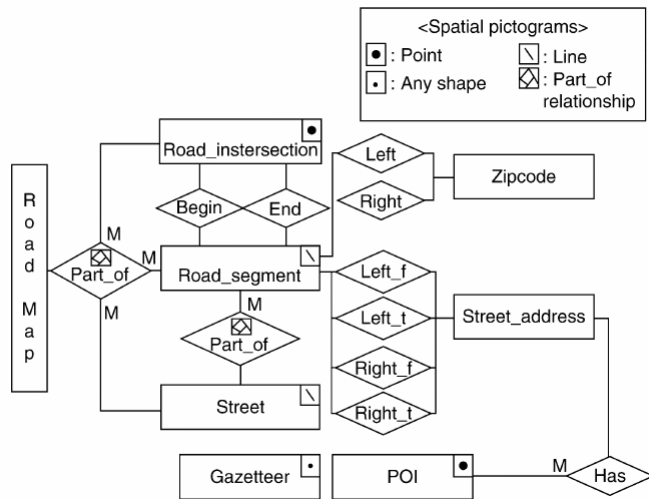


- Σε εννοιολογικό επίπεδο (μοντέλο ER) υπάρχουν 3 βασικές έννοιες
 - Οι **οντότητες** (Entities) έχουν ανεξάρτητη ύπαρξη
 - Παραδείγματα: Street, Road_segment, Road_intersection, POI, ...
 - Χαρακτηρίζονται από **γνωρίσματα** (Attributes)
 - Παράδειγμα: Ένα POI έχει ως γνωρίσματα το όνομά του, κάποιο URL, κλπ.
 - Οι οντότητες αλληλεπιδρούν μέσω **συσχετίσεων** (Relationships)
 - Ένα Street αποτελείται από πολλά Road_segments (συσχέτιση “Part_of”)
- Επέκταση ER με **εικονογράμματα** (Pictogram Enhanced ER – PEER)
 - Περιγράφουν τον τύπο των χωρικών οντοτήτων με κατάλληλη ετικέτα
 - Επιτρέπουν εξαγωγή συμπερασμάτων χωρικών σχέσεων και περιορισμών

24

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα PEER για Road Network DB



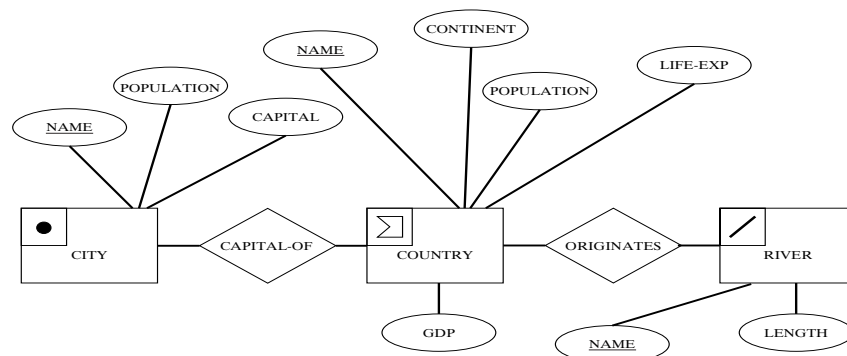
25

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Άλλο παράδειγμα PEER: World DB



- 3 Οντότητες: Country, City, River
- 2 συσχετίσεις: capital-of, originates-in



26

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σχεδιασμός χωρικής ΒΔ σε τρία βήματα



- Οι εφαρμογές χωρικών ΒΔ μοντελοποιούνται βάσει μίας διαδικασίας τριών βημάτων
 - Εννοιολογικοί τύποι δεδομένων, σχέσεις και περιορισμοί (μοντέλο οντοτήτων-συσχετίσεων - ER model)
 - Λογική απεικόνιση σε ένα σχεσιακό μοντέλο και διασύνδεση με γλώσσα ερωτήσεων (π.χ. SQL)
 - Φυσική αποθήκευση, δεικτοδότηση

27

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Λογικό μοντέλο



- 3 σχέσεις (Relations)
 - **Country**(Name, Cont, Pop, GDP, Life-Exp, Shape)
 - **City**(Name, Country, Pop, Capital, Shape)
 - **River**(Name, Origin, Length, Shape)
- Κλειδιά
 - Primary keys: Country.Name, City.Name, River.Name
 - Foreign keys: River.Origin, City.Country
- Δεδομένα
 - Βλ. Επόμενη διαφάνεια

28

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Πίνακες δεδομένων World database



COUNTRY	Name	Cont	Pop (millions)	GDP (billions)	Life-Exp	Shape
	Canada	NAM	30.1	658.0	77.08	Polygonid-1
	Mexico	NAM	107.5	694.3	69.36	Polygonid-2
	Brazil	SAM	183.3	1004.0	65.60	Polygonid-3
	Cuba	NAM	11.7	16.9	75.95	Polygonid-4
	USA	NAM	270.0	8003.0	75.75	Polygonid-5
	Argentina	SAM	36.3	348.2	70.75	Polygonid-6

(a) Country

CITY	Name	Country	Pop (millions)	Capital	Shape
	Havana	Cuba	2.1	Y	Pointid-1
	Washington, D.C.	USA	3.2	Y	Pointid-2
	Monterrey	Mexico	2.0	N	Pointid-3
	Toronto	Canada	3.4	N	Pointid-4
	Brasilia	Brazil	1.5	Y	Pointid-5
	Rosario	Argentina	1.1	N	Pointid-6
	Ottawa	Canada	0.8	Y	Pointid-7
	Mexico City	Mexico	14.1	Y	Pointid-8
	Buenos Aires	Argentina	10.75	Y	Pointid-9

(b) City

RIVER	Name	Origin	Length (kilometers)	Shape
	Rio Parana	Brazil	2600	LineStringid-1
	St. Lawrence	USA	1200	LineStringid-2
	Rio Grande	USA	3000	LineStringid-3
	Mississippi	USA	6000	LineStringid-4

(c) River

Χρήση χωρικών λειτουργιών στο SELECT clause



- **Query:** Βρες όνομα, πληθυσμό και εμβαδόν κάθε χώρας του πίνακα Country

```
SELECT C.Name, C.Pop, Area(C.Shape) AS "Area"
FROM Country C
```

Χρήση χωρικού τελεστή Distance



- **Query:** Βρες το GDP και την απόσταση κάθε πρωτεύουσας όλων των χωρών από τον ισημερινό.

```
SELECT Co.GDP, Distance(Point(Ci.Shape.x, 0),Ci.Shape) AS "Distance"
FROM Country Co,City Ci
WHERE Co.Name = Ci.Country
AND Ci.Capital = 'Y'
```

Co. Name	Co. GDP	Dist-to-Eq (in Km).
Havana	16.9	2562
Washington, D.C.	8003	4324
Brasilia	1004	1756
Ottawa	658	5005
Mexico City	694.3	2161
Buenos Aires	348.2	3854

31

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Χρήση χωρικών λειτουργιών στο WHERE clause



- **Query:** Βρες τα ονόματα όλων των χωρών που είναι γείτονες των USA

```
SELECT C1.Name AS "Neighbors of USA"
FROM Country C1,Country C2
WHERE Touch(C1.Shape,C2.Shape)=1
AND C2.Name = 'USA'
```

32

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Χωρική ερώτηση με πολλαπλούς πίνακες



- **Query:** Για όλα τα ποτάμια στον πίνακα River, βρες τις χώρες από τις οποίες περνάνε.

```
SELECT R.Name, C.Name
FROM River R, Country C
WHERE Cross(R.Shape,C.Shape)=1
```

- **Άσκηση:** να τροποποιήσετε την ερώτηση ώστε να βρίσκει το μήκος του ποταμιού στο εσωτερικό κάθε χώρας

33

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα χωρικής ερώτησης...Buffer and Overlap



- **Query:** Το ποτάμι St. Lawrence μπορεί να προμηθεύει με νερό πόλεις σε απόσταση 300 χλμ. Βρείτε τις πόλεις αυτές.

```
SELECT Ci.Name
FROM City Ci, River R
WHERE Overlap(Ci.Shape, Buffer(R.Shape,300))=1
AND R.Name ='St.Lawrence '
```

34

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Χρήση χωρικής λειτουργίας σε μία αθροιστική ερώτηση



- **Query:** Αναφέρατε τις χώρες, ταξινομημένες ανά πλήθος γειτονικών χωρών

```
SELECT Co.Name, Count(Co1.Name)
FROM Country Co, Country Co1
WHERE Touch(Co.Shape,Co1.Shape)
GROUP BY Co.Name
ORDER BY Count(Co1.Name)
```

35

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Χρήση χωρικών λειτουργιών σε ένθετες ερωτήσεις



- **Query:** Για κάθε ποτάμι, βρες την κοντινότερη πόλη.

```
SELECT C1.Name, R1.Name
FROM City C1, River R1
WHERE Distance (C1.Shape,R1.Shape) <= ALL
    ( SELECT Distance(C2.Shape)
      FROM City C2
      WHERE C1.Name <> C2.Name
    )
```

- **Άσκηση:** Είναι ενδιαφέρον να παρατηρήσετε ότι η εύρεση της κοντινότερης απόστασης από κάθε ποτάμι στην κοντινότερη πόλη είναι πολύ απλούστερη και δεν απαιτεί ένθετη ερώτηση.

36

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Εμφωλευμένη χωρική ερώτηση



- **Query:** Βρες τις χώρες που έχουν μόνο μία γειτονική χώρα. Μία χώρα θεωρείται γειτονική με μία άλλη όταν έχουν εν μέρει κοινά σύνορα των ηπειρωτικών περιοχών τους (άρα, οι χώρες νησιά, όπως η Ισλανδία, δεν έχουν γείτονες).

```
SELECT Co.Name
FROM Country Co
WHERE Co.Name IN (SELECT Co.Name
FROM Country Co,Country Co1
WHERE Touch(Co.Shape,Co1.Shape)
GROUP BY Co.Name
HAVING Count(*)=1)
```

37

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Επανεγγραφή εμφωλευμένης χωρικής ερώτησης με Views



```
CREATE VIEW Neighbor AS
SELECT Co.Name, Count(Co1.Name) AS num neighbors
FROM Country Co,Country Co1
WHERE Touch(Co.Shape,Co1.Shape)
GROUP BY Co.Name

SELECT Co.Name, num neighbors
FROM Neighbor
WHERE num neighbor = ( SELECT Max(num neighbors)
FROM Neighbor )
```

38

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σχεδιασμός χωρικής ΒΔ σε τρία βήματα



- Οι εφαρμογές χωρικών ΒΔ μοντελοποιούνται βάσει μίας διαδικασίας τριών βημάτων
 - Εννοιολογικοί τύποι δεδομένων, σχέσεις και περιορισμοί (μοντέλο οντοτήτων-συσχετίσεων - ER model)
 - Λογική απεικόνιση σε ένα σχεσιακό μοντέλο και διασύνδεση με γλώσσα ερωτήσεων (π.χ. SQL)
 - Φυσική αποθήκευση, δεικτοδότηση

Φυσικό μοντέλο σχεδιασμού



- Τι είναι ένα φυσικό μοντέλο ΒΔ;
 - Πώς θα υλοποιηθεί αυτό που περιγράφεται στο λογικό μοντέλο
 - Με αποτελεσματικό και χωρίς σφάλματα τρόπο
- Λόγοι μάθησης εννοιών φυσικών μοντέλων
 - Σωστή επιλογή εμπορικών ΣΔΒΔ
 - Μερικά ΣΔΒΔ δεν υποστηρίζουν ευρετήρια πάνω σε χωρικά δεδομένα !
 - Σωστή διαχείριση δυνατοτήτων ΣΔΒΔ με στόχο τον καλύτερο συντονισμό της απόδοσης του (performance tuning)
 - Παράδειγμα: μία ερώτηση εκτελείται αργά → Ίσως χρειάζεται να δημιουργηθεί ένα ευρετήριο

Φυσικό μοντέλο δεδομένων για SDBMS



- Είναι το μοντέλο των relational DBMS κατάλληλο για τα χωρικά δεδομένα;
 - Τα relational DBMS παρέχουν απλούς τύπους τιμών, όπως αριθμοί και συμβολοσειρές
 - Έννοια της ολικής διάταξης
 - Η έννοια της ολικής διάταξης δεν έχει νόημα για τα χωρικά δεδομένα (π.χ. σημεία στο επίπεδο)
- Εναλλακτικές λύσεις:
 - Επαναχρησιμοποίηση εννοιών σχεσιακού μοντέλου
 - Καινοτόμες χωρικές τεχνικές

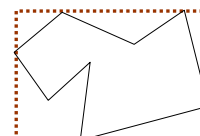
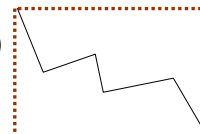
41

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Υποθέσεις για το φυσικό μοντέλο SDBMS



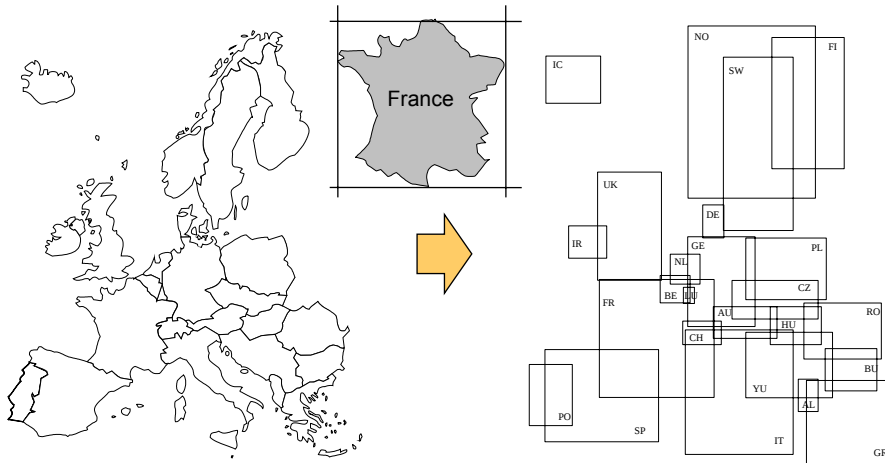
- Χωρικά δεδομένα
 - Ο αριθμός διαστάσεων του χώρου είναι μικρός π.χ. 2 ή 3
 - Τύποι δεδομένων: συμβατοί με το OpenGIS πρότυπο
 - Προσεγγίσεις αντικειμένων με έκταση (π.χ. γραμμές, πολύγωνα)
- Minimum (Orthogonal) Bounding Rectangle (MOBR ή MBR)
 - MBR(obj) είναι το μικρότερο ορθογώνιο παραλληλόγραμμο που περιβάλλει το αντικείμενο obj και είναι παράλληλο με το καρτεσιανό σύστημα συντεταγμένων
 - Ανάγκη φιλτραρίσματος και εκλέπτυνσης (filter and refinement)
- Χωρικές λειτουργίες πάνω στα MBRs



42

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα εφαρμογής MBR



43

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Απλές χωρικές ερωτήσεις και λειτουργίες



- Το φυσικό μοντέλο παρέχει απλές λειτουργίες που χρειάζονται στην επίλυση χωρικών ερωτήσεων!
- Τυπικές ερωτήσεις
 - **Point query**: Βρες όλα τα αντικείμενα που περικλείουν ένα δοθέν σημείο
 - **Range query**: Βρες όλα τα αντικείμενα που κείνται μέσα σε μια (συνήθως ορθογώνια) περιοχή
 - **Nearest neighbor query**: Βρες το πλησιέστερο αντικείμενο σε σχέση με ένα σημείο
 - **Spatial Join** (μεταξύ δύο συνόλων αντικειμένων): Βρες όλα τα ζεύγη αντικειμένων (ένα από κάθε σύνολο) που ικανοποιούν μια χωρική συνθήκη (συνήθως, overlap)

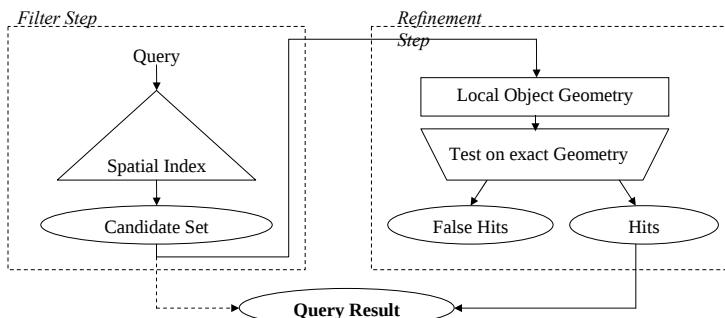
44

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Η διαδικασία «Filter-Refinement»



- Επεξεργασία μίας χωρικής ερώτησης
 - βήμα φιλτραρίσματος** (Filter step): «σίγουρες» και «υποψήφιες» απαντήσεις βάσει της προσέγγισης MBR του αντικειμένου
 - βήμα εκλέπτυνσης** (Refinement step): έλεγχος των «υποψήφιων» απαντήσεων βάσει της πλήρους γεωμετρίας του αντικειμένου



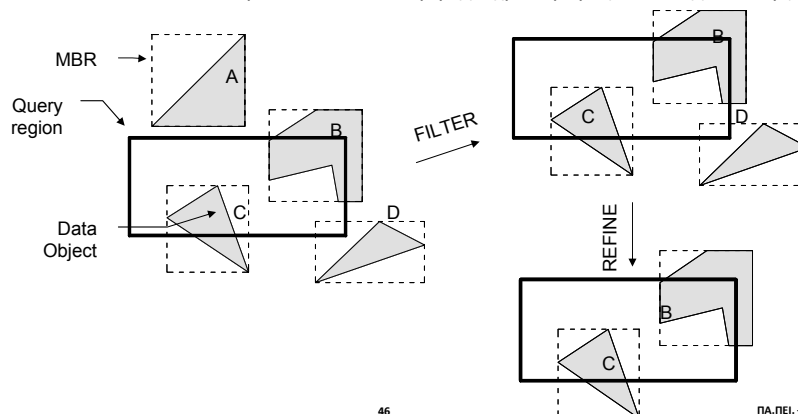
45

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα filter – refinement



- Range queries (find all data objects A that overlap query object Q)
 - filter step: εκτελείται ο προσεγγιστικός / ταχύς τελεστής overlap(MBR(A), Q)
 - refinement step: εκτελείται ο ακριβής / χρονοβόρος τελεστής overlap(A, Q)



46

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Οργάνωση αρχείων & ευρετήρια



- 2 εναλλακτικές:
 - Επαναχρησιμοποίηση εννοιών σχεσιακού φυσικού μοντέλου – η έννοια της διάταξης (ordering)
 - Οι καμπύλες διάσχισης χώρου (space filling curves) ορίζουν μία ολική διάταξη για τα σημεία
 - Άρα μπορούν να χρησιμοποιηθούν διατεταγμένα αρχεία ή ευρετήρια τύπου B-trees
 - Καινοτόμες τεχνικές
 - Χωρικά ευρετήρια, π.χ. δένδρα περιοχών (R-trees), δένδρα τεταρτημορίων (Quadtree), αρχεία πλέγματος (Gridfile)
 - Παρέχουν καλύτερες αποδόσεις στους υπολογισμούς

47

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Διατεταγμένα αρχεία

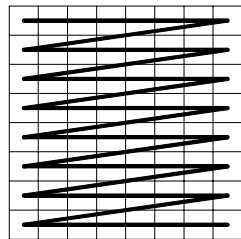


- Πρόβλημα:
 - Η διάταξη δεν έχει κάποιο φυσικό νόημα στα χωρικά δεδομένα
 - Είναι αδύνατο να βρεθεί απεικόνιση «1-1» και «επί» που να μετασχηματίζει τα σημεία του 2d χώρου σε μία διάσταση
- Πιθανή λύση:
 - Μετασχηματισμός των χωρικών αντικειμένων από δύο σε μία διάσταση χρησιμοποιώντας μία «καμπύλη διάσχισης χώρου» (space filling curve). Η λύση μπορεί να εφαρμοστεί μόνο σε διακριτό και όχι συνεχή χώρο.
 - Επιθυμούμε τη διατήρηση των σχέσεων εγγύτητας των αντικειμένων στο μετασχηματισμένο χώρο
 - Παραδείγματα: Z-curve, Hilbert-curve

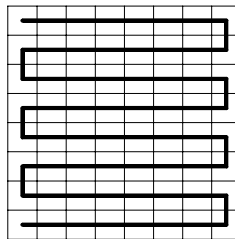
48

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

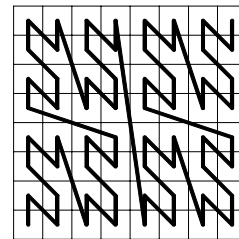
Καμπύλες Διάσχισης Χώρου



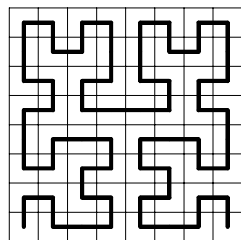
Row



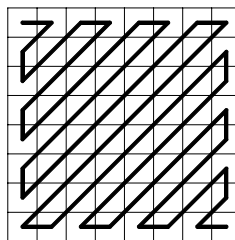
Row-prime



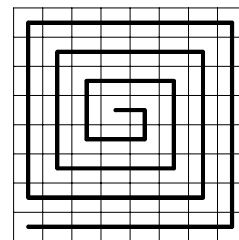
Z-Order (Morton)



Hilbert



Cantor diagonal



Spiral

49

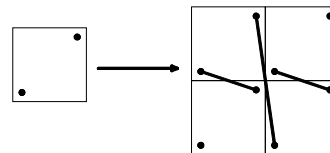
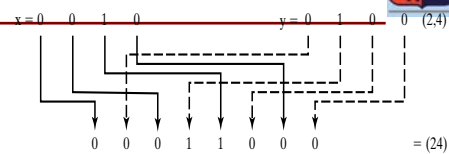
ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Διάσχιση Z



■ Τι είναι μία Z-curve?

- Μία καμπύλη διάσχισης χώρου
- Παραγόμενη από διαπλεκόμενα (interleaved) bits
 - x, y συντεταγμένες
- Προκύπτει μια διάσχιση που μοιάζει με N ή Z



n=0

n=1

n=2

n=3

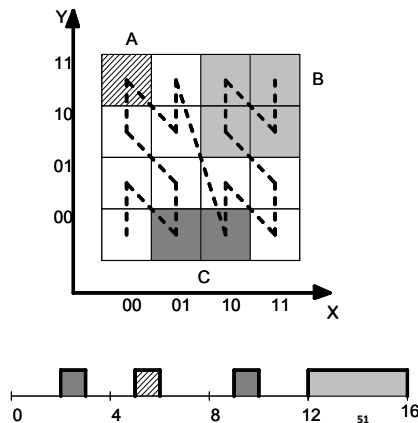
50

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα Z-τιμών



- Ένας χάρτης με τα χωρικά αντικείμενα A, B, C, και οι Z-τιμές που προκύπτουν
 - το C παίρνει z-τιμές 2 και 8, οι οποίες δεν είναι γειτονικές!
 - Άσκηση: υπολογίστε τις Z-τιμές για το B



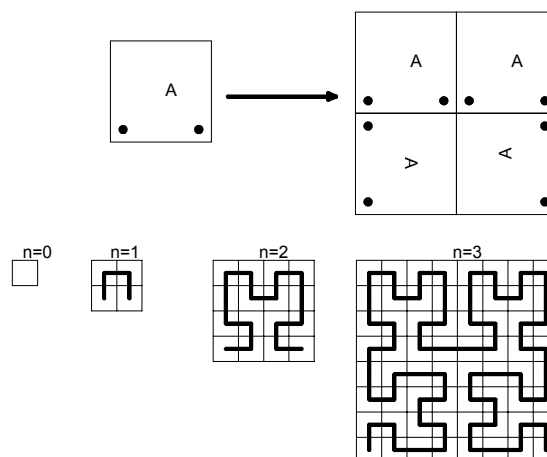
Object	Points	x	y	Interleave	z-value
A	1	00	11	0101	5
B	1				
	2				
	3				
	4				
C	1	01	00	0010	2
	2	10	00	1000	8

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Διάσχιση Hilbert



- Τι είναι μία Hilbert curve?
 - Μία καμπύλη διάσχισης χώρου
 - Πιο πολύπλοκη στη δημιουργία της από τη Z
 - Λόγω περιστροφών (στα 2 κάτω τεταρτημόρια)



52

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Χωρικά ευρετήρια



- Η δομή ενός ευρετηρίου (index) εξαρτάται από τον τύπο δεδομένων του δεικτοδοτούμενου χαρακτηριστικού
- Χαρακτηριστικά με ολική διάταξη
 - αριθμοί, συμβολοσειρές, σημεία ταξινομημένα με καμπύλες διάσχισης χώρου
 - Το B-tree είναι μία δημοφιλής δομή οργάνωσης τέτοιων τύπων
- Χωρικά αντικείμενα (π.χ. πολύγωνα)
 - Η χωρική οργάνωση είναι πιο αποτελεσματική
 - Δύο βασικές οικογένειες: Τετραδικά δέντρα (**Quadtrees**) και R-δέντρα (**R-trees**)

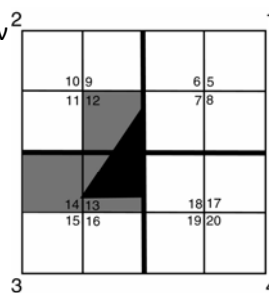
53

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

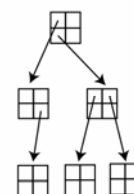
Quadtree



- Ιεραρχική αποσύνθεση του χώρου σε τεταρτημόρια (quadrants)
- Κάθε επίπεδο στο Quadtree αναπαριστά το αντικείμενο ως σύνολο MBRs με βάση τα τεταρτημόρια ανάλυσης του χώρου.
 - Κάθε επίπεδο είναι μια πιο ακριβής αναπαράσταση του αντικειμένου.
- Ο αριθμός των απαιτούμενων επιπέδων εξαρτάται από την επιθυμητή ακρίβεια.
- Μη ισοζυγισμένα δένδρα



a) Representing Triangle With Quadrants



b) Quad Tree

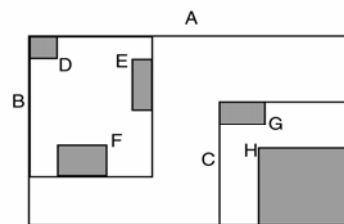
54

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

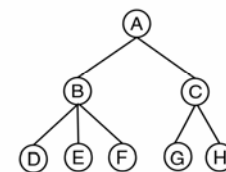
R-tree



- Βασική ιδέα: Γενίκευση B-tree σε χωρικά δεδομένα
- Ιδιότητες των R-trees
 - Ισοζυγισμένα / Οι κόμβοι είναι ορθογώνια / Το ορθογώνιο-παιδί περιέχεται πλήρως εντός του ορθογωνίου-πατέρα
 - Μπορεί να προκύψει χωρική επικάλυψη μεταξύ των κόμβων!



a) Partitioning with MBRs

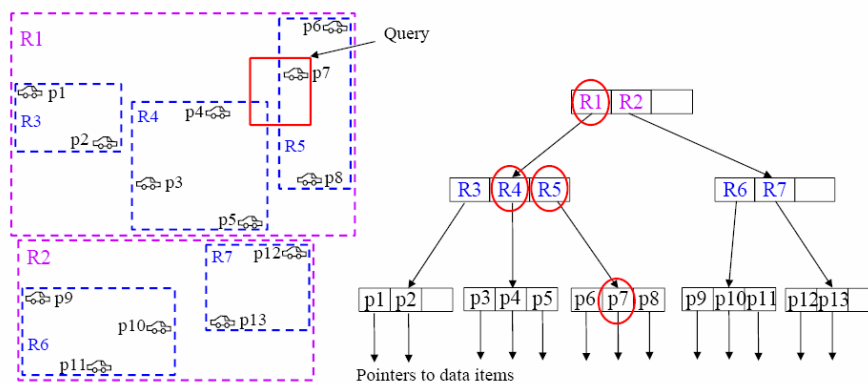


b) R-Tree

55

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Παράδειγμα αναζήτησης σε R-tree: range query



56

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Επεξεργασία & βελτιστοποίηση ερωτήσεων



- Query Processing and Optimization (QPO)
- Βασική ιδέα QPO
 - Στην SQL, οι ερωτήσεις εκφράζονται σε μία υψηλού επιπέδου δηλωτική μορφή
 - Η διαδικασία QPO μεταφράζει μία SQL ερώτηση σε ένα «σχέδιο εκτέλεσης» (execution plan)
 - Πάνω στο φυσικό μοντέλο δεδομένων
 - Χρησιμοποιώντας λειτουργίες σε δομές αρχείων, ευρετήρια κτλ.
 - Ένα ιδανικό σχέδιο εκτέλεσης απαντά στην ερώτηση Q στο συντομότερο δυνατό χρόνο

57

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Τρεις έννοιες-κλειδιά για QPO



- Στρατηγικές επεξεργασίας ερωτήσεων
 - Τα relational DBMS διατηρούν ένα μικρό αριθμό στρατηγικών για κάθε δομικό συστατικό
 - Π.χ. ένα match query (δηλ. μια συνθήκη ισότητας στο Where clause) μπορεί να απαντηθεί μέσω ενός index (αν υπάρχει) ή διατρέχοντας όλο το αρχείο δεδομένων (αν δεν υπάρχει index)
- Βελτιστοποίηση ερωτήσεων
 - Επιλογή της πιο αποτελεσματικής στρατηγικής εκτέλεσης, βάσει κάποιων παραμέτρων ή παραδοχών
 - Παραδείγματα παραμέτρων: μέγεθος πίνακα, διαθέσιμα ευρετήρια, ...
 - Παράδειγμα παραδοχών: μέγεθος ενδιάμεσου αποτελέσματος

58

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Προκλήσεις QPO σε SDBMS



- Επιλογή δομικών συστατικών
 - Πλούσιο σύνολο χωρικών τύπων δεδομένων, λειτουργιών
 - Έλλειψη κοινά αποδεκτών δομικών συστατικών
 - Ποικίλουν ανάλογα με τον κατασκευαστή και το προϊόν
 - Ενδεικτικά: spatial selection (point / range / nearest neighbor query), spatial join (overlap / distance join)
- Επιλογή στρατηγικών
 - Τα μοντέλα κόστους είναι πιο πολύπλοκα καθώς
 - Οι χωρικές ερωτήσεις έχουν υψηλές απαιτήσεις, τόσο σε CPU όσο και σε I/O

59

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Επιλογή δομικών συστατικών



- Αντιπροσωπευτικά δομικά συστατικά (με παραδείγματα από χάρτη Ευρώπης)
 - Point Query – π.χ. να βρεθεί σε ποιο κράτος ανήκει ένα τονισμένο σημείο στο χάρτη
 - Range Query – π.χ. να βρεθούν όλα τα κράτη που εμφανίζονται στο παράθυρο
 - Nearest Neighbor Query – π.χ. να βρεθεί το κράτος που βρίσκεται πλησιέστερα σε ένα τονισμένο σημείο
 - Spatial Join Query – π.χ. να βρεθούν τα ζεύγη των κρατών που συνορεύουν μεταξύ τους



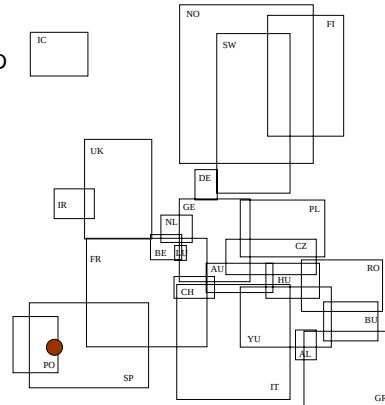
60

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Στρατηγικές για Point Queries



- Point Query – π.χ. να βρεθεί σε ποιο κράτος ανήκει ένα τονισμένο σημείο στο χάρτη
 - Είσοδος: ένα σύνολο χωρικών αντικειμένων D και ένα σημείο Q
 - Έξοδος: ένα ή περισσότερα αντικείμενα στα οποία ανήκει το σημείο
- Στρατηγικές:
 - Αν το D είναι διατεταγμένο (π.χ. σε Z-order)
 - δυαδική αναζήτηση στη Z-order
 - Αν υπάρχει χωρικό ευρετήριο (π.χ. R-tree) πάνω στο D
 - λειτουργία point query του ευρετηρίου



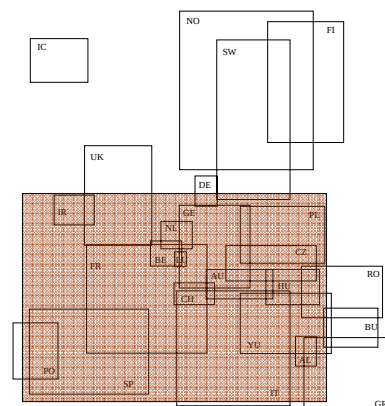
61

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Στρατηγικές για Range Queries



- Range Query - π.χ. να βρεθούν όλα τα κράτη που εμφανίζονται στο παράθυρο Q
 - Είσοδος: ένα σύνολο χωρικών αντικειμένων D και ένα παράθυρο Q
 - Έξοδος: τα αντικείμενα που έχουν επικάλυψη με το παράθυρο
- Στρατηγικές:
 - Αν το D είναι διατεταγμένο (π.χ. σε Z-order)
 - Προσδιορισμός του εύρους των τιμών του Z-order που ικανοποιούν την ερώτηση
 - δυαδική αναζήτηση για την εύρεση της μικρότερης αποδεκτής τιμής Z
 - γραμμική αναζήτηση μέχρι τη μεγαλύτερη αποδεκτή τιμή Z
 - Αν υπάρχει χωρικό ευρετήριο (π.χ. R-tree) πάνω στο D
 - λειτουργία range query του ευρετηρίου



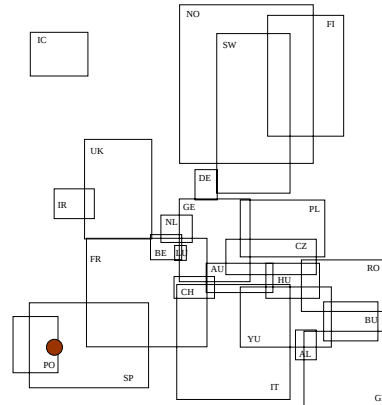
62

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Στρατηγικές για NN Queries



- NN Query - π.χ. να βρεθεί το κράτος που βρίσκεται πλησιέστερα σε ένα τονισμένο σημείο
 - Είσοδος: ένα σύνολο χωρικών αντικειμένων D και ένα σημείο Q
 - Έξοδος: ένα αντικείμενο, αυτό που βρίσκεται πλησιέστερα στο σημείο
 - Παραλλαγή: k-NN Query
- Στρατηγικές:
 - Αν το D είναι διατεταγμένο (π.χ. σε Z-order)
 - Εντοπισμός της σελίδας δίσκου του D που περιέχει την τοποθεσία του σημείου Q (Point Query)
 - M = minimum distance (Q, αντικείμενα που φορτώθηκαν)
 - Έλεγχος όλων των αντικειμένων σε απόσταση M από το Q (Range Query)
 - Αν υπάρχει χωρικό ευρετήριο (π.χ. R-tree) πάνω στο D
 - λειτουργία NN query του ευρετηρίου



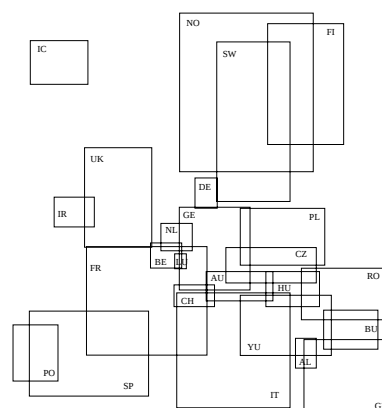
63

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Στρατηγικές για Spatial Joins



- Spatial Join Query – π.χ. να βρεθούν τα ζεύγη των κρατών που συνορεύουν μεταξύ τους
 - Είσοδος: δύο σύνολα χωρικών αντικειμένων D1 και D2 και ένας χωρικός τελεστής σύνδεσης τ
 - Έξοδος: ένα σύνολο από ζεύγη αντικειμένων (r, s) , $r \in D1$ και $s \in D2$, για τα οποία ικανοποιείται ο τελεστής τ
- Στρατηγικές:
 - φωλιασμένοι βρόγχοι (nested loops):
 - Έλεγχος όλων των πιθανών ζευγαριών (r, s) βάσει του τελεστή τ
 - Διαμερισμός χώρου (space partitioning):
 - Έλεγχος ζευγαριών (r, s) μόνο από κοινές περιοχές
 - Αν υπάρχει χωρικό ευρετήριο πάνω στο D1 ή/και το D2
 - Διάφορες τεχνικές spatial join, με βάση το αν υπάρχουν και τα δύο ευρετήρια ή μόνο το ένα.



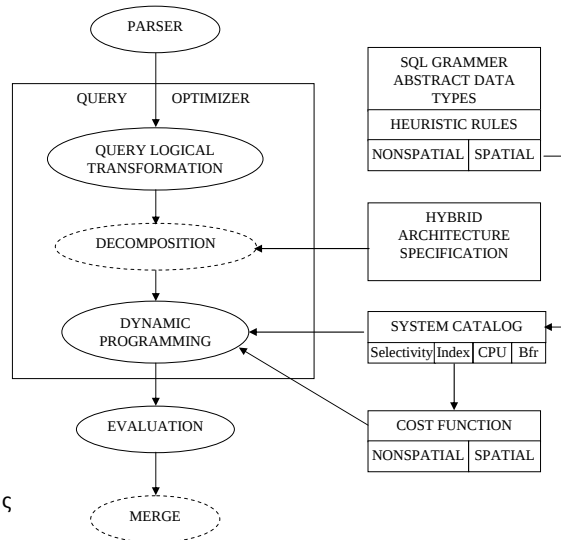
64

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Διεργασία QPO

■ Η διαδικασία

- Αφετηρία : Μία SQL ερώτηση
- Προορισμός: Ένα «καλό» σχέδιο εκτέλεσης
- Ενδιάμεσες στάσεις
 - Δένδρο ερώτησης
 - Λογικός μετασχηματισμός δένδρου
 - Επιλογή στρατηγικής



65

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

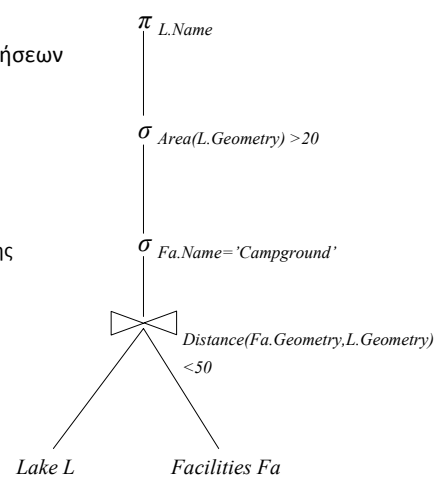
Δένδρα ερωτήσεων

- Κόμβοι = δομικά συστατικά χωρικών ερωτήσεων
- Παιδιά = είσοδος στα δομικά συστατικά
- Φύλλα = Πίνακες
- Παράδειγμα

- Βρες όλες τις λίμνες με εμβαδόν > 20 km² σε ακτίνα 50 km από το χώρο κατασκήνωσης

```

SELECT L.Name
FROM Lake L, Facilities Fa
WHERE Area(L.Geometry) > 20
AND Fa.Name = 'Campground'
AND Distance (Fa.Geometry, L.Geometry) < 50
  
```



66

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Λογικοί μετασχηματισμοί δένδρων ερώτησης



- Ο μετασχηματισμός δεν αλλάζει την απάντηση

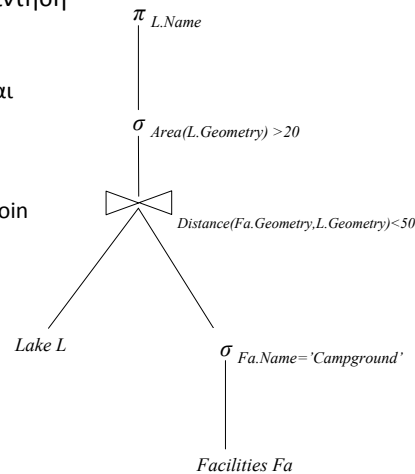
- ... αλλά αλλάζει το υπολογιστικό κόστος
 - Μειώνοντας τα δεδομένα που παράγονται από τα ενδιάμεσα αποτελέσματα

- Παράδειγμα μετασχηματισμού

- Εκτέλεση της λειτουργίας select μετά το join
- Μειώνει το μέγεθος του πίνακα για τη λειτουργία join

- Άλλοι συνήθεις μετασχηματισμοί

- Μετάθεση της λειτουργίας project
- Επαναδιάταξη λειτουργιών join
- ...



67

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Λογικοί μετασχηματισμοί και χωρικές ερωτήσεις

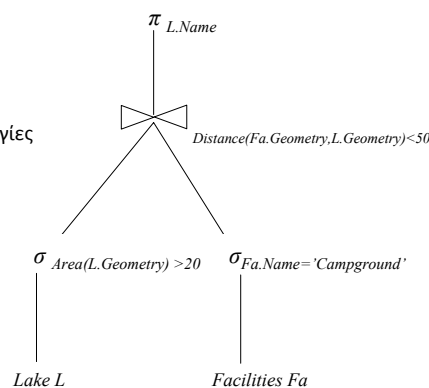


- Παραδοσιακοί κανόνες λογικών μετασχηματισμών

- Για σχεσιακές ερωτήσεις με απλούς τύπους δεδομένων και λειτουργίες
 - Τα CPU κόστη είναι πολύ μικρότερα από τα I/O κόστη
- Ανάγκη επαναθεώρησης για τις χωρικές ερωτήσεις
 - Πολύπλοκοι τύποι δεδομένων, λειτουργίες
 - Το κόστος CPU είναι υψηλότερο

- Παράδειγμα:

- Μετάθεση spatial selection κάτω από το spatial join
- Μπορεί να μη μειώσει το κόστος, αν η λειτουργία area() είναι ακριβότερη από την distance()



68

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σχέδια εκτέλεσης

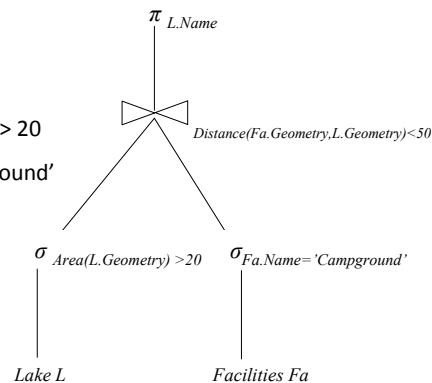


- Ένα σχέδιο εκτέλεσης (execution plan) αποτελείται από τρία μέρη

- Ένα δένδρο ερώτησης και
- Μία στρατηγική εκτέλεσης κάθε μη-τερματικό κόμβου

- Παράδειγμα

- Αναζήτηση για $\text{Area}(\text{L.Geometry}) > 20$
- Αναζήτηση για $\text{Fa.Name} = \text{'Campground'}$ με χρήση index
- Αναζήτηση για $\text{Distance}(\text{Fa}, \text{L}) < 50$ με χρήση space-partitioning join
- Προβολή on-the-fly



69

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Περιεχόμενα



- Εισαγωγή
 - GIS - παραδείγματα λογισμικού και εφαρμογών, αρχιτεκτονική
- Spatial DBMS
 - Εισαγωγή – Τύποι χωρικών δεδομένων, Συστατικά ενός SDBMS
 - Σχεδίαση χωρικών ΒΔ – Εννοιολογικός, Λογικός, Φυσικός σχεδιασμός
 - Εμπορικά SDBMS – Oracle Spatial, PostgreSQL
- Σύνοψη, ερευνητικά θέματα

70

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Συστήματα χωρικών βάσεων δεδομένων



Εισαγωγή στην Oracle Spatial

(εκδόσεις 8i, 9i, 10g)

71

ΠΑ.ΠΕΙ. – Γιάννης Θεοδορίδης

Oracle Spatial 8i, 9i, 10g



- Η Oracle Spatial παρέχει:
 - ένα σχήμα βάσης σε SQL και λειτουργίες για την τήρηση, ενημέρωση, ανάκτηση χωρικών οντοτήτων μέσα σε μια βάση δεδομένων σε περιβάλλον Oracle (εκδόσεις 8i, 9i, 10g).
- Η Oracle Spatial περιλαμβάνει:
 - σχήμα (MDSYS) για την αποθήκευση, σύνταξη και σημασιολογία των γεωμετρικών τύπων δεδομένων
 - μηχανισμό δεικτοδότησης χωρικών δεδομένων (spatial indexing)
 - σύνολο τελεστών (operators) και συναρτήσεων (functions) για την διατύπωση χωρικών ερωτημάτων (queries)
 - εργαλεία διαχείρισης της βάσης

72

ΠΑ.ΠΕΙ. – Γιάννης Θεοδορίδης

Δύο εναλλακτικά μοντέλα



- Η Oracle Spatial υποστηρίζει δύο μοντέλα για την αναπαράσταση γεωμετρικών δεδομένων:
 - το αντικειμενο-σχεσιακό (object-relational model)
 - πίνακας με ένα πεδίο τύπου MDSYS.SDO_GEOMETRY
 - μια εγγραφή για κάθε γεωμετρική οντότητα
 - το σχεσιακό (relational model)
 - πίνακας με προκαθορισμένα αριθμητικά πεδία (τύπου NUMBER)
 - μια ή περισσότερες εγγραφές για κάθε γεωμετρική οντότητα.
- και τα δύο μοντέλα υπακούουν στις προδιαγραφές ODBC/SQL του OpenGIS Consortium για χωρικές οντότητες

73

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σύγκριση μοντέλων



- Το αντικειμενο-σχεσιακό μοντέλο
 - επικρατεί από την έκδοση 9i κι έπειτα
 - πλήρης υποστήριξη replication και κατανεμημένων αντικειμένων
- Το σχεσιακό μοντέλο
 - θεωρείται πλέον παρωχημένο
 - διατηρείται για λόγους συμβατότητας

74

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Πρωταρχικοί γεωμετρικοί τύποι

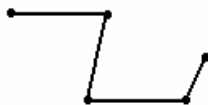


- Σχεσιακό μοντέλο

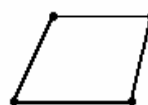
Point



Line String



Polygon



75

ΠΑ.ΠΕΙ. – Γιάννης Θεοδορίδης

Πρωταρχικοί γεωμετρικοί τύποι



- Αντικειμενο-σχεσιακό μοντέλο (πρόσθετοι τύποι)

Arc Line String



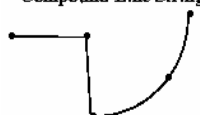
Arc Polygon



Compound Polygon



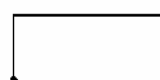
Compound Line String



Circle



Rectangle



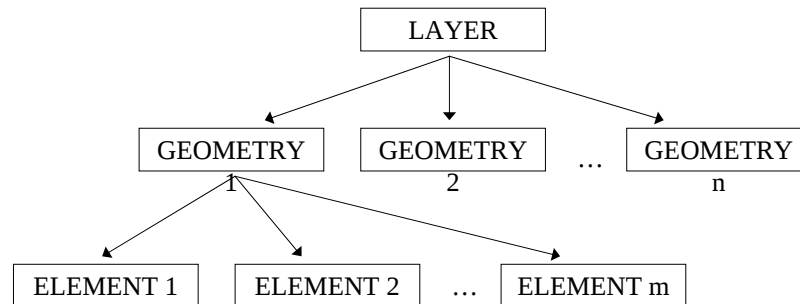
76

ΠΑ.ΠΕΙ. – Γιάννης Θεοδορίδης

Μοντέλο δεδομένων



- Ιεραρχική δομή (επίπεδα, γεωμετρίες, στοιχεία)



77

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



- Στοιχείο (Element)
 - βασική δομική μονάδα της γεωμετρίας
 - σημεία, πολυγραμμές (ακολουθίες ευθ. τμημάτων), πολύγωνα
 - οι συντεταγμένες των στοιχείων καταχωρούνται ως ζεύγος $\langle X, Y \rangle$
 - πολλαπλά στοιχεία μπορεί να αναφέρονται στην ίδια γεωμετρική οντότητα

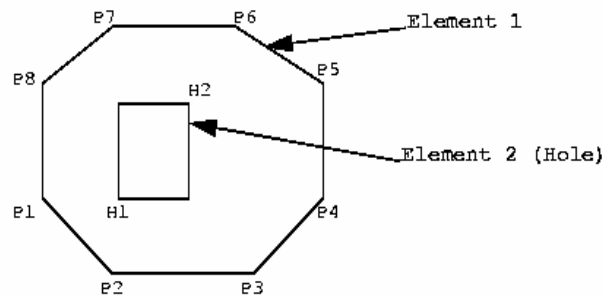
78

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



Geometry OBJ_1:



- ο εξωτερικός και ο εσωτερικός δακτύλιος ενός πολυγώνου με σπές
- δύο διακριτά στοιχεία που σχηματίζουν μαζί το σύνθετο πολύγωνο

79

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



- Ορισμός στοιχείων:
 - σημείο: ένα ζεύγος συντεταγμένων
 - γραμμή: κάθε ευθύγραμμο τμήμα AB ορίζεται από δύο ζεύγη συντεταγμένων (για τα άκρα A και B)
 - πολύγωνο: κάθε ακμή του πολυγώνου ορίζεται από το ζεύγος των αντίστοιχων κορυφών
 - οι συντεταγμένες καταχωρούνται διατεταγμένες γύρω από το πολύγωνο
 - κατά τη φορά των δεικτών του ρολογιού για κάθε εσωτερικό δακτύλιο,
 - ανάστροφα για τον εξωτερικό δακτύλιο

80

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



- Γεωμετρία (Geometry)
 - η αναπαράσταση του σχήματος μιας χωρικής οντότητας σε ένα σύστημα συντεταγμένων
 - διατεταγμένη ακολουθία από πρωταρχικά στοιχεία
 - Στο σχεσιακό μοντέλο
 - κάθε γεωμετρία πρέπει να προσδιορίζεται μοναδικά από έναν κωδικό (geometry identifier, GID) για το συσχετισμό με τα υπόλοιπα γνωρίσματα (attributes)
 - Στο αντικειμενο-σχεσιακό μοντέλο
 - συμπαγής αναπαράσταση σε έναν πίνακα

81

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



- Εναλλακτικές μορφές γεωμετρίας:
 - ένα μόνο στοιχείο: στιγμιότυπο ενός από τους πρωταρχικούς γεωμετρικούς τύπους POINT, LINE, POLYGON
 - μια ομοιογενής συλλογή στοιχείων (λ.χ. MULTIPOLYGON για συστάδα νησιών),
 - μια ετερογενής συλλογή στοιχείων (λ.χ. COLLECTION για πολύγωνα και τα κεντροειδή τους)
- Στο σχεσιακό μοντέλο
 - μια σύνθετη γεωμετρία, λ.χ. ένα πολύγωνο με οπές, πρέπει να αποθηκευτεί ως ακολουθία πολυγωνικών στοιχείων
 - τα επιμέρους στοιχεία ενός τέτοιου πολυγώνου πρέπει να περιέχονται εντός του στοιχείου για τον εξωτερικό δακτύλιο.

82

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



■ Θεματικό επίπεδο (Layer)

- ετερογενής συλλογή γεωμετρικών οντοτήτων με το ίδιο σύνολο γνωρισμάτων (attributes)
- τα γεωμετρικά στοιχεία και οι σχετικοί χωρικοί δείκτες αποθηκεύονται σε πίνακες της βάσης
- Παράδειγμα:
 - Νομοί
 - Δρόμοι
 - Πόλεις



83

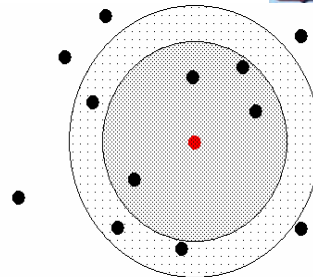
ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Μοντέλο δεδομένων



■ Ανοχή (Tolerance)

- εκτίμηση σχετικά με την ακρίβεια ή την ορθότητα των χωρικών δεδομένων
- Αν η απόσταση δύο σημείων είναι μικρότερη από την τιμή της ανοχής, τα σημεία θεωρείται ότι ταυτίζονται
- Παράμετρος ανοχής σε χωρικές συναρτήσεις κατά τη διατύπωση των ερωτημάτων
- Παράμετρος ανοχής στα μεταδεδομένα των χωρικών επιπέδων (αν θεωρείται εξαρχής γνωστή)
- Για γεωδαιτικά δεδομένα η ανοχή εκφράζεται στην αντίστοιχη μονάδα μέτρησης (συνήθως σε μέτρα)



84

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ Τύπος αντικειμένου SDO_GEOMETRY

- Το πρωτεύον κλειδί ενός πίνακα δεν μπορεί να οριστεί στο πεδίο τύπου SDO_GEOMETRY

- ένας τέτοιος πίνακας (geometry table) πρέπει να περιλαμβάνει και άλλα πεδία για τον ορισμό του κλειδιού

■ Ορισμός τύπου αντικειμένων SDO_GEOMETRY:

```
CREATE TYPE sdo_geometry AS OBJECT (
    SDO_GTYPE NUMBER,
    SDO_SRID NUMBER,
    SDO_POINT SDO_POINT_TYPE,
    SDO_ELEM_INFO MDSYS.SDO_ELEM_INFO_ARRAY,
    SDO_ORDINATES MDSYS.SDO_ORDINATE_ARRAY);
```

85

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ Αποδεκτοί γεωμετρικοί τύποι (SDO_GTYPE):

Value	Geometry Type	Description
d/00	UNKNOWN_GEOMETRY	Spatial ignores this geometry.
d/01	POINT	Geometry contains one point.
d/02	LINE or CURVE	Geometry contains one line string that can contain straight or circular arc segments, or both. (LINE and CURVE are synonymous in this context.)
d/03	POLYGON	Geometry contains one polygon with or without holes. Not 1
d/04	COLLECTION	Geometry is a heterogeneous collection of elements. Not 2 COLLECTION is a superset that includes all other types.
d/05	MULTIPOINT	Geometry has one or more points. (MULTIPOINT is a superset of POINT.)
d/06	MULTILINE or MULTICURVE	Geometry has one or more line strings. (MULTILINE and MULTICURVE are synonymous in this context, and each is a superset of both LINE and CURVE.)
d/07	MULTIPOLYGON	Geometry can have multiple, disjoint polygons (more than one exterior boundary). (MULTIPOLYGON is a superset of POLYGON.)

86

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ SDO_GTYPE

- $d = \{ 2, 3, 4 \}$ ο αριθμός των διαστάσεων
 - λ.χ., τιμή 2003 σημαίνει 2-διάστατο πολύγωνο.
- $l = \{ 0, 3, 4 \}$ η διάσταση που χρησιμοποιείται για τιμές μετρήσεων σε γεωμετρικά στοιχεία με γραμμική αναφορά (Linear Referencing System)
 - λ.χ., τιμή 2302 σημαίνει ότι η 3η διάσταση κάθε κορυφής 2-διάστατης γραμμής αντιστοιχεί στην τιμή μέτρησης (απόσταση από αφετηρία).
- Ο αριθμός των διαστάσεων αντιστοιχεί στο πλήθος των συντεταγμένων για την αναπαράσταση κάθε κορυφής
 - Αρκεί ένα ζεύγος $\langle X, Y \rangle$ για 2-διάστατα αντικείμενα

87

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ SDO_SRID

- χρησιμεύει στον προσδιορισμό συστήματος αναφοράς για τις χωρικές συντεταγμένες
 - αναγκαίο για τη σωστή ερμηνεία της γεωμετρίας
- Αν είναι NULL, οι γεωμετρικές οντότητες δεν προσδιορίζονται από κάποιο σύστημα αναφοράς
 - απλές καρτεσιανές συντεταγμένες
- Αν είναι NOT NULL, τότε πρέπει να συμπληρωθεί με την κατάλληλη τιμή από το πεδίο SRID του πίνακα των συστημάτων συντεταγμένων MDSYS.CS_SRS

88

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ SDO_POINT

- Ορίζει έναν τύπο αντικειμένου με αριθμητικά γνωρίσματα X, Y, Z
- Οι τιμές X, Y (και Z, αν συμπληρωθεί) θεωρούνται ως συντεταγμένες ενός σημειακού αντικειμένου
 - Μόνο αν τα σύνολα SDO_ELEM_INFO και SDO_ORDINATES είναι NULL και το SDO_POINT είναι NOT NULL
 - Σε κάθε άλλη περίπτωση, η τιμή του SDO_POINT αγνοείται στην Oracle Spatial
- Πολύ αποδοτικό όταν το θεματικό επίπεδο περιλαμβάνει μόνο σημειακά αντικείμενα.

89

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ SDO_ELEM_INFO

- Μεταβλητό σύνολο από τριάδες τιμών:
 - SDO_STARTING_OFFSET
 - δείχνει την θέση στο διατεταγμένο σύνολο SDO_ORDINATES όπου έχει αποθηκευτεί η πρώτη τιμή συντεταγμένων γι' αυτό το αντικείμενο
 - SDO_ETYPE
 - φανερώνει τον τύπο του αντικειμένου, λ.χ.:
 - 1003: εξωτερικός δακτύλιος πολυγώνου (ανάστροφη φορά δεικτών)
 - 2003: εσωτερικός δακτύλιος πολυγώνου (κατά τη φορά δεικτών ρολογιού)
 - SDO_INTERPRETATION
 - ερμηνεία των συντεταγμένων που αποθηκεύονται για το αντικείμενο
 - το αντικείμενο στο SDO_ETYPE μπορεί να είναι σύνθετο (compound)

90

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ SDO_ORDINATES

- μεταβλητού μήκους ακολουθία αριθμητικών τιμών (NUMBER)
- αποθηκεύει τις τιμές των συντεταγμένων που σχηματίζουν το περίγραμμα του χωρικού αντικειμένου
- αυτό το διατεταγμένο σύνολο τιμών ερμηνεύεται σε συνδυασμό με την πληροφορία στο SDO_ELEM_INFO
- Παράδειγμα:
 - ένα πολύγωνο με τέσσερις 2-διάστατες κορυφές πρέπει να δηλωθεί ως {X1,Y1, X2,Y2, X3,Y3, X4,Y4, X1,Y1}

91

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Το αντικειμενο-σχεσιακό σχήμα



■ Παράδειγμα

```
INSERT INTO regions VALUES (
  92,
  'PARK',
  MDSYS.SDO_GEOMETRY(
    2003, -- 2-dimensional polygon
    8307, -- SRID for Longitude / Latitude (WGS84 coordinate system)
    NULL,
    MDSYS.SDO_ELEM_INFO_ARRAY(1,1003,3),
    MDSYS.SDO_ORDINATE_ARRAY(23.78, 38.76, 23.89, 38.84)
  )
);
```

92

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Συστήματα χωρικών βάσεων δεδομένων



Εισαγωγή στην PostgreSQL

(εκδόσεις 7.3, 8)

93

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

PostgreSQL



- Ελεύθερα αναπτυσσόμενο λογισμικό
 - εκδόσεις για Linux, αλλά και σε Windows
- Υποστήριξη χωρικών στοιχείων:
 - Γεωμετρικοί τύποι δεδομένων
 - Χωρικοί δείκτες
 - Γεωμετρικοί τελεστές
 - Γεωμετρικές συναρτήσεις
 - Μετατροπές γεωμετρικών στοιχείων

94

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

PostgreSQL



■ Γεωμετρικοί τύποι δεδομένων

Τύπος	Αποθήκευση	Περιγραφή	Αναπαράσταση
point	16 bytes	Σημείο στο επίπεδο	(x,y)
line	32 bytes	Ευθεία γραμμή (δεν έχει υλοποιηθεί πλήρως)	((x1,y1),(x2,y2))
lseg	32 bytes	Πεπερασμένο ευθύγραμμο τμήμα	((x1,y1),(x2,y2))
box	32 bytes	Ορθογώνιο	((x1,y1),(x2,y2))
path	16+16n bytes	Κλειστό μονοπάτι (όμοιο με πολύγωνο)	((x1,y1),(x2,y2),...)
path	16+16n bytes	Ανοιχτό μονοπάτι	[(x1,y1),(x2,y2),...]
polygon	40+16n bytes	Πολύγωνο (όμοιο με κλειστό μονοπάτι)	((x1,y1),(x2,y2),...)
circle	24 bytes	Κύκλος	<(x,y),r> (κέντρο και ακτίνα)

95

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

PostgreSQL



■ Χωρικά ευρετήρια

- R-δένδρα για πολυγωνικά δεδομένα (μόνο)
- Τετραδικά δένδρα δεν υποστηρίζονται
- Γενικευμένα δένδρα αναζήτησης (GiST):
 - για δεικτοδότηση όλων των χωρικών τύπων
 - υλοποίηση των R-δένδρων ως ειδικών περιπτώσεων
 - δυνατότητα παραμετροποίησης των χωρικών λειτουργιών

96

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

PostgreSQL



■ Γεωμετρικοί τελεστές

Τελεστής	Περιγραφή	Παράδειγμα
+	Translation	box '((0,0),(1,1))' + point '(2,0,0)'
-	Translation	box '((0,0),(1,1))' - point '(2,0,0)'
*	Scaling/rotation	box '((0,0),(1,1))' * point '(2,0,0)'
/	Scaling/rotation	box '((0,0),(2,2))' / point '(2,0,0)'
#	Point or box of intersection	'((1,-1),(-1,1))' # '((1,1),(-1,-1))'
#	Number of points in path or polygon	# '((1,0),(0,1),(-1,0))'
@-@	Length or circumference	@-@ path '((0,0),(1,0))'
@@	Center	@@ circle '((0,0),10)'
##	Closest point to first operand on second operand	point '(0,0)' ## lseg '((2,0),(0,2))'
<->	Distance between	circle '((0,0),1)' <-> circle '((5,0),1)'
&&	Overlaps?	box '((0,0),(1,1))' && box '((0,0),(2,2))'
&<	Does not extend to the right of?	box '((0,0),(1,1))' &< box '((0,0),(2,2))'
&>	Does not extend to the left of?	box '((0,0),(3,3))' &> box '((0,0),(2,2))'

PostgreSQL



■ Γεωμετρικοί τελεστές (συνέχεια)

Τελεστής	Περιγραφή	Παράδειγμα
<<	Is left of?	circle '((0,0),1)' << circle '((5,0),1)'
>>	Is right of?	circle '((5,0),1)' >> circle '((0,0),1)'
<^	Is below?	circle '((0,0),1)' <^ circle '((0,5),1)'
>^	Is above?	circle '((0,5),1)' >^ circle '((0,0),1)'
?#	Intersects?	lseg '((-1,0),(1,0))' ?# box '((-2,-2),(2,2))'
?-	Is horizontal?	?- lseg '((-1,0),(1,0))'
?-	Are horizontally aligned?	point '(1,0)' ?- point '(0,0)'
?	Is vertical?	? lseg '((-1,0),(1,0))'
?	Are vertically aligned?	point '(0,1)' ? point '(0,0)'
?⊥	Is perpendicular?	lseg '((0,0),(0,1))' ?⊥ lseg '((0,0),(1,0))'
?	Are parallel?	lseg '((-1,0),(1,0))' ? lseg '((-1,2),(1,2))'
~	Contains?	circle '((0,0),2)' ~ point '(1,1)'
@	Contained in or on?	point '(1,1)' @ circle '((0,0),2)'
~=	Same as?	polygon '((0,0),(1,1))' ~= polygon '((1,1),(0,0))'

PostgreSQL



■ Γεωμετρικές συναρτήσεις

Συνάρτηση	Επιστρέφει	Περιγραφή	Παράδειγμα
area(object)	double precision	area	area(box '((0,0),(1,1)'))
box_intersect(box, box)	box	intersection box	box_intersect(box '((0,0),(1,1)') , box '((0.5,0.5),(2,2)'))
center(object)	point	center	center(box '((0,0),(1,2)'))
diameter(circle)	double precision	diameter of circle	diameter(circle '((0,0),2.0)')
height(box)	double precision	vertical size of box	height(box '((0,0),(1,1)'))
isclosed(path)	boolean	a closed path?	isclosed(path '((0,0),(1,1),(2,0)'))
isopen(path)	boolean	an open path?	isopen(path '((0,0),(1,1),(2,0)'))
length(object)	double precision	length	length(path '(((-1,0),(1,0)'))
npoints(path)	integer	number of points	npoints(path '((0,0),(1,1),(2,0)'))
npoints(polygon)	integer	number of points	npoints(polygon '((1,1),(0,0)'))
pclose(path)	path	convert path to closed	pclose(path '((0,0),(1,1),(2,0)'))
popen(path)	path	convert path to open	popen(path '((0,0),(1,1),(2,0)'))
radius(circle)	double precision	radius of circle	radius(circle '((0,0),2.0)')
width(box)	double precision	horizontal size of box	width(box '((0,0),(1,1)'))

PostgreSQL



■ Μετατροπές γεωμετρικών στοιχείων

Συνάρτηση	Επιστρέφει	Περιγραφή	Παράδειγμα
box(circle)	box	circle to box	box(circle '((0,0),2.0)')
box(point, point)	box	points to box	box(point '(0,0)', point '(1,1)')
box(polygon)	box	polygon to box	box(polygon '((0,0),(1,1),(2,0)'))
lseg(point, point)	lseg	points to line segment	lseg(point '(-1,0)', point '(1,0)')
path(polygon)	point	polygon to path	path(polygon '((0,0),(1,1),(2,0)'))
point(double precision, double precision)	point	construct point	point(23.4, -44.5)
point(box)	point	center of box	point(box '((-1,0),(1,0)'))
point(circle)	point	center of circle	point(circle '((0,0),2.0)')
point(lseg)	point	center of lseg	point(lseg '((-1,0),(1,0)'))
point(lseg, lseg)	point	intersection	point(lseg '((-1,0),(1,0)') , lseg '((-2,-2),(2,2)'))
point(polygon)	point	center of polygon	point(polygon '((0,0),(1,1),(2,0)'))
polygon(box)	polygon	box to 4-point polygon	polygon(box '(0,0),(1,1)')
polygon(circle)	polygon	circle to 12-point polygon	polygon(circle '((0,0),2.0)')
polygon(npts, circle)	polygon	circle to npts-point polygon	polygon(12, circle '((0,0),2.0)')
polygon(path)	polygon	path to polygon	polygon(path '((0,0),(1,1),(2,0)'))

πίδες

PostgreSQL – παραδείγματα ορισμού δεδομένων



```
CREATE TABLE zones(poly_id integer,
                    name varchar(30),
                    sector polygon );

INSERT INTO network.zones VALUES (1, 'PARK', '(479243,4204000,
477728,4202750, 477559,4202100, 476271,4204750)::polygon);

CREATE TABLE locations(point_id integer,
                        name varchar(30),
                        pos point);

INSERT INTO locations VALUES (52, 'SYNTAGMA', '(476600,4202800)::point);
```

101

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

PostgreSQL – παραδείγματα αναζήτησης δεδομένων



```
SELECT point_id,(pos<->Point '(475750,4201500)') as distance
FROM locations
WHERE (pos <-> Point '(475750,4201500)' )<=200
```

Αναζήτηση βάσει
απόστασης από
δεδομένο σημείο

```
SELECT point_id, name
FROM locations
WHERE (pos @ polygon '(479243,4204000,476271,4204750)')=TRUE
```

```
SELECT point_id, name
FROM locations
WHERE (pos @ (SELECT sector FROM zones WHERE name='PARK'))
```

Αναζήτηση εντός
δεδομένης περιοχής

102

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Περιεχόμενα



- Εισαγωγή
 - GIS - παραδείγματα λογισμικού και εφαρμογών, αρχιτεκτονική
- Spatial DBMS
 - Εισαγωγή – Τύποι χωρικών δεδομένων, Συστατικά ενός SDBMS
 - Σχεδίαση χωρικών ΒΔ – Εννοιολογικός, Λογικός, Φυσικός σχεδιασμός
 - Εμπορικά SDBMS – Oracle Spatial, PostgreSQL
- Σύνοψη, ερευνητικά θέματα

103

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Σύνοψη



- Ένα SDBMS χρησιμεύει για την αποθήκευση και διαχείρισης χωρικών δεδομένων για χρήση από GIS και άλλες εφαρμογές
- Τα συστατικά ενός SDBMS περιλαμβάνουν
 - Μοντέλο δεδομένων, τύπους δεδομένων και τελεστές για χωρικά δεδομένα,
 - Γλώσσες ερωτήσεων, μηχανισμούς επεξεργασίας και βελτιστοποίησης
- Τα προϊόντα SDBMS αναπτύσσεται ταχύτατα, λόγω και της ραγδαίας ανάπτυξης των προϊόντων GIS
- Μεγάλη ποικιλία εφαρμογών και χρηστών
- Κύριες περιοχές τρέχουσας ανάπτυξης
 - Διαδίκτυο (βλ. Google Maps κ.α.)
 - Κινητές συσκευές (PDA, PNA, smartphones)

104

ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

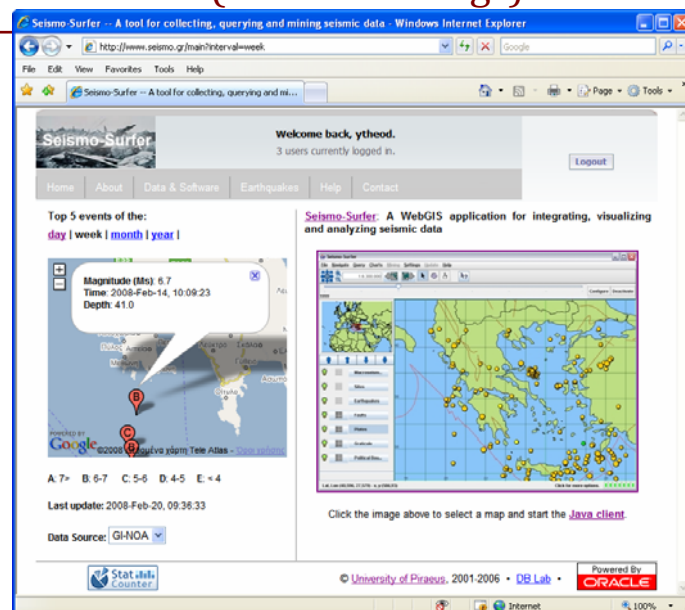


- Τέλος, κάποιες εφαρμογές SDBMS που αναπτύσσουμε στο InfoLab ...

105

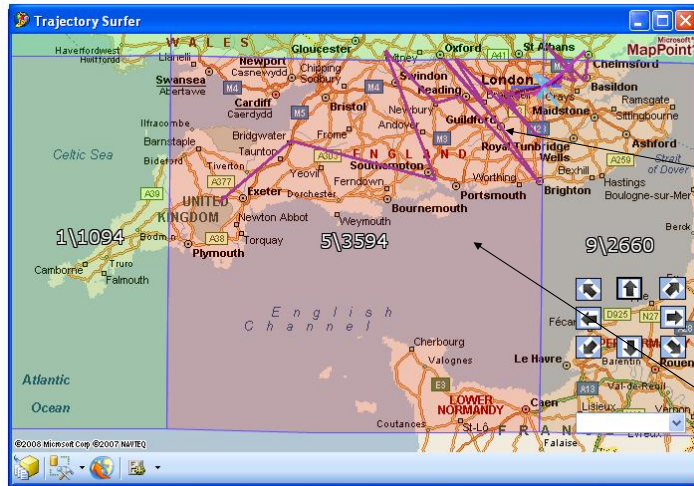
ΠΑ.ΠΕΙ. – Γιάννης Θεοδωρίδης

Seismo-Surfer (www.seismo.gr)



– Γιάννης Θεοδωρίδης

T-Warehouse



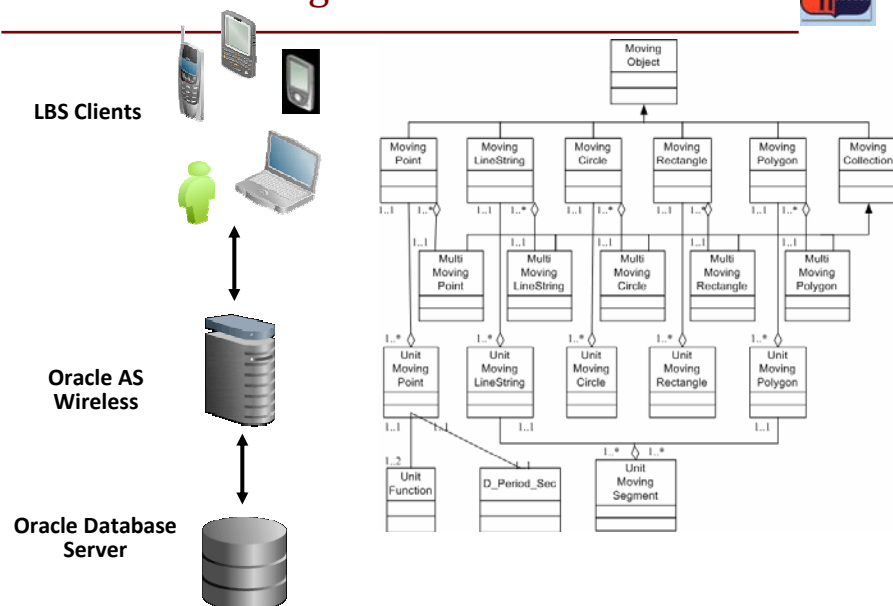
Αντιπροσωπευτικές τροχιές

Κελί πλέγματος με ID =5 και συνολικό αριθμό τροχιών 3594.

Μενού επιλογών:
OLAP ανάλυση δεδομένων
Εξόρυξη γνώσης (συσταδοποίηση)
Ρυθμίσεις προγράμματος

ΠΑ.ΠΕΙ. – Γιάννης Θεοδοωρίδης

Hermes MOD engine



108

ΠΑ.ΠΕΙ. – Γιάννης Θεοδοωρίδης

Προτεινόμενη βιβλιογραφία



- A. Guttman. R-trees: a dynamic index structure for spatial searching. Proc. ACM SIGMOD 1984.
- R.H. Güting. An Introduction to Spatial Database Systems. VLDB Journal, 3 (1994), 357-399.
- V. Gaede & O. Guenther. Multidimensional access methods. ACM Computing Surveys, 30 (1998), 170–231.
- S. Shekhar et al. Navigation systems: a spatial database perspective. Chapter 3 in Location-Based Services, J. Schiller, A. Voisard (eds.). Morgan Kaufmann, 2004.