Spatial Indexing and Cars

Dr John Fawcett
Indexing

• Index = additional data structure

• Contains pointers to objects that are often stored in databases

• Index entries are arranged/grouped/sorted to speed up insertion/deletion/searching

• *Spatial* indexes speed up operations involving geometric spaces in 1, 2, or 3 dimensions
Nature of Spatial Datasets

• Type of data
  – point data (e.g. spot-height samples)
  – continuous spatial properties (e.g. contours)
  – discrete spatial data/regions (e.g. RCL databases)

• Use of data
  – location as place/context for activity
  – spatial relationships/relative positioning

Geographic Information Systems

• GIS databases store data and construct indexes based on spatial attributes.
• Market leaders:
  – AutoCAD—architecture and scientific modelling/drawing/rendering
  – Oracle *Spatial*—models co-ordinate data and uncertainty/error. 2D: Cartesian and Lat/Long. Fuzzy comparison operators.
Spatial Indexing

• Typical situation:
  – Static data (e.g. road map) but mobile focus
  – Searches are “range queries”

• Region-Trees (R-Trees)
  – Like B⁺-Trees
R-Trees [1]

- Shapes of arbitrary objects too complex for efficient processing
  - Use Bounding Boxes or “Minimum Bounding Rectangles” (MBRs)

- Leaf node of R-Tree:
  - Contains tuples: (MBR, pObject)

- Internal node of R-Tree:
  - Contains tuples: (MBR, pChildNode)
Example R-Tree Leaf Node

Tuples:
1. (          , objA)
1. (          , objB)
1. (          , objC)
Example R-Tree

Root node tuples:

1. (, child-1)
2. (, child-2)
(m,M) R-Trees [2]

• Leaf and internal nodes contain between m and M entries where $2 \leq m \leq \text{ceil}(M/2)$
  – Similar storage flexibility to B-Trees
  – There’s a special case...

• Tree with only one node (the root) may house fewer than m items.

• Terminology: as items are added and removed, a node might overflow or underflow and require splitting or merging
R-Trees [3]

• Tree is height-balanced

• Bounding boxes of internal nodes are:
  – hierarchically nested
  – permitted to overlap
  – found to offer max searching efficiency when:
    • coverage is minimised
    • overlap is minimised
Another R-Tree

Level 0 ➔

Level 1 ➔

All objects at depth 2
Constructing an R-Tree

• Build it dynamically...
  – Insert objects in order of arrival into system
  – Need a splitting strategy

• ...or build it statically
  – Remember objects as they appear
  – Then construct an optimally-divided tree
  – “Packing”
Algorithms on R-Trees

• Insertion, splitting strategies
  [demonstration given in lectures]
• Searching
  [demonstration given in lectures]

• Exercises: work out how to…
  – Delete an item from R-Tree
  – Merge two R-Trees
Optimising R-Trees: m

• Consider large m...
  – At least m of M entries used—few unused positions in data structure
  – Splitting a node can be expensive

• Consider small m...
  – (m,M) gives wider population range—fewer over/under flows expected

• Which best suits static/changing datasets?
Variants

- R+-Trees (1987): hierarchically nested containers, overlap not permitted, objects indexed in each container spanning their position
- R*-Trees (1990): hierarchically nested containers, overlap permitted, objects indexed in each container spanning their position
- TV-Tree (1994), X-Tree (1997), …
Special-Purpose Variants

- SMR-Trees: range queries are infrequently demanded
- QSF-Trees: dataset very big; not even the index fits in main memory. Arranged to minimise page faults when descending levels of the tree.
- What if my dataset is dynamic?
  - Predictable changes can be pre-optimised!
  - See Kollios, Gunnopulos and Tsotras [96]
Fun Stuff

- Install in your car: GPS, a PC, sensors, GSM/GPRS/3G phone, Wireless LAN, Bluetooth, USB hub in dashboard, CanBUS adaptor, EMC link, ByteFlight host, …
- Naïve to think of the PC as just a means of collecting sensor data
- Instead, consider the car to be a piece of the Internet…
Navstar GPS in 1 slide

- All-weather, round-the-clock, timing and ranging to an unlimited number of simultaneous users with anti-jamming
- Block 1 SVs formed Demo System. Block 2 achieved FOC. 24 SVs in space segment+spares. 3 orbital planes.
- SV=3-Axis stabilized, nadir pointing using reaction wheels. Dual solar arrays 400W+ NiCd batteries. S-Band (SGLS) communications for control and telemetry. UHF cross-link between spacecraft. Hydrazine propulsion system. *Translation to English*=>*flying atomic clock.*
- Fix complex—corrections for relativistic velocity, photon pressure, local G anomalies, ionosphere diffraction...
Portable?