SPATIAL DATABASE ISSUES, SUCCESSES, AND CHALLENGES

HANAN SAMET

COMPUTER SCIENCE DEPARTMENT AND
CENTER FOR AUTOMATION RESEARCH AND
INSTITUTE FOR ADVANCED COMPUTER STUDIES
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742-3411 USA

E-MAIL: HJS@CS.UMD.EDU

SUCCESSES

1. GPS
2. Google Maps/Earth and Microsoft Virtual Earth have made spatial data a first class citizen in databases and applications
    - General mapping
    - Routing
3. Profusion of location-based services
4. Iphone and handheld computing devices for mapping applications
    - Traffic congestion
    - Parkme and other location-based applications in AppStore
5. Spatial spreadsheets like the SAND Internet Spatial Browser

ISSUES IN SPATIAL DATABASES

1. Representation
    - bounding boxes versus disjoint decomposition
2. How are spatial integrity constraints captured and assured?
    - edges of a polygon link to form a complete object
    - line segments do not intersect except at vertices
    - contour lines should not cross
3. Interaction with the relational model
    - spatial operations don't fit into SQL
        a. buffer
        b. nearest to ...
        c. others ...
    - difficult to capture hierarchy of complex objects (e.g., nested definition)
4. Spatial input is visual
    - need a graphical query language

5. Spatial output is visual
    - unlike conventional databases, once operation is complete, want to browse entire output together rather than one tuple at-a-time
    - don't want to wait for operation to complete before output
        a. partial visual output is preferable
            - e.g., incremental spatial join and nearest neighbor
        b. multiresolution output is attractive
6. Functionality
    - determining what people really want to do!
7. Performance
    - not enough to just measure the execution time of an operation
    - time to load a spatial index and build a spatially-indexed output is important
    - sequence of spatial operations as in a spatial spreadsheet
        a. output of one operation serves as input to another
            - e.g., cascaded spatial join
        b. spatial join yields locations of objects and not just the object pairs

INTERESTING AREAS (NOT NECESSARILY "HOT"!)

1. Spatial data mining
2. Integration of spatial and nonspatial data
3. Identification set of possible spatial queries
4. Explore interaction techniques with spatial data such as gesturing
5. Accurate geocoding
6. Exploiting new computing architectures for spatial applications
    - GPU and parallel architectures
    - Distributed problem-solving architectures such as map-reduce and dryad
7. Spatio-textual reference disambiguation
    - E.g., which London is it?
8. Accurate distance computation
    - Spherical geometry
    - Network distances instead of "as the crow flies"

CHALLENGES:

1. Incorporation of geometry into database queries without user being aware of it!
   - find geometric analogs of conventional database operations (e.g., ranking semi-join yields discrete Voronoi diagram)
   - extension of browser concept to permit more general browsing units based on connectivity (e.g., shortest path), frequency, etc.
2. Spatial query optimization
   - different query execution plans
   - use spatial selectivity factors to choose between them
3. Graphical query specification instead of SQL
4. Incorporation of time-varying data
   - how to represent rates?
5. Incorporation of imagery
6. Develop spatial indices that support both location-based (“what is at X”?) and feature-based queries (“where is Y”?)
7. Incorporate rendering attributes into database objects or relations
   - queries based on the rendering attributes
   - Ex: find all red regions
   - query by content (e.g., image databases)
8. GIS on the Web and distributed data and algorithms
9. Knowledge discovery
10. Interoperability