HTM2: Spatial Toolkit for the Virtual Observatory

Gyorgy Fekete[1], Alex Szalay[1], Jim Gray[2].
1 Johns Hopkins University, 2 Microsoft Research

GOAL: Efficient Management of Spherically Distributed Spatial Information

Catalogs contain hundreds of millions of objects
Spatial correlation of pairs, triples, etc of objects and regions of interest requires geometric computation

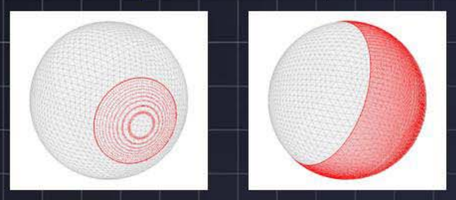
HTM scheme provides a coarse and efficient way of discovering potential spatial matches (or mismatches)

METHOD:
: Quantize the sphere into trixels
: Decompose any region (shape) into primitive building blocks
Associate shape primitives to trixel sets
Formulate a spatial query in terms of Boolean operations on regions

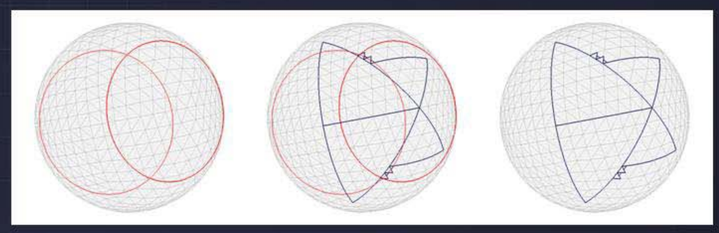
CRITERIA:
Rapid computation of trixel sets covering a region
Emphasis on minimizing the number of trixels
All of this transparent to the user
HTM2 (HTM Version 2 software) has significant performance improvements over previous versions

Regions, Shapes
Circles
Triangles
Rectangles
Bands
Polygons

Basic Shape *Constraint* = Spherical Cap



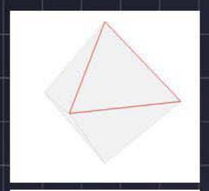
Intersection of 2 circles



Intersection of 2 Constraints

32	62	141	253
9120	9121	16144	16146
36494	36608	64589	65408

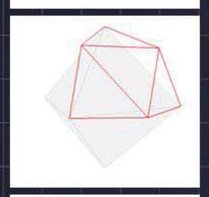
Start with a platonic solid, such as the Octahedron with all vertices at unit distance from the origin



N1

Each face (trixel) has an "address" in the form:
N0, N1, N2, N3, S0, S1, S2, S3

Push the midpoints of edges to unit distance from the origin

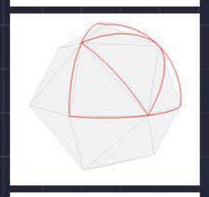


Symbolic addresses are formed by labeling each child trixel with a single digit from 0,1,2,3

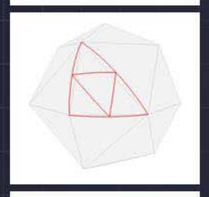
Connect the new points to yield four smaller triangles

N11
N13
N12 N10

The edges are actually great circle segments…

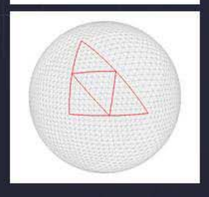


Repeat the process for each spherical triangle



N122
N123
N120 N121

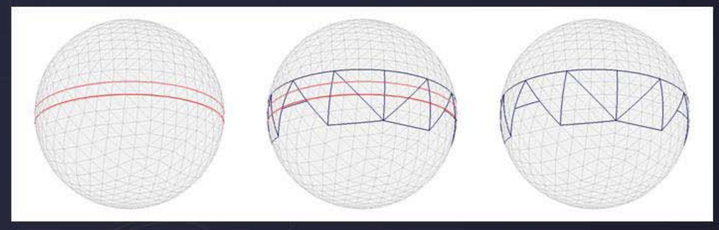
Continue until desired resolution is reached



An address can be represented as an integer called HTMID

```
N1 = 13
N12 = 54
N120 = 216, N121 = 217…
Large numbers represent small areas
```

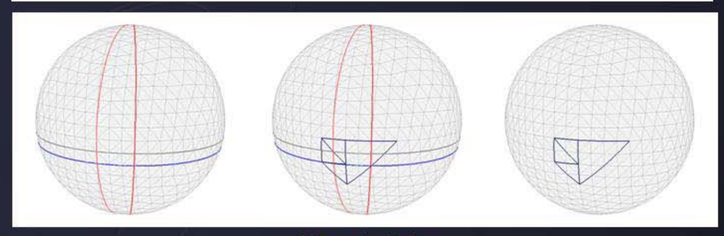
Band



Band = Intersection of 2 Constraints

```
128 130 131 136 137 139
144 146 147 152 153 155
160 162 163 168 169 171
176 178 179 184 185 187
564 628 692 756
```

Rectangle



Rectangle =
Intersection of band and 2 Constraints =
Intersection of 4 Constraints

```
144 544 546 547 144 544
546 547
```

Constraints | **Constraints and HTMIDs** | **HTMIDs** | **HTMID numbers**

Any single connected shape is an intersection of a finite number of *constraints*. This is called a *convex*. Any shape can be represented by a finite union of convexes.

The hierarchical triangular mesh (HTM) is a discrete foundation for describing location, size and shape on the (celestial) sphere. Indices derived from HTM descriptors are used in a relational database for managing spatial information. Algorithms are implemented as extended stored procedures accessible to the database engine, Microsoft SQL Server 2000. A language to support describing shapes processed by functions added to t-sql. These functions provide adequate encapsulation of the HTM methods, so that users need not be aware of the workings of HTM algorithms. Familiar shapes, like rectangles, circles, bands, are transformed into an internal normal form based on the union of *convexes*, which, in turn are intersections of so called *constraints* (caps). In a computer program, the region is an object that contains the HTMIDs of the trixels that represent the region. These are generated by the library from descriptions in terms of familiar shapes, such as circles, rectangles, arbitrary polygons. If a user needs to know whether an observation is outside of a region of interest, a simple call to the HTM object with the coordinates of the observation provides the answer.

HTM objects can be combined with set-theoretical operations

THE US NATIONAL VIRTUAL OBSERVATORY

Supported by the National Science Foundation's Information Technology Research Program
under Cooperative Agreement AST0122449 with The Johns Hopkins University