Notes on Shapely

Ivan Marin

September 13, 2012

The objective of this notebook is to define all geometric operations used on GIS, both in spatially enabled databases and on code that needs to use this definitions. These definitions are mostly from [?, ?].

1 OpenGEOS Simple Features Access

1.1 Definitions

1.1.1 Coordinate System

Coordinate System is a set o mathematical rules for specifying how coordinates are to be assigned to each point.

1.1.2 Coordinate Reference System

Coordinate Reference System is a coordinate system that is related to the real world by a datum. Same thing as Spatial Reference System [?].

1.2 Simple Feature Access

Simple Feature Access specifies a common storage model of geographical data, using both well-known text or binary formats [?, ?, ?].

- [?] is also named ISO 19125

- Simple Feature Geometry is an object model, represented by UML, *Unified Modeling Language*

- Each geometric object has a Spatial Reference System (UTM, etc) associated

- Basic Geometry class has Point, Curve, Surface and GeometryCollection subclasses

- Extended geometry model has MultiPoint, MultiLineString and MultiPolygon for collections of Points, LineStrings and Polygons, plus MultiCurve and MultiSurfaces

1.3 WKT

WKT, *well-known text*, is a markup language to represent vector geometry on a map, representing the spatial objects in spatial reference systems and its transformations, also defined by OGC [**?**].

1.4 Abstract Class Geometry

- Geometry is a root class.

- Geometry is not instantiable

- $R^2 : (x, y)$

- $R^3 : (x, y, z)$ or (x, y, m)

- $R^4 : (x, y, z, m)$

- All points are associated with a coordinate reference system

- All points in the same Geometry are associated with the same coordinate reference system, either directly or through its containing Geometry

- the coordinate m represents a measurement

- All Geometries are topologically closed - the boundaries are points included in the coordinate sets.

1.4.1 Methods that a Geometry has:

The representations on this section and the following is Method(ArgumentType):ReturnType - Description.

- Dimension(): Integer - $R^1 = 1$, $R^2 = 2$, etc

- GeometryType(): Geometry - type of Geometry, string

- SRID(): Integer - ID of the Spatial Reference System, a pointer (index) in a list of reference systems stored somewhere.

- Envelope(): Geometry - the minimum bounding box, as two positions, one containing the MINX, MINY and other as MAXX, MAXY.

- AsText(): String - Exports this geometric object to a Well-known Text Representation.

- AsBinary(): Binary - Exports this geometric object to a Well-known Binary Representation.

- IsEmpty(): Integer - if returns TRUE (int 1), is an empty set in the coordinate space.

- IsSimple(): Integer - returns True (int 1) if there are no self intersections or self tangency.

- Is3D(): Integer - if it has z points

- IsMeasured(): Integer - if the geometry includes m measure points.

- Boundary(): Geometry - returns the boundary geometry.

1.4.2   Methods for testing spatial relations in a Geometry:

- Equals(otherGeometry): Integer - test two geometries, returns 1 if they are equal.

- Disjoint(otherGeometry): Integer - returns 1 if spatially disjoint. Tests if the two geometries does not overlap, touches or are within.

- Intersects(otherGeometry): Integer - returns 1 if this intersect another geometry.

- Touches(otherGeometry): Integer - returns 1 if this geometry touches another one.

- Crosses(otherGeometry): Integer - returns 1 if this geometry crosses another one.

- Within(otherGeometry): Integer - returns 1 if this geometry is inside another one.

- Contains(otherGeometry): Integer - returns 1 if this geometry contains other geometry.

- Overlaps(otherGeometry): Integer - returns 1 if this geometry overlaps other geometry.

- Relate(otherGeometry): hum, not clear

- LocateAlong(double): not clear, need to define measurement

- LocateBetween(double) same thing as above

1.4.3   Methods that support spatial analysis

These methods depend on the coordinate representations and its precision and linear interpolation.

- Distance(otherGeometry): returns a double with the shortest distance between the two geometries (this and otherG) in the spatial reference system of this geometry.

- Buffer(distance: double): Geometry - returns a geometry with all the points with the same distance or less, on the spatial reference system of this geometry.

- ConvexHull():Geometry - Geometry returns a geometric convex hull of this geometry.

- Intersection (otherGeometry): Geometry - returns a geometry with the point set intersection of this geometric object with otherGeometry

- Union(otherGeometry): Geometry - returns a geometry with the point set of the union of this geometry with otherGeometry.

- Difference (otherGeometry):Geometry - returns a geometry with the difference point set between this geometry and otherGeometry.

- SymDifference(otherGeometry):Geometry - returns a geometry with the symmetric difference between this and otherGeometry

1.4.4 The coordinate z or m

A Point may include z or m. In GIS, z may be the height above or below sea level. m is a measure from some reference point. Note that usually the z or m coordinates are *not* used for calculations of distance, length, etc, that happens in 2D, so the objects are projected on the (x, y) plane, on "shadow" style.

1.5 GeometryCollection

A GeometryCollection is a geometric object that is a collection of some number of geometric objects. All geometric objects ina GeometryCollection shall be in the same Spatial Reference System.

1.5.1 Methods

- NumGeometries(): Integer - Number of Geometries present.

- GeometryN(N: integer): Geometry - returns the Geometry number N.

2 Geometric Objects Definitions

2.1 Point

Methods: X(), Y(), Z(), M()

2.2 MultiPoint

Collection of points not connected.

2.3 Curve

1D, stored as sequence of points, with some interpolation algorithm. Is simple if it does not pass through the same Point twice, with exception of the endpoints. Methods: Length, StartPoint, EndPoint, IsClosed, IsRing. A Curve is a Ring when it's simple and closed.

2.4 LineString, Line, LinearRing

A LineString is a Curve with linear interpolation between Points. Each consecutive pair of Points defines a Line Segment. A Line is a LineString with exactly 2 Points. A LinearRing is a LineString that is closed and simple. Has NumPoints and PointN as methods.

2.5 MultiCurve

A MultiCurve is a collection of Curves. It is not instantiable, only defines methods. Is simple if all Curves are simple, and intersections only happens on the boundary of both geometries.

2.6 MultiLineString

Collection of LineStrings.

2.7 Surface

A Surface is a 2D geometric object, a patch associated with one exterior boundary. Stitching together surfaces generates a polyhedron. Two subclasses: Polygon and PolyhedralSurface.

2.8 Polygon, Triangle

A Polygon is a simple surface that is planar. A Triangle is a polygon with 3 distinct, non-collinear vertices and no interior boundary. Polygons are topologically closed, the boundary of a polygon are a set of LinearRings to make the exterior and interior boundaries, and two Rings in the boundary can intersect only on a tangent Point. The interior of a Polygon is a connected point set.
  Methods:

- ExteriorRing (), returns a LineString with the exterior ring of this Polygon.

- NumInteriorRing() returns the number of interior rings in this Polygon

- InteriorRingN(N): LineString returns the N interior Ring of this Polygon as LineString.

2.9 PolyhedralSurface

Composed by a continuous collection of polygons, that share a common boundary segment. Each polygon that touchs the interface must be represented by a finite number of LineStrings. Note that the same top orientation must be the same.

2.9.1 Methods

NumPatches, PatchN, BoundingPolygons, IsClosed.

2.10 MultiSurface

GeometryCollection of Surfaces, all with the same coordinate reference system. The geometric interior of the surfaces must not intersect each other, except the boundary. It is instatiable.

2.10.1 Methods

Area, Centroid and PointOnSurface.

2.11 MultiPolygon

A MultiPolygon is a MultiSurface whose elements are Polygons. Some conditions: The interior of 2 Polygons in the MultiPolygon must not intersect. The boundaries of 2 Polygons may not cross and can only touch in a finite number of Points. Also, it is topologically closed. The boundary of the Polygons are LineStrings (Curves).

3 Shapely

Shapely [?] is a Python Package for manipulation of planar features importing the GEOS [?] via Python ctypes. (Note that GEOS is itself a port of the Java Topology Suite to C++). Shapely is the geometry of PostGIS, the spatial extension for the PostgreSQL RDBMS.