Chapter 2
Some Philosophical Issues Regarding Geometric Modeling for Geographic Information and Knowledge Systems

Robert Laurini

Abstract It is common to state the importance of geometry in geographic information systems. But with the advent of the knowledge society, it is important to revisit some philosophical aspects that were traditionally the backbone of GIS. Indeed, the necessity to build robust systems for automatic geographic reasoning implies that several issues must be reexamined, especially due to the existence of new types of sensors which continuously measure some phenomena under interest: two sensors which will measure any phenomenon will give values a little bit different for various reasons. And we have to integrate those aspects. Now, with the appearance of new systems based on geographic knowledge, mathematic modeling of reality is again in the critical path of research. In this paper, we will examine rapidly the philosophical background of the common modeling used in GIS and try to propose new directions especially in the vision of requirements for geographic knowledge systems.

Keywords GIS philosophy · Computational geometry · Query processing · Geographic knowledge

Introduction

From a conceptual point of view, the use of mathematics in geographic information systems is fundamental for several reasons. Etymologically speaking, the word geometry means, in Ancient Greek γεωμετρία, the measurement of the Earth, or the measurement of terrains. It could be traced to Babylonians and Egyptians for surveying, two millennia BC. Remember that in geography (γεωγραφία), the Greek work γραφία means both drawing and writing about the Earth.

In other words, geometry can be considered as a basis and an ancestor of Geographic Information Systems. But, now with the use of information technologies

R. Laurini (✉)
KSI, Overland Park, KS, USA
e-mail: roberto.laurini@gmail.com

University of Lyon, Lyon, France

T. Tambassi (ed.), *The Philosophy of GIS*, Springer Geography,
https://doi.org/10.1007/978-3-030-16829-2_2

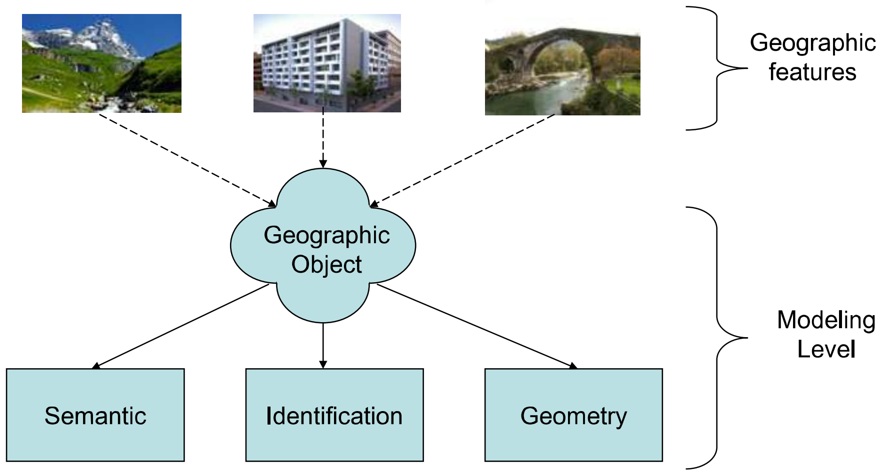


Fig. 2.1 The various points of view for modeling geographic features

and especially of knowledge engineering, it is important to revisit the mathematical backgrounds of geoprocessing.

In a previous book (Laurini 2017a), I have shown that any geographic feature can be modeled according to three points of views (Fig. 2.1):

- A **semantic point of view**, that is, the nature of the feature (road, river, mountain, etc.) which can be categorized or even subcategorized; the role of geographic ontologies is to offer adequate and relevant categorizations; do not forget that those categorizations have different cultural backgrounds leading to various categorizations in various languages;
- An **identification point of view**, that is, the name of the feature (Germany, Lady Liberty, Eiffel Tower, California, etc.); again, some linguistic aspects can be considered since features can have different names in different languages; for instance, the city of “Venice”, Italy, is also known as “Venezia”, “Venise”, and “Venedig”, respectively, in Italian, French, and German;
- A **geometric point of view** corresponding to the shape and location of the feature.

In this chapter, only the philosophical issues of the geometric point of view will be developed, more precisely, will be examined the mathematical backgrounds for modeling the characteristics of geographic features and their relationships. Then, the importance of spatial analysis will be developed leading to automatic geographic reasoning.

About Geometric Modeling of Geographic Features

It is common to state that geometric objects can be modeled at 0D (point), 1D (lines), 2D (areas) and 3D (volumes), and we can add the temporal dimension when necessary. But, is anybody able to show 0D and even 1D objects on the Earth? As far as I know, we can only mention the North and the South poles as points. Regarding lines, the so-called linear objects as roads and rivers have some width, leading to consider them as areas. However, concerning geodesy, equator, parallels, and meridians are theoretical lines which have no visible reality: they could be considered as geographic objects, but not as geographic features.

Remember that before the expression "GIS", the specialists were speaking about "Computer-Aided Cartography." The idea was to easily generate new maps by changing scales and colors. But the data were essentially coming from digitizing existing planar maps. Then, some other acquisition devices were created and used. But the original background is still on planar maps, planar plotters, and planar screens. In other words, the world looks planar. After, it becomes apparent that the major interest was not only to generate maps, but to store geographic information in the so-called GIS.

As for practical problems, for instance, at urban level, it is acceptable to consider a flat world. But what is the limit?

Earth Rotundity

The Earth is not a sphere; indeed, due to centrifugal force, it must be considered as an ellipsoid whose larger radius is located at the Equator. Now it is modeled as a geoid.

Consider now a projection plane tangent to the Earth at point *C*, and a point *P* in the Earth. There are several methods to project a spherical point onto a plane. Figure 2.2 shows two examples of projections, *T* as the intersection of an Earth radius and the projection plane, and *S* as an orthogonal projection. It is easy to see that the position of *T* is given, respectively, by $Rtg(\theta)$ and of *S* by $Rsin(\theta)$ in which *R* is the radius of the Earth. Remember that the real distance of *P* to *C* is $R\theta$. By doing so, we commit an error, respectively $\varepsilon_T = |R\theta - Rtg(\theta)|$ and or $\varepsilon_S = |R\theta - Rsin(\theta)|$. Using the first terms of Taylor series, namely, $tg(\theta) = \theta - \theta^3/3 + 2\theta^5/15 + \cdots$ and $sin(\theta) = \theta - \theta^3/6 + \theta^5/120 + \cdots$ we can compute an approximation of the errors; or conversely, by accepting any error level, we can compute the size of a place in which the planar assumption is acceptable. So, $\varepsilon_T = R\theta^3/3$ and $\varepsilon_S = R\theta^3/6$ and consequently $\theta = \sqrt[3]{3\varepsilon_T/R}$ or $\theta = \sqrt[3]{6\varepsilon_S/R}$.

Now, based on those models, let's consider a squared place and let's, for instance, tolerate an error of 1 cm. Respectively, we obtain 20 and 25 km wide places; and for 1 mm, we get 9.4 km and 11.8 km.

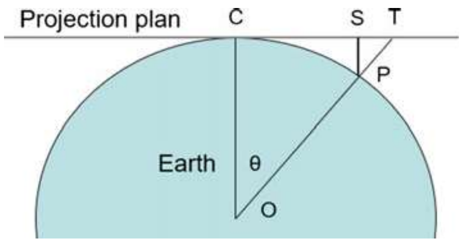


Fig. 2.2 Example of some projections

As a conclusion, we can state that if we consider a town or city with such a size, and accept those error limits, we can consider a flat world for this city or region. But, outside, Earth rotundity must be taken into account. Of course, based on other models of projections, the results can be a little bit different.

Regarding GIS, if the jurisdiction of the owner is small enough, the planar assumption is valid.

A second aspect is the origin of the measurement unit. Historically speaking, length measurement seems to be linked to the human body. A traditional tale tells the story of Henry I (1100–1135) who decreed that the yard should be "the distance from the tip of the King's nose to the end of his outstretched thumb".[1] So, this unit varied in time. To avoid such problems, during the eighteenth century, some people argue to have a more rationalistic definition , for instance, based on the Earth. During the French Revolution, in 1791, the French National Assembly decided in favor of a standard that would be one ten-millionth part of a quarter of the earth's circumference: the well-known meter unit was thus created. But now, some more precise definition is standardized, based on physics. Finally, from a philosophical point of view, the origin of length measurement was based on a geographic reasoning.

Characteristics of Geographic Features

According to Prolegomenon #1 (3D +T objects): "*All existing objects are tridimensional and can have temporal evolution; lower dimensions (0D, 1D and 2D) are only used for modeling (in databases) and visualization (in cartography)*" (Laurini 2017a). Indeed, rivers can change their bed, mountains can have earth slides, continents move (continental drift), roads can be enlarged, buildings can be demolished, and so on. As a consequence, their mathematical model must integrate also a temporal dimension. But in this chapter, this important problem will not be addressed.

[1] http://www.npl.co.uk/educate-explore/factsheets/history-of-length-measurement/.

Euclidean and Spherical Geometry

Remember that what we usually call 2D geometry or rather Euclidean geometry is not a true 2D geometry since the eyes are located in the third dimension. Suppose your eye is really located in the 2D plane, everything will be modeled by segments, whether it is a line or an area. As a consequence, the objects are no more distinguishable. As previously shown, Euclidean planar geometry can be applied in smaller zone, whereas spherical geometry in larger places, perhaps with the help of projection.

Remember that, by definition, cadastral data only use *x* and *y* coordinates, never the elevation. In other, it means that the cadastral surface of a sloping terrain is not the soil surface, but its projection unto a plane, say the horizontal surface. For instance, a cliff has a horizontal surface close to zero, whereas the vertical surface can be important. But, now, some countries intend to build 3D cadasters.

About Points, Lines, and Areas in Our World

At the beginning of geoprocessing, several models of polygons were in competition; among them, some were based on a set of points and other on a set of segments. In this regard, the role of SORSA (Segment-Oriented Referencing System Association, then Spatially Oriented Referencing System Association) was important in the 70s by promoting the segment-oriented approach essentially because it was a more efficient model to deal with consistent tessellations. Alas, the object orientation mode in the 80s imposes a definition of geographic object as delimited by a set of points, and it was the basis of the OGC geographic model (See Fig. 2.3). But this approach leads to difficulties to tessellations and overall to secure for their consistency.

As previously told, in the nature, points and lines do not exist. In a previous work, I have introduced the concept of ribbons (Laurini 2014) which can be defined as a line with a width. Remember that in mathematics, lines have no width, as ribbons can be seemed as a good starting point for modeling roads, rivers, etc. But rivers and roads can have various widths. So, several types of ribbons must be defined, rectangular ribbons, curved ribbons, and loose ribbons as given Fig. 2.4 with an application in urban context in Fig. 2.5.

Crisp Boundaries and Indeterminate Boundaries

Smith and Varzi (2000) distinguish two categories of geographic objects, those whose boundaries are natural, such as island, continents, rivers, and those which were defined by humans, for instance, property line, national boundaries, etc. Using Latin expressions, they call the former, *fiat*, and the latter *bona fide* objects.

But the situation is a little more complex. For instance, some states in the USA have some *fiat* boundaries along rivers and *bona fide* boundaries sometimes defined by parallels or meridians.

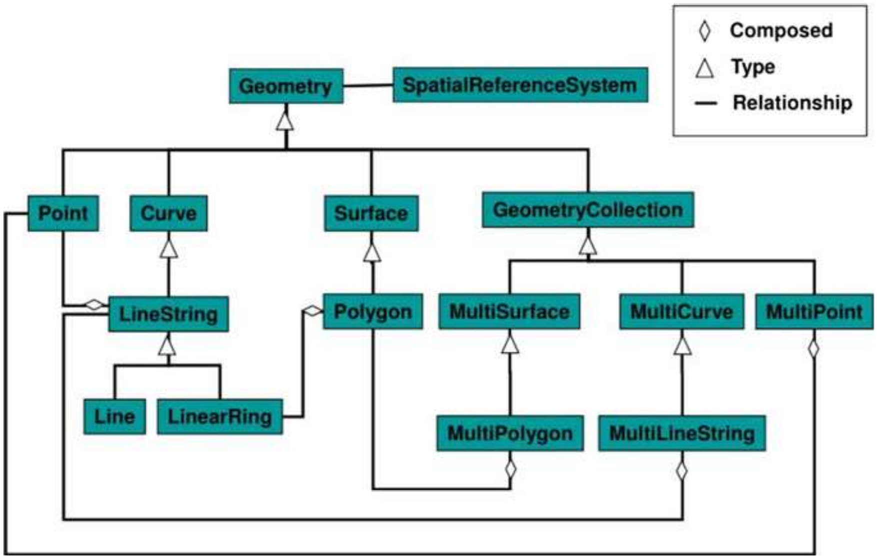
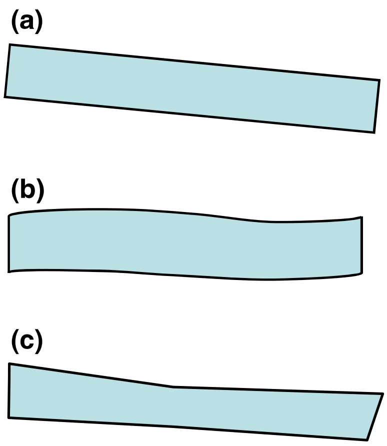


Fig. 2.3 OGC spatial model

Fig. 2.4 Various types of ribbons. **a** rectangular ribbon; **b** curved ribbon; **c** loose ribbon



Now, consider mountains. Where do they begin, where do they end? Similarly, for deserts, mangroves, etc. Here, fuzzy set theory (Zadeh 1965; Pantazis and Donnay 1998) can help model those objects with indeterminate boundaries: in the very core of the mountains, points can be said to belong 100%, farther 80%, or even 20%. As a consequence, some features can be modeled by means of fuzzy sets (Fig. 2.6).

But the manipulation of fuzzy sets in geography is not so easy. For those objects, fuzzy set theory can be used in which some membership grades can be defined (Fig. 2.7) (Zadeh 1965). An interesting model (Cohn and Gotts 1996) is the "egg-yolk" model with two parts, the core (the yellow part) and the extension, the white

Fig. 2.5 Identification of ribbons in an urban context

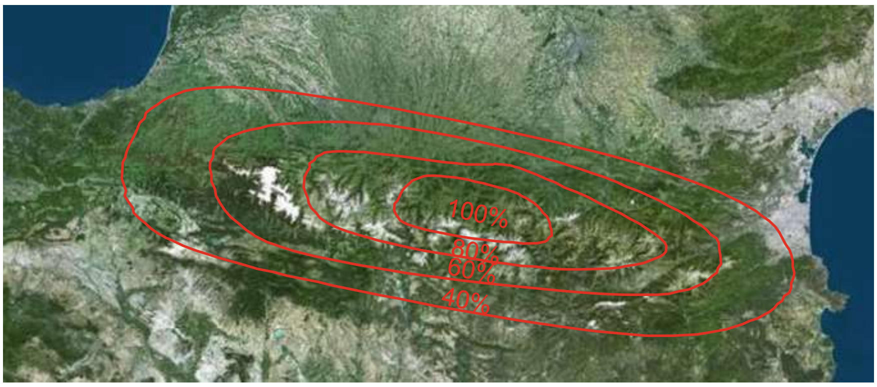


Fig. 2.6 Example of mountains modeled by fuzzy sets

part of the egg. For instance, for a river, the "yolk" represents the minor bed, whereas the "egg" modeled its major bed. Another example is given in Fig. 2.8 in which the mangrove and the jungle are modeled with the egg-yolk representation.

Again, the egg-yolk model can be used to model ribbons: so, a fuzzy ribbon can have a wider ribbon including a narrower ribbon. This model can be applied to modeling rivers each of them with their minor bed and major bed as exemplified in Fig. 2.9.

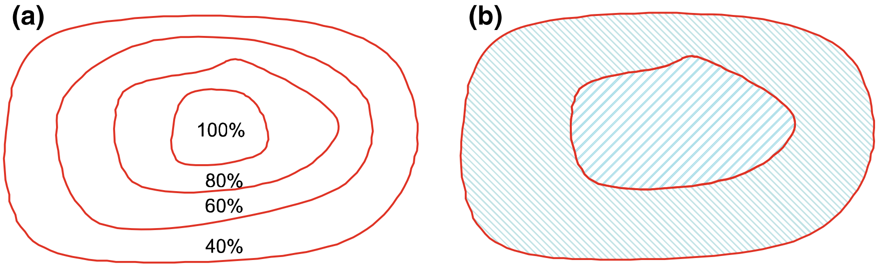


Fig. 2.7 Fuzzy geographic object. **a** Different membership grades. **b** The egg-yolk representation

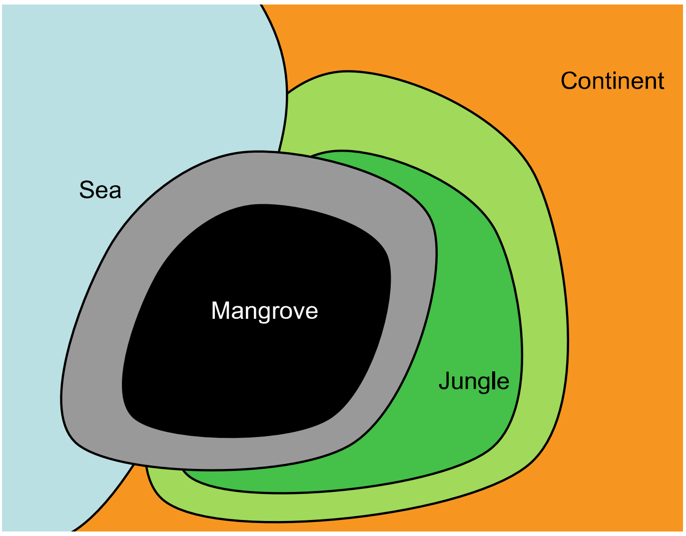
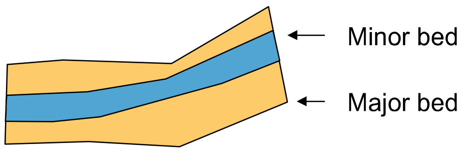


Fig. 2.8 Fuzzy geographic features

Fig. 2.9 River modeled with the egg-yolk representation, emphasizing the minor bed and the major bed



Multiple Representations

Now, continue for example to consider a river (Fig. 2.10). Various mathematical models can be assigned depending on the context. In cartography, this is generally a line; for navigation, an area and a volume. And considering tributaries, a so-called hydrographic network, or more exactly a hierarchical network can be defined. But if we add canals, this network is a little bit more complex. Moreover, during floods, the river bed is enlarged leading the existence of minor bed and major beds. As a

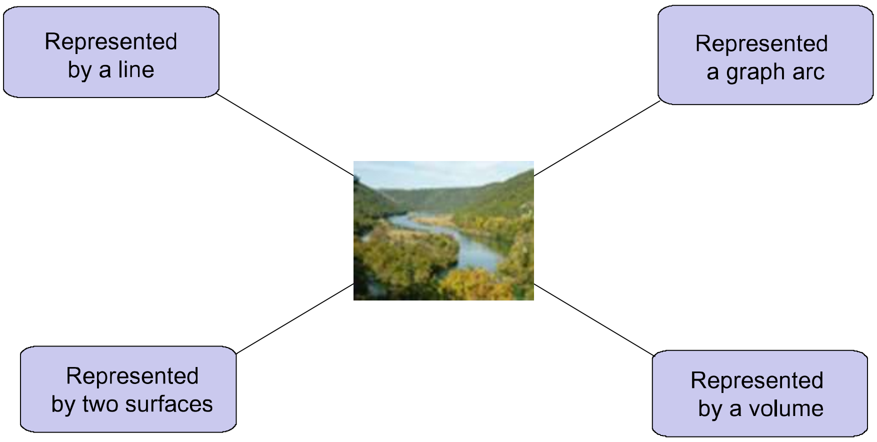


Fig. 2.10 Multiple representations of a river

consequence, multiple mathematical representations of the same geographic feature can be offered.

Now, consider an island, and ask several people to measure this island. Some, a little bit lazy, will only give the coordinates of a 100-point polygon, whereas other can give 1000 or more points. Even if two people are giving both 100 points, those points could be different. Finally, we get different measures for the same feature, leading to different computer geographic objects.

Suppose that we have now two different databases in which we have in both the representation of the same feature, namely, O_1 and O_2. Obviously, the stored data will be different, but representing the same object. If the stored measures are exactly the same, we can easily write $O_1 = O_2$. But, if only one data is a little bit different, this equality does not hold anymore! Whereas they represent the same feature! As a consequence, equality is a too strong concept for comparing geographic objects, and we need to weaken it.

Between two objects, A and B, a homology relation is a relation that is reflexive and symmetric which defines a sort of similarity between two things. Let us denote ₪ this relation, so that one can write A ₪ B. Therefore, both (A ₪ B) and (B ₪ A) hold. Remark that an equivalence relation ($\equiv$) is a homology relation, which is also transitive. Figure 2.11a illustrates two homologous polygons and Fig. 2.11b two homologous lines.

As an example, consider two persons (Mark and Fran) taking the coordinates of the same island as exemplified in Fig. 2.12. When we overlay those two polygons, we can see that they are not exactly the same, thus a relation of homology must hold between those two polygons.

Another aspect must be taken into consideration, namely linked to specifications. Back to the previous example, consider that one is measuring at low tide and the other at high tide: if the island has no beaches, no additional problem, but when

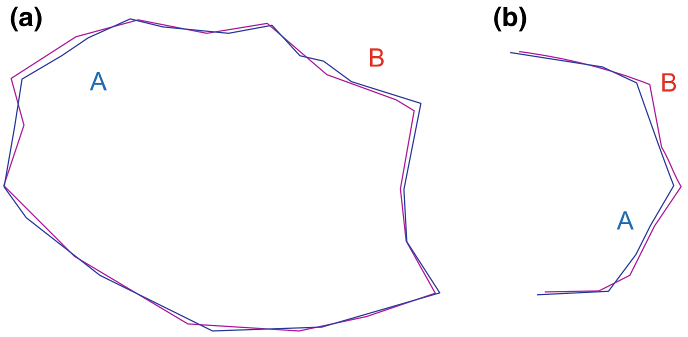


Fig. 2.11 Two homologous geometric descriptions of the same geographic objects. **a** Polygon homology. **b** Line homology

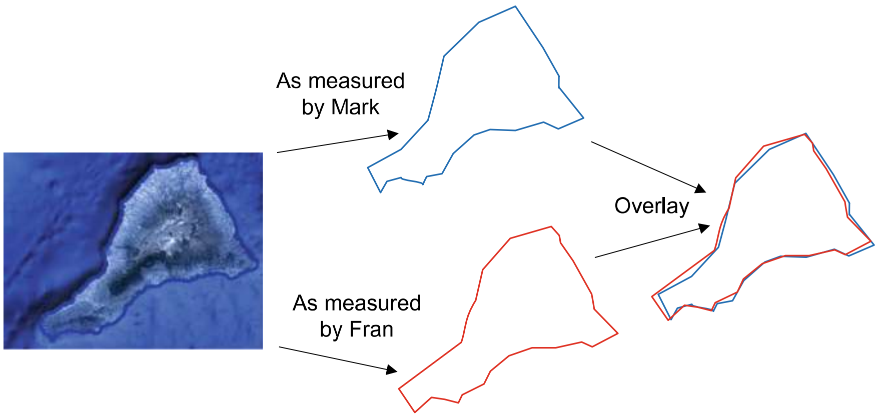


Fig. 2.12 Two descriptions of the same island, one by Mark and one by Fran and their overlay

there are several beaches, the size of the foreshore will change the measured data. Similarly, consider a house (Fig. 2.13); in the cadaster, measures are taken in the soil (soil coordinates), whereas in aerial photo, there are roof coordinates for which the difference can reach easily one foot.

Remember fractal geometry. In his seminal book, Mandelbrot (1967) shows that, depending on the length of a yardstick, the perimeter of an island varies: the shorter the yardstick, the longer the perimeter. And finally, the perimeter tends toward infinity whereas the area converges to a finite value.

As a consequence, depending on the application, the necessary number of points of a polygon vary. Suppose that for cartography, the limit is decided 0.1 mm, and we have a 10 m yardstick for measuring an island. The threshold scale s_0 will be 0.1 mm/10 m $= 10^{-5}$ for using all acquired points. If the yardstick is shorter, we can remove points because they become useless for this application.

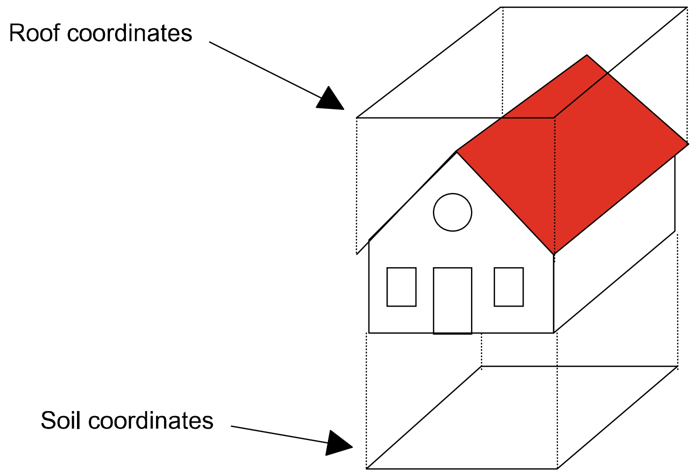


Fig. 2.13 For a house, soil, and roof coordinates

Connected and Non-connected Objects

A connected object is defined by the following: if we consider any couple of points belonging to this object, there is a path inside the object. Otherwise, it is said non-connected.

Naively speaking, a geometric object is defined with different sub-objects and holes. Consider Italy: it has several islands (Sicily, Sardinia, etc.) and two holes (San Marino and Vatican), so this country is said non-connected from a geometric point of view.

Back to the discourse about *fiat* and *bona fide* objects, we can state the following: fiat objects are always connected whereas bona fide can be non-connected, such as countries having islands or containing various pieces. Another example is the USA, with Puerto Rico and Alaska and other smaller territories throughout the world. In contrast, an archipelago is a set of isolated islands with a single name.

About Polygons and Tessellations

There are several ways to encode a polygon, and all of them have consequences regarding tessellations. Remind that each polygon may be defined either as a set of points connected by segments of lines or by a set of segments. Let call them, respectively, point-oriented polygons and segment-oriented polygons.

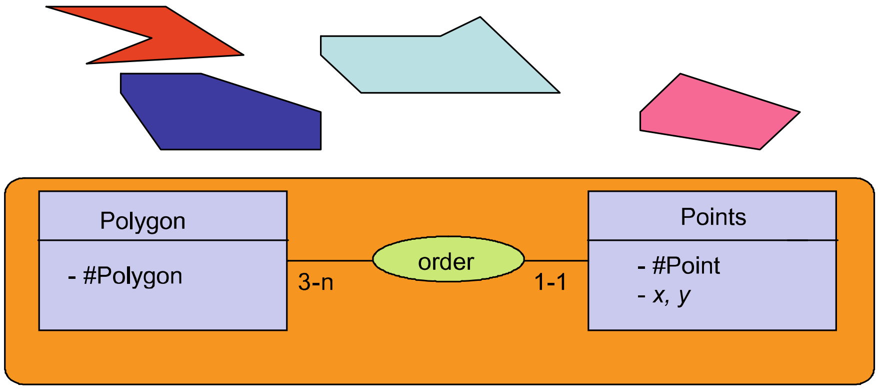


Fig. 2.14 Description of isolated polygons by a set of points by means of the entity–relationship model

Point-Oriented Polygons

A very simple way to encode polygons is to give an ordered set points, for instance, either in the clockwise order, or in the trigonometric order. As this representation is excellent for a non-connected polygon, there are difficulties for connected polygons. For example, consider again Italy; in addition to the main body of this country, we need additional polygons for islands such as Sicily, Sardinia, etc., and for the two holes, i.e., San Marino and Vatican City.

This representation is very common in GIS, but the great disadvantage is the difficulty to tackle consistent tessellations especially due to errors forming the so-called sliver polygons. An example is given in Fig. 2.14 using the entity–relationship approach (See Laurini and Thompson 1992) for more details. So, for checking the consistency of tessellations consisting of point-oriented polygons is a very complex task.

The more common way to store the coordinates is to consider a new abstract data type generally named "geometry". Anyhow, in the well-known standard OGC model,[2] this representation was chosen.

Segment-Oriented Polygons

In this representation, a polygon is seen as an unordered set of segments, each segment being limited by two points. Among advantages, we can see that there are no problems regarding islands and holes. But one of the disadvantages is that the reconstruction of a polygon is a little more complex task. See Fig. 2.15. Another advantage is that the tessellations are consistent.

[2] http://www.opengeospatial.org/.

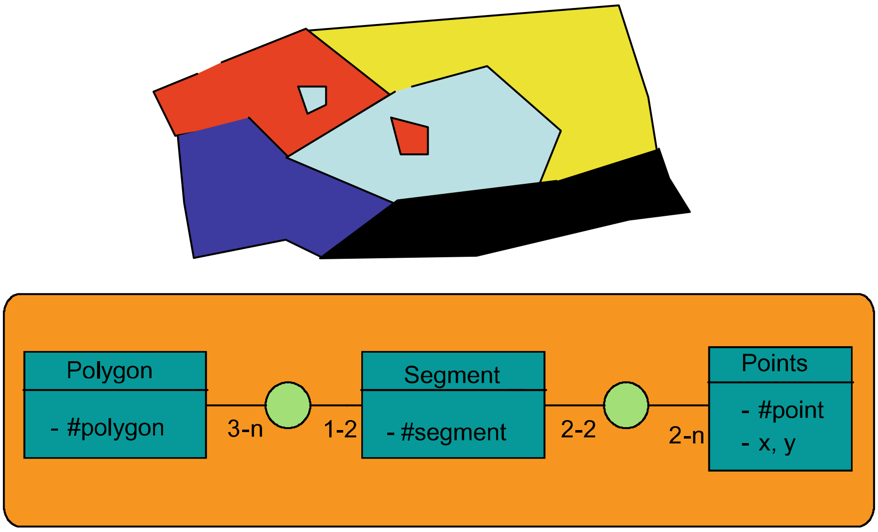


Fig. 2.15 Description of a tessellation of non-connected polygons with the segment orientation

Conclusions About Polygons

The OGC standards regarding geographic object modeling, based on point representations, have pushed to solved efficiently many practical problems. But by facing new applications, some models must be revisited. In the past, several other models have been proposed, either based on segments or allowing to manage consistent tessellations more easily.

Another aspect is the well-known Euler–Poincaré formula stating that $V + F = E + \mathrm{S}$ in which

- V is the number of vertices (previously called points),
- F is the number of faces (previously called polygons),
- E is the number of edges (previously called segment), and
- S is the number of disconnected tessellation sub-objects (holes and islands).

This formula can be used as an integrity constraint to state whether the tessellation is correct. Alas, this condition is necessary, but not sufficient because an extra vertex can balance an extra edge. Anyhow, such a formula must be extended to secure tessellations, and overall to take errors into account. In a cadaster, this formula can be written as $P+V=CB+E$ in which CB and P stand respectively for the number of city-blocks and parcels. While this formula is easy to check in the segment-oriented representation, it is very complex for the object-oriented one.

As a conclusion, all the discarded models must be examined again to test whether they can be more powerful for solving new salient problems.

Geometric-Type Mutation

In cartography, depending on the scale, the type of geographic objects can vary. Consider a city, at one scale, it is an area, but at a smaller scale, it becomes a point, and again at a smaller and smaller scale, as it is no more mentioned, it disappears. This can be formalized by the following chains:

- Area ==> Point ==> null object,
- Ribbon ==> Line ==> null object.

What About Shape Grammars

Shape grammars were initially conceptualized by the Italian architect Palladio (albeit not using this expression) for the systematic generation of rooms in buildings and façades, although repetitions can be easily seen in Ancient Egyptians or Babylonians constructions and cultural artifacts.

More generally, shape grammars allow the defining of iterative geometric objects (Stiny 1978, 1980), whereas fractal geometry defines recursive objects. Indeed, repetitions are commonly found in man-made environment. Look, for instance, the models of cities (Fig. 2.16) as defined initially by Hippodamos of Milet, successively refined by L'Enfant for Washington D.C. or Benoît for La Plata in Argentina. Other models do exist (Laurini 2017a, b) for designing populated communities, schools, barracks, hospitals, campuses, etc. (See Halatsch et al. 2008 for an example in Master Planning or Schirmer and Kawagishi 2011).

Beyond the Vector/Raster Debate

Two classes of representations exist, vector or raster. As the vector representation is based on points, lines, etc., the raster representation is based on grids and generally on squared cells organized into rectangular grids: this is the case for satellite images and photographs. The main theoretical problem is that with collections of squared and rectangles, it is impossible to continuously cover the geoid when squared are big. But if you accept the tolerance as explained in section "Euclidean and Spherical Geometry", this becomes possible.

Another important aspect of using the raster model is for the representation of phenomena generally modeled by field orientation (terrains, meteorology, etc.). In other terms, we can state that for any point, a function $f(x, y, z)$ can determine a value, for instance, in meteorology. This aspect can be extended from scalar fields to vector fields.

Remember that a 3D continuous field is governed by Laplace equation:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

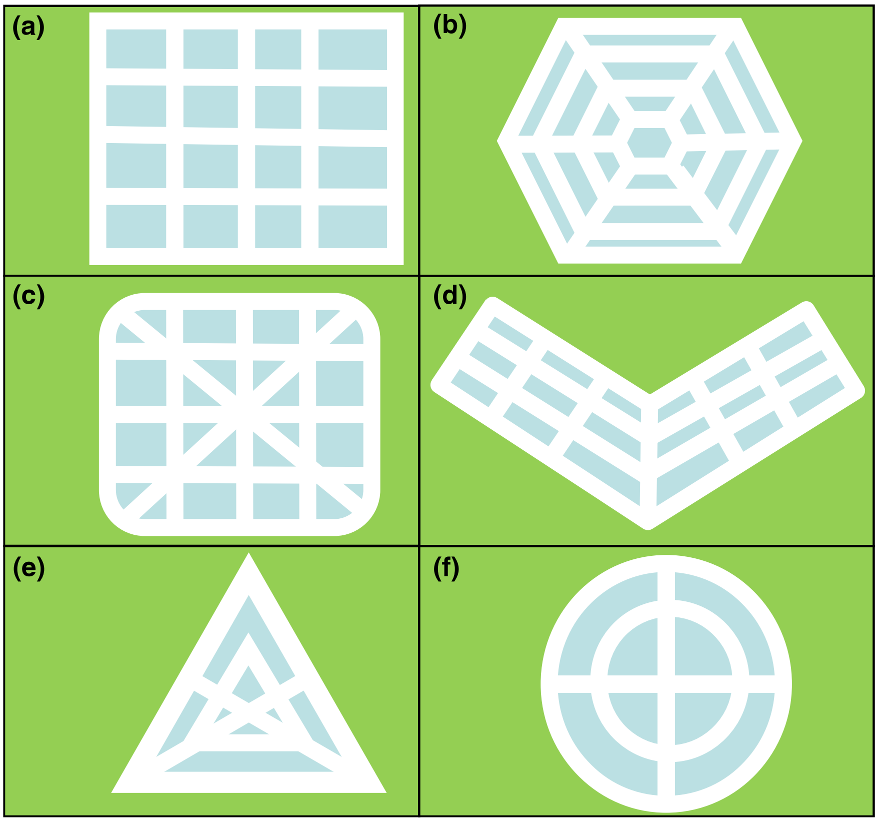


Fig. 2.16 Illustration of spatial patterns in cities. **a** Plan designed by Hippodamos of Milet. **b** Palmanova. **c** La Plata. **d** Brasilia. **e** Erice in Sicily. **f** Beijing's ring roads

Couclelis (1992), by stating "is the world ultimately made up of discrete, indivisible elementary particles, or is it a continuum with different properties at different locations?" showed that field orientation as can be a nice way to model environmental phenomena. Then Kemp (1996) explicated more precisely the variables, and Laurini et al. (2001) proposed a complete model for applications in meteorology. See Fig. 2.17.

Eventually, three GIS models are in competition, the vector model for applications such as cadaster, traffic, etc.; the raster model especially for remote sensing; and finally, the field-oriented approach for modeling environmental phenomena. Anyhow, two tracks of research may be followed:

1. a theoretical path which must tend toward a unique and integrative mathematical model; for the moment being and as far as we know, no clue is given to explore this solution;

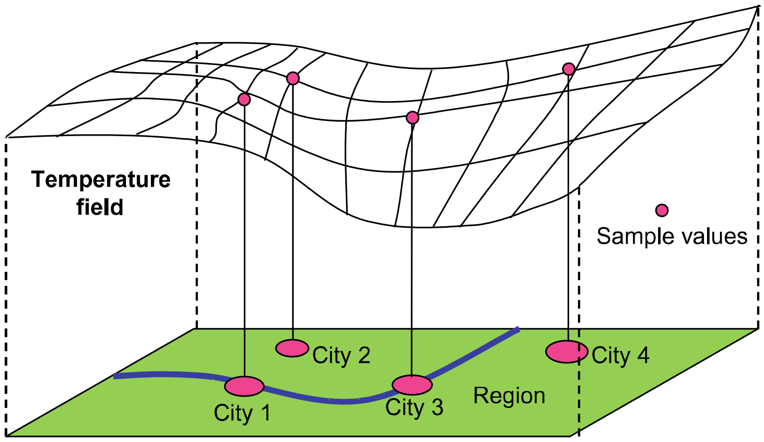


Fig. 2.17 Example of a field-oriented model for temperature

2. pragmatic solutions based on interoperability software products, essentially by trying to make bridges when needed.

And 3D

In the continuation of mapping which prefers 2D objects, it is more and more important to study 3D aspects. As an intermediary to sore terrains, 2,5D models have been created in which the altitude *z* is considered as an attribute. As it is possible to store easily typical terrains and mountains with this solution, it is impossible to model caves and also some cliffs, because to one couple (*x*, *y*), there may correspond several *z*'s.

Different methods exist to acquire terrain models. From conventional theodolites, one can measure altitudes of points, for instance, organized along a squared grid as illustrated in Fig. 2.18a, whereas Fig. 2.18b depicts the case of terrain altitudes acquired through a distance laser. In the first case, we deal with point heights and in the second case with "pixel" heights.

Afterward, the resolution must be taken into account. If the resolution is less than 10 cm, the two models appear equivalent; but when the resolution is larger, for instance, 100 m, the situation is totally different leading to "scalp" mountain summits.

Many authors have faced full 3D models. See for instance (Van Oosterom et al. 2008). The objective is not only to offer 3D models of the Earth but also to handle practical 3D models for cities and inside buildings. The present challenge is to offer seamless models for outdoor and indoor applications. For the description of building, the BIM standard (Building Information Modeling)[3] is presently mainstream, but

[3] https://www.nationalbimstandard.org/.

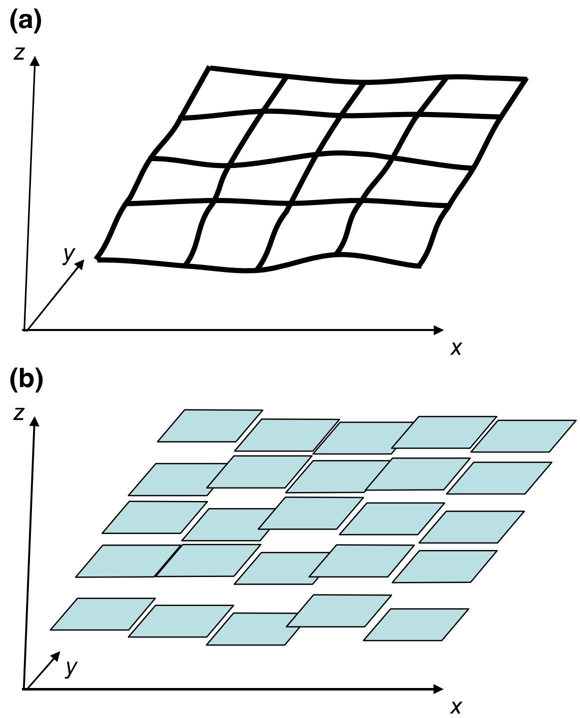


Fig. 2.18 Examples of terrain models. **a** Based on points located along a grid. **b** Based on squares (pixels) whose altitudes come from laser beam

sometimes modified by additional national options. One of the issues is the link with Computer-Aided Design (CAD); as the links with CAD in architecture is obvious, it is not the case with mechanics. When designing a new plant with a lot of robots, the possible connections between architectural and mechanical CAD software products must be envisioned, within a sort of interoperability mechanism.

So, the big question is "what is the limit from a geometric point of view?" Presently, the limit is something like 0.1 mm, but the interoperability of building modeling with the objects located inside buildings can lead to other solutions.

After having reexamined geographic objects and their mathematical modeling, let's revisit topology.

Topological Relations

Topology is the study of the relative positioning of two geometric objects. The word topology comes from the Greek τόπος, place, and λόγος, study. From a mathematical point of view, there are several domains of usage of this term sometimes with different meanings. In this chapter, I don't want to annoy the reader with too complicated notions, but only to show some practical difficulties.

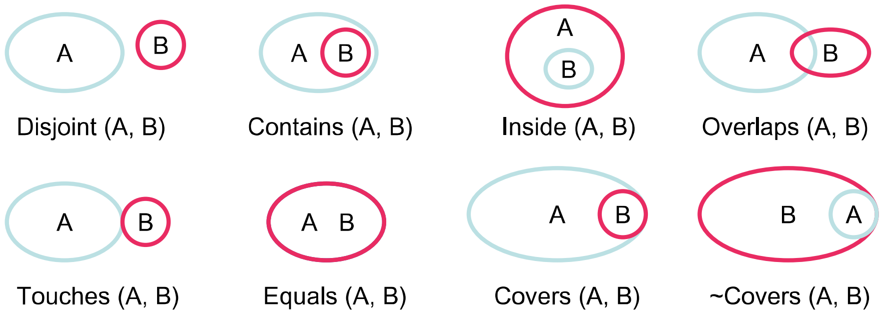


Fig. 2.19 Egenhofer topological relations at 2D (Egenhofer 1994)

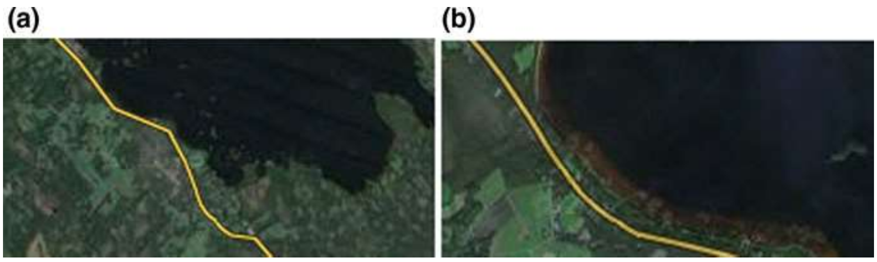


Fig. 2.20 According to scale, the road *Touches* or not the sea

Usually, in GIS, the Egenhofer (Egenhofer and Franzosa 1991; Egenhofer 1994) model is used as exemplified in Fig. 2.19 for defining the relationships between two objects *A* and *B*. Sometimes, the so-called RCC is also in use (Randell et al. 1992), but their characteristics are similar.

But, due to both the difficulties of measuring and the scaling consequence, we need to revisit this model.

Scaling Effects

Consider a road going along the sea; so, implying a *Touches* relation between the road and the sea. But if we carefully consider those features, sometimes there are small beaches between the road and the sea (Fig. 2.20). From a cartographic point of view, the type of relation will vary: indeed, at a scale of 1:1000, the relation is *Disjoint,* whereas at 1:100,000, there is a *Touches*, since the beach is discarded. More generally, the concept of granularity of interest will enlarge the concept of scale.

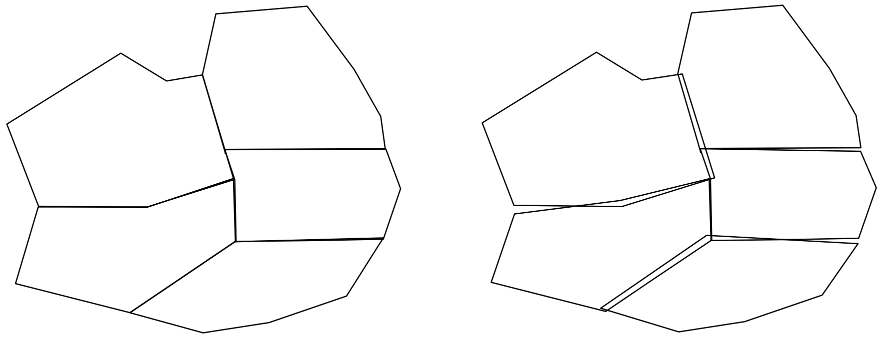


Fig. 2.21 Examples of irregular tessellations. **a** A mathematical good-looking tessellation (valid). **b** A practical tessellation (loose tessellation) with sliver polygons in which errors are voluntarily exaggerated

Measuring Effects

Consider two neighboring countries. Officially, there is a *Touches* relation between them. But since the boundary coordinates were acquired differently, they are different; as a consequence, small overlaps or holes may exist.

Back on Tessellations

Let's continue the reflection considering tessellations, i.e., composed of many polygons. Considering that each of them has errors, all the boundaries between them do not coincide. In Fig. 2.21, two cases are illustrated, the first case (Fig. 2.21a) of a good-looking or consistent tessellation, and in (Fig. 2.21b) the case of a loose tessellations in which boundaries do not coincide. Thus, specific algorithms must be run to correct this tessellation.

Indeed, by applying the OGC standards for polygons, this problem is common.

Another problem in tessellation comes from very small polygons. For instance, in a small European map, it is common not to consider small countries some as Andorra.

In reality, the situation is a little bit more complex, because different cases can occur. Figure 2.22 depicts those cases.

1. The first one is a very small polygon inside a bigger (let's call it Vatican style). In this case, this polygon can be discarded.

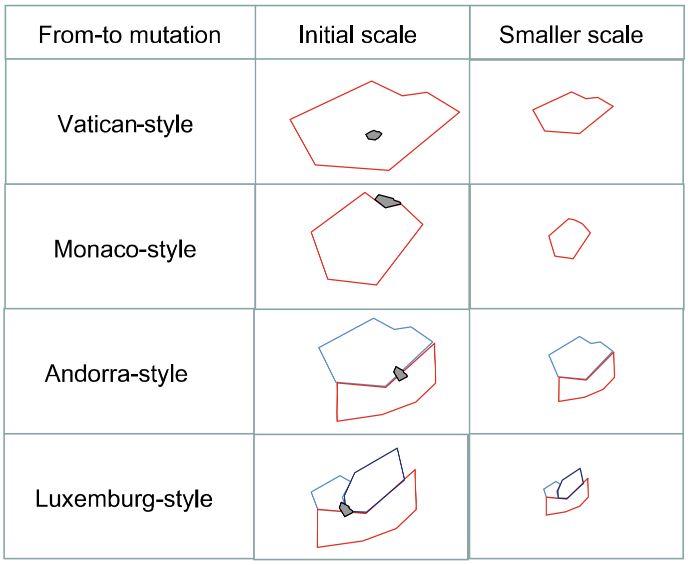


Fig. 2.22 Different cases of polygon disappearance due to scaling in tessellations

2. The second is when the small polygon is located at the boundary of the tessellation (Monaco-style); in this case, it could be either discarded or absorbed by the neighbor.
3. The third one is when the smaller polygon has only two neighbors (Andorra-style). In this case, for instance, each neighbor absorbs 50% of the smaller polygon.
4. The latter case is when the smaller polygon has three neighbors (Luxembourg style). Likewise, it can be split and absorbed.

To conclude this paragraph, let's say that novel algorithms must be designed.

Toward New Topological Relations

Taking these characteristics into account, some new kinds of relations must be defined. Starting from Egenhofer relations, the relation "*Equals*" is already transformed into a homology. When objects are very far, the "*Disjoint*" relation holds on; but when the boundaries are very close or overlapping a little bit, the problem is more getting complex. Figure 2.23 illustrates those new relations.

From a formal point of view, we have the following statements, provided that some thresholds are provided:

1. DISJOINT: If *A* far from *B*, no problem; but if A is very close to *B*, the relation can become TOUCHES.

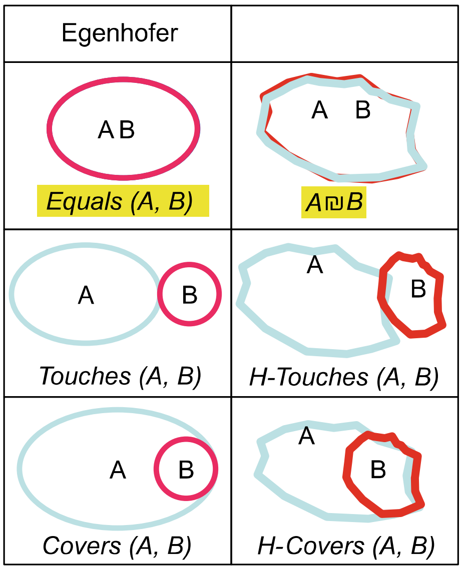


Fig. 2.23 New types of topological relations

2. TOUCHES: taking measuring difficulties into account, this relation practically never holds.
3. OVERLAP: if this is a very small overlap, the new TOUCHES relation can hold.
4. COVERS: taking measuring difficulties into account, this relation practically never holds.
5. CONTAINS: if the distance to the boundary from the smallest object to the biggest object is very small, maybe a COVERS can hold.
6. INSIDE: similar as CONTAINS, but exchanging *A* and *B*.
7. COVERBY: similar as COVERS, but exchanging *A* and *B*.
8. EQUALS: see homology.

Defining exactly those new relations is outside the goal of this chapter; nevertheless, we can say that one or several thresholds must be defined. The big difficulties stay in their values. Are those values unique for any kind of geographic objects, or several must be given? For instance, when comparing parcels and countries the areas of which are very different, a threshold given as a percentage could be of interest; perhaps 3%, less or more. The question is delicate and implies more investigations: they are known as sliver polygons in tessellation.

Now, let examine the problem of encoding geospatial knowledge.

Mathematical Requirements for Automatic Geospatial Reasonings

Two very different aspects must be considered, first, the actual way of encoding geospatial knowledge and second, the requirements for automatic reasoning.

Encoding Geospatial Knowledge

In conventional logic, assertions (or statements) are usually written based on predicates (Boolean expressions). Let us defined A_i and B_j some predicates (Boolean conditions), a rule is expressed as

$$A_1 \wedge A_2 \wedge A_3 \wedge A_n \Rightarrow B_1 \vee B_2 \vee B_3 \vee B_m$$

that is a conjunction of predicates implies a disjunction of other predicates. As the conjunctions are clear from a practical point of view, the disjunctions are not very clear. Does it mean that all B_i must be true, or only some of them? So, a clarification of the semantics is needed.

In Business Intelligence, those assertions are encoded as rules. According to Ross (2011), two types of rules exist, IF-THEN-fact and IF-THEN-action in which the first one corresponds to the creation of a new fact, for instance, the value of an attribute or even the existence of a predicate, and the second for an action to run, maybe by a computer, a human or any kind of machine. But in our case, those statements are too limited. Hence, in Laurini (2017a, b), concerning geoprocessing, new other types of rules can be distinguished:

- IF-THEN-Zone, for the creation of a zone from scratch, for instance, the administrative creation of a recreational park;
- Metarules such as “IF some conditions hold, THEN apply *RuleC*”;
- among the latter a special case is located rules such as “IF in the place A, THEN apply *RuleB*”, meaning that when we are in the place A, the *RuleB* holds;
- colocation rules the meaning of which is “if something here, then another thing nearby”;
- bilocation rules such as “IF something holds in place P, then something else in place Q”; in other domains, this rule is similar to the well-known butterfly effect.

Regarding colocation rules, since in a lot of towns, a church is located at the vicinity of the town hall (say within 500 m), the encoding can be as follows:

$$\forall T \in GO, \exists C \in GO, Type(T) = \text{“Town hall”}, Type\,(C) = \text{“Church”} : \\ Distance\,(Centroid(Geom(T)), Centroid\,(Geom(C)) < 500 \\ \Rightarrow Colocation(T, C)$$

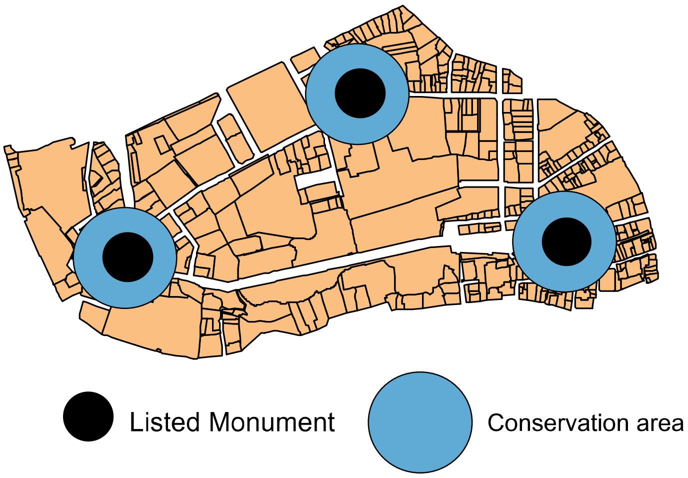


Fig. 2.24 At the vicinity of listed monuments, it is prohibited to construct a new building within the conservation area

In which

- *GO* corresponds to the set of Geographic Objects,
- *Type*, to a type as defined in an ontology,
- *Geom*, a function for the geometry of an object, and *Centroid* for defining its centroid,
- *Distance*, an operator to compute the distance between two points, and
- *Colocation*, a relation of colocation between two objects.

As example in urban planning, let us consider the case of somebody having a project to construct a new building within the conservation area of a listed monument. Practically, in all countries, such new construction is prohibited (Fig. 2.24).

To deny the approbation of this building, the rule can be encoded as follows (distance equals 100 m):

$$\begin{gathered}
\forall Terr \in EARTH, \forall B \in PROJECT, \forall M \in GO, \\
Type\,(B) = \text{“Building”}, \\
Type\,(M) = \text{“Listed_Monument”} : \\
Inside\,(Geom(B), Terr) \\
\wedge Inside(Geom(M), Terr) \\
\wedge Inside(Geom(B), Union(Buffer(Geom(M), 100))) \\
\Rightarrow \\
State\,(B) = \text{“LM_Denied”}
\end{gathered}$$

In which

- *Terr* represents the territory onto which this rule applies,
- *PROJECT*, the set of projects,
- *Inside*, a topological relation,
- *Union* and *Buffer*, geometric functions.

Requirements

With the increasing use of artificial intelligence and knowledge engineering in a great variety of domains, it could be interesting to re-examine how conventional mathematical background must be revisited to allow automatic reasoning. In fact, as it was thought that this problem was answered decades ago, mathematical issues must be considered again to be at the edge in the critical path of research in geoprocessing. Among those problems in automatic reasoning, the more salient to be solved are as follows:

- Encoding of geospatial rules and mechanisms to deduce new knowledge chunks or to suggest new actions to be made (Laurini et al. 2016);
- Considering big data in smart cities, create an efficient framework for deep learning, i.e., starting from examples and observations to derive mechanisms for better solutions.

So, a research program must be set to exhibit novel solutions in the following domains:

1. find a unified representation covering all aspects of geographic features (2D, 3D, time, multi-representations, etc.), robust enough to take rid of measurement uncertainties;
2. based on this new model, design powerful algorithms for all conventional geographic queries (point-in-a-polygon, topological queries, graph queries (minimum path, etc.), spatial analysis, what-if models);
3. present and experiment models for encoding geographic rules, able to overpass uncertainties, to allow deduction;
4. innovate in geovisualization; and
5. allow deep learning.

Conclusions

Historically speaking, the future will be on smart cities and territorial intelligence. Facing this evolution, the goal of this chapter was to revisit geometric modeling for geographic applications from a philosophical point of view. It has been shown that, as geometry was born to measure land (i.e., for geoprocessing before the word was existing), practical problems usually issued from data acquisition imply to revisit the main concepts of geometry applied to geoprocessing, and geographic reasoning.

Geographic reasoning demands very robust theories to get rid of difficulties derived from the various devices of data acquisition. In the chapter, it has been shown that comparing geometric entities imply the weakening of conventional mathematical operators (for instance from "=" to "₪"), of topological relations between geographic objects.

Four main remarks must be presented concerning GIS geometry.

The first remark is the necessity to revisit the well-known quadruplet (point, line, area, volume) by integrating the concept of ribbons which is more appropriate to model the so-called GIS linear objects.

The second remark is the necessity to create more robust mechanisms to compare geographic objects based on the multiplicity of representations and measurements.

The third remark concerns topology. Conventional topology (f.i. Egenhofer topology) must be revisited to properly integrate differently measured objects.

To define a library of adequate functions and relations is to be easily handled for automatic reasoning.

To conclude this chapter, let me make some final comments:

1. In the context of smart cities, and especially of geometric reasoning, it is important to revisit geometric modeling to provide more robustness (to get rid of measuring uncertainties) and more independence from geometric representations.
2. The increasing role of sensors in urban context must imply some new real time approaches in spatial analysis.
3. From big data, we need to design new technologies to integrate reasoning from deep learning in geoprocessing applications, not only to understand but also to explore new scenarios of development.

References

Cohn A, Gotts N (1996) The "Egg-Yolk" representation of regions with indeterminate boundaries. In: Burrough P, Frank A (eds) Geographic objects with indeterminate boundaries. Taylor & Francis, London-Bristol (PA)

Couclelis H (1992). People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In: Proceedings of the conference: theories and methods of spatio-temporal reasoning in geographic space, international conference GIS—from space to territory: theories and methods of spatio-temporal reasoning, Pisa, Italy, 21–23 September 1992, pp 65–77

Egenhofer M, Franzosa RD (1991) Point-set topological spatial relations. Int J GIS 5(2):161–174

Egenhofer M (1994) Deriving the combination of binary topological relations. J Vis Lang Comput (JVLC) 5(2):133–149

Halatsch J, Kunze A, Schmitt G (2008) Using shape grammars for master planning. In: Gero JS, Goel AK (eds) Design computing and cognition DCC'08. Proceedings of the third international conference on design computing and cognition. Springer, Dordrecht. https://doi.org/10.1007/978-14020-8728-8_34

Kemp K (1996) Fields as a framework for integrating GIS and environmental process models. Part 2: Specifying field variables. Trans GIS 1(3):235–246

Laurini R (2014) A conceptual framework for geographic knowledge engineering. J Vis Lang Comput 25:2–19. https://doi.org/10.1016/j.jvlc.2013.10.004i
Laurini R (2017a) Geographic knowledge infrastructure: applications for territorial intelligence and smart cities. Elsevier and ISTE, London
Laurini R (2017b) Nature of geographic knowledge bases. In: Fazi S, Mahmoudi K (eds) Handbook of research on geographic information systems applications and advancements. Hershey (PA), IGI-Global
Laurini R, Gordillo S, Mostaccio C (2001) Architecture of a new field-oriented database system for meteorology. In: XXVI general conference of the European geo-physical society, Nice, France, 25–30 March 2001, EGS Newsletter, Number 78
Laurini R, Servigne S, Favetta F (2016) An introduction to geographic rule semantics. In: Proceedings of the 22nd international conference on distributed multimedia systems (DMS 2016), Salerno, Italy, 25–26 Nov 2016. Published by Knowledge Systems Institute
Laurini R, Thompson D (1992) Fundamentals of spatial information systems. Academic Press, London
Mandelbrot B (1967) How long is the coast of Britain? Statistical self-similarity and fractional dimension. Science 156(3775):636–638. https://doi.org/10.1126/science.156.3775.636
Pantazis D, Donnay JP (1998) Objets géographiques à limites indéterminées. Modélisation et Intégration dans un modèle conceptuel de données. Revue Internationale de Géomatique 7(2):159–186
Randell DA, Cui Z, Cohn AG (1992) A spatial logic based on regions and connection. In: Nebel B, Swartout W, Rich C (eds) Proceedings of 3rd international conference on knowledge representation and reasoning. Morgan Kaufmann, Los Altos (CA)
Ross RG (2011) More on the If-Then format for expressing business rules: questions and answers. Bus Rules J 12(4). http://www.BRCommun2002ity.com/a2011/b588.html
Schirmer P, Kawagishi N (2011) Using shape grammars as a rule-based approach in urban planning—a report on practice. In: 29th conference on education in computer aided architectural design in Europe (eCAADe 2011), Ljubljana, Slovenia, Sept 21–24
Smith B, Varzi AC (2000) Fiat and bona fide boundaries. Philos Phenom Res 60:401–420
Stiny G (1980) Introduction to shape and shape grammars. Environ Plan 7(3):343–351
Stiny G, Mitchell WJ (1978) The palladian grammar. Environ Plann B 5:5–18
Van Oosterom P, Zlatanova S, Penninga F, Fendel EM (eds) (2008) Advances in 3D geoinformation systems. Lecture notes in geoinformation and cartography. Springer, Berlin-Heidelberg
Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353