An Introduction to Geographic Rule Semantics

Robert Laurini
Knowledge Systems Institute, USA
and LIRIS, INSA-Lyon,
University of Lyon, France
Robert.Laurini@insa-lyon.fr

Sylvie Servigne
LIRIS,
INSA-Lyon,
University of Lyon, France
Sylvie.Servigne@insa-lyon.fr

Franck Favetta
LIRIS,
Claude Bernard University of Lyon,
University of Lyon, France
Frank.favetta@liris.cnrs.fr

Abstract — As business rules are very common in business information systems, it is fundamental to analyze geographic rules in order to include them into geographic knowledge systems. The goal of this paper is, starting from examples taken overall from urban and environmental planning, to examine geographic rule semantics, and in particular to model space by combining logic and computational geometry. They can come from laws, physical laws, socio-economic considerations, from best practices, or from data mining. By defining new concepts such as metarules, superseded rules and jurisdictions, a conceptual model is proposed.

Keywords-component: geographic information systems, geographic knowledge systems, geographic knowledge, geographic rules.

I. INTRODUCTION

According to Graham [7] and Morgan [11], rules (business rules) should be considered as first-class citizens in computer science. In enterprises, the ‘craft’ of expert know-how is capitalized in an information system in the form of “business” rules. These rules can then be explained and implemented in applications such as business intelligence in software architectures integrated type ERP (Enterprise Resource Planning) or not. For example, the SAP integrated software package is based on a declarative formalism for the description of the job tasks with lists of rules, such as “automobile insurance does not cover drivers who have been recognized guilty of driving while intoxicated over the past two years”, or “when its monthly invoices are sent, should be included the supplementary documents that match the profile of the client”. On the other hand, the explanation and the formalization of business rules is still a hot topic as a new standard from the OMG issued in September 2015 entitled Decision Model and Notation. The objective is to provide a language of formalization of business rules including exploited in the decision-making process (OMG - DMN 2015). For more details refer to http://www.omg.org/spec/DMN/

Thus, a rule is a basic element of a strategy to build reasoning. In contrast to algorithms, they are expressed declaratively. Among business rules, Dietz [6] distinguishes between three categories:

- rejectors typically those related to quality control, that allow a rejection (rejection rules),
- producers such as those determining new values (ex VAT calculation); they can be considered as rules of production of information,
- and projectors such as those related to the replenishment of stocks.

In our case, a rule is not necessarily a legal regulation, but only an inference (implication) between elements or phenomena whose origin can also be physical, statistical or best practice type, or from data mining. Mathematically, a rule will be written in the form of Modus Ponens $A \Rightarrow B$.

Few studies have been conducted on the automatic reasoning in geoprocessing. Nevertheless one can include the works presented in the book edited by Kim et al. [9] and the special issue of the CEUS (Computers, Environment and Urban Systems) journal including the editorial of Batty [1] which were very typical of the time. But these works are old and the knowledge engineering framework has drastically changed. More recently, mention the paper written by Jain and Payal [8], but it does not address the reasoning part based on declarative rules.

Include also the paper [13] which operates a GIS on minerals data to discover characterization rules that can facilitate data mining and spatial analysis. In the same spirit, Biasotti et al. [2] present methods for extracting the characteristics of a field (peaks, passes, edges, etc.) from the raw digital terrain models, which could easily be described in the form of rules, including those based on the value of the Gaussian curvature of terrains.

However in GIS, beyond administrative rules, other statements may lead to geographic rules. In fact, let us look at some of them:

- in the United Kingdom, we drive on the left;
- in Canada, the majority of the population lives along the border with the United States;
- each capital city has an international airport nearby;
- between the two capitals, in general, there are direct flights;
- in the Northern Hemisphere, the more you are going to the north, the colder (but locally this is not always true);
- the more you climb a mountain, the colder;
- heavy rain upstream, downstream flooding;
- mosques are oriented towards Mecca;
- if a zone is a swamp, it is necessary to prohibit construction;
- if there is unemployment, the creation of companies or industrial areas must be encouraged;
- If a plot is adjacent to an airport, it is necessary to limit the height of buildings;
- It is forbidden to open a new pharmacy within 500 meters of an another already existing;
- A good practice in Mexico is to use a bus to go from Puebla to Oaxaca City.

DOI reference number: 10.18293/DMS2016-015
Another example is derived from the use of electrical plugs in the countries. This table could be considered as a Visual decision table (Figure 1).

Figure 1. Types of electrical plugs (a) and the countries where they are standardized. Source: http://wikitravel.org/en/Electrical_systems.

Among the rules, there are several categories. Take two examples “if it rains, I get wet” and “if it rains, I take an umbrella”. In the first case, it is a rule of physical type whose consequence is systematic provided that I am outside. For the second, it reveals a good practice that I am not obliged to follow. Therefore, we can see that if the premises are identical, the status of the conclusions can be totally different.

Thus, the goal of this paper will be to review the rules in geoprocessing in order to extract particular semantics and to lay the foundations of a machine-processable language and model. Voluntarily, we will treat neither temporal aspects nor 3D aspects although sometimes it is necessary to use these dimensions in certain rules.

The areas of applications fall within geoprocessing. Specifically, we are interested in the rules about not only geographic objects, but also about all objects whose knowledge of the location is important.

This paper is thus built: after having presented an introductory example and detailed computer modeling of rules, many of them will be examined. Finally a sketch of model will be given and some perspectives will be open. And to conclude this analysis, a definition of geographic rules will be presented.

II. INTRODUCTORY EXAMPLE: STREET NAMING

In [15], an example of rule encoding is given concerning road naming in Australia in order to automate the process. Rules are defined in the form of ontological vocabularies using SWRL a Semantic Web Rule Language based on a combination of the OWL DL and OWL. For more details, refer to https://www.w3.org/Submission/SWRL/.

In this paragraph, only some rules are presented. Rule R1 automatically infers information with the help of a road link between proposed and existing roads; this rule is necessary as every road needs to link with at least one other road to allow access. Rule R2 checks road length against road type; checking the road length for shortest road types (‘Place’, ‘Close’ and ‘Lane’) is necessary to avoid confusion with the preference for road usage. Rules R3 and R4 check the compatibility between road usage and road links; for example an open-ended road must have a road link at both start and end points of the road. And finally Rule R5 checks whether or not the proposed road has a wide panoramic view across surrounding areas.

However several remarks can be done because this is logic reasoning, not geographic reasoning:

- Rule 1 tests whether the new proposed road is linked to another existing road; but the link can be by a set of new proposed roads. The adapted solution is concerned by the order of presentation of new streets. A very general solution must be based on graph theory.
- In Rule 2, road length which must be taken into account, is given as a given attribute, not computed from road coordinates.
- In Rule 5, panoramic view is also given as an attribute, not calculated taken terrain morphology into account by 3D computational geometry.

As a consequence from this rapid analysis, classical logic is not sufficient to represent geographic rules since other mathematical domain must be integrated for reasoning.

III. GEOGRAPHIC KNOWLEDGE AND REASONING

The objective of this section is to brush the outline between the bases of geographic knowledge and automatic reasoning. After some generalities, we will quickly review the structuring of the knowledge bases and computer modeling of rules.

A General information

The purpose of a declarative rule-based model is to allow automatic reasoning. In contrast to expert systems of the past that were using simply logic, in our case, it will be very different. Here are some examples:

- defining the location of a new airport, a new hospital, a new stadium, social housing, etc. (Computational Geometry and Operation Research),
- checking the compliance of a building vis-à-vis building regulations (Topology and Computational Geometry),
- determining the best way to go from A to B, (Graph Theory or Computational Geometry in terrains),
- organizing a policy about urban green spaces (Spatial Analysis),
- determining transit policy (Spatial Analysis),
- flood analysis (Differential Equations, Computational Geometry)
- limiting crime in a city (Spatial Analysis),
- building garbage collection circuits (Graph Theory and Spatial Analysis),
- determining the most polluted, noisiest, places (Spatial Analysis and Operation Research),
- organizing the evacuation of the people during a volcanic eruption (Spatial Analysis and Operation Research),
- etc.
From these examples, we find that the geographic reasoning must invoke other mathematical disciplines such as:

- spatial reasoning by the integration of topology and computational geometry,
- graph theory in some cases like search for routes or circuits,
- spatial analysis,
- simulation techniques,
- fuzzy logic and reasoning,
- multi-criteria decision theory,
- and operations research.

These elements can be integrated in the form of procedures which will be invoked at the right time.

### B Geographic knowledge bases

Once selected a human language and a territory, remember (Figure 2) that a geographic knowledge base may include an ontology, a gazetteer, geographical objects, the relationship between these objects, rules and mathematical models. Figure 3 shows such a base linked to a geographic inference engine that will allow reasoning.

![Figure 2. Contents of a geographic knowledge Base.](image)

In input, there will be a territorial project to study so as to determine, through the geographic inference engine, the consequences in output. These consequences will be grouped in a result of feasibility that may be in the form of maps, diagrams, texts, etc.

![Figure 3. Linking a geographic inference engine to a geographic knowledge base.](image)

To conclude this section, all the components of a geographical knowledge base can be used to write not only a rule but also in their activation context.

IV. **STUDY OF THE SEMANTICS OF SOME GEOGRAPHIC RULES**

Now let us examine certain rules in different areas. First, we will discuss the rules across the globe (often as physical nature), then those only valid in certain places (often of an administrative nature) and those from data mining. We will continue by rules related to cartographic input and presentation data.

**A Global rules**

In this category, one may include geodetic and physical geography rules.

**A.1 Geodetic rules**

The rules of this type are valid all over the world because we consider cardinal points. If \( A \) is north of \( B \) and \( B \) is located north of \( C \), then \( A \) is north of \( C \), thus by applying a rule of transitivity. But there is nothing north of the North Pole; so therefore transitivity is limited. A similar rule can be written for the South. But for the East and West, transitivity is partial because of the rotundity of the Earth. If Rome is East of Los Angeles, and if Los Angeles, is East of Beijing, Rome should be to the East of Beijing, but it is actually to the West. The comprehensive rule will be written thus:

\[
\forall A, B \in \text{Earth} : \text{East}(A, B) \land \text{East}(B, C) \land \text{Longitude}(A, C) < 180 \Rightarrow \text{East}(A, C)
\]

\[
\forall A, B \in \text{Earth} : \text{East}(A, B) \land \text{East}(B, C) \land \text{Longitude}(A, C) > 180 \Rightarrow \text{West}(A, C)
\]

Remember that rarely there are lists of type \( \text{North}(R, S) \) or \( \text{East}(U, V) \) as such information is hidden in the coordinates of the \( R, S, U \) and \( V \).

If the application of such rules is easy for points, it is not the case for areas. If it is common to hear that the Switzerland is East of France, it remains that, given the geometry of the border, some places of the Switzerland are west of some points of the France. In the case of automatic reasoning, more sophisticated definitions should be developed for areas.

**A.2 Rules of physical geography**

In this area, the rules should represent natural phenomena and their consequences. For example following tsunamis, volcanic eruptions, storms, heavy rain, we must consider some automatic consequences. But, in addition, we must consider more recently, prevention, protection or mitigation, and effective real-time monitoring systems.

But due to the local topography, some rules may be invalidated. In the northern hemisphere, the more you go north, the colder. A certain scale, this type of rule is valid, but there are places where this reasoning is no longer valid. Therefore, one must distinguish between local and global rules. Here the local rule supersedes the global rule as, for example also when one is dealing with microclimates. The rules of the spatial distribution of flora and fauna, hydrology, etc. from climatology, meteorology fall into this category. Figure 4 gives an example of vegetation layers in the Alps.

Suppose that we are on the southern slopes (sunsided). Two options to write a rule:

DOI reference number: 10.18293/DMS2016-015
IF Lichen THEN Altitude > 3000;
IF Altitude > 3000 THEN Lichen.

In the first case, it would be a study linking a type of vegetation to elevation, while the second shows since we are at some level from the sea, what are the types of flora that we can meet.

Now consider the case of mathematical models and assume that we have $A = M(B, C, D)$. This formula can be easily transformed into a rule in the following manner: one must:

- write a procedure or a function representing $M$, which will be encapsulated,
- search or determine to $B$, $C$ and $D$ in the antecedent part,
- then run $M$ in the consequent part in order to determine $A$.

Another variant would be that $A$ may intervene into a condition; therefore $M$ will be invoked in the antecedent part.

B Local rules

Here, we will discuss only the rules applied on a restricted territory, namely the administrative rules and those relating to specific spaces.

B.1 Rules coming from laws

Each country has its own rules, not only from the administrative point of view, but also from location. For instance, when analyzing road traffic from aerial photos, it is important to know that in the United Kingdom the cars drive on the left. In addition, concepts such as language and currency can impact geographic rules.

Typically in addition in each country there are a Constitution and many laws governing the geographic aspects in relation to urban and environmental planning. Take the example of pharmacies in some countries. Figure 5 shows an example of this rule. To establish where it is possible to create a new pharmacy, several geometric operations such as determination of buffer zones and geometric difference have to be implemented in an encapsulated way. Two rules must be defined: (i) for determining the zone where possible, and (ii) whether the projected place belongs or not to this zone.

One of the peculiarities of the rules of administrative origin is the existence of sanctions whenever they are not met. But there may be cases of exemptions, exemption which should be taken into account in one way or another.

On the other hand, although there are international standards for the Highway Code, each country has its own peculiarities concerning for instance priority in the intersections, roundabouts, etc.

B.2 Urban planning rules

Generally, in each country, there are also laws that govern urban planning. Let us take a small example taken from building licenses as presented in Figure 6 where you can see a building that must follow certain rules.

Among the urban planning rules, there may be good practices as "bury the air engineering networks (electricity, telephone)" or even "before to create an underground metro line, we must move sewerage networks".

B.3 Local socio-economic rules

Due to its great inertia, demography generates rules of the type: “the more of children, the more of schools.” But may this statement be considered as an administrative rule, a good practice or a recommendation? And what about time frame?
A majority of countries have developed rules for the organization of the economy and companies. These rules have a significant impact on the use of the soil. For instance, consider the rule “along the edges of sea, the greater the distance from the sea, the lower are prices of homes.”

A particular case regards rules related to the flow of people or goods. In these two places or two families of places may intervene, either as origin or destination.

B.4 Good or best practices

Among the rules of good practice, there are those related to the description of itineraries. In most countries where the landscape is charged, description is often done according to the directions, cities and villages to cross. In other countries where the landscape is lightly loaded, the description is made by taking into account numbers and route directions, North, South, East and West. In the deserts and seas, it was common to use the positions of stars as landmarks. These good practices had been used for centuries; but now there are other ways to do so. However this type of knowledge is out of the topic of our research because it uses information outside the terrestrial globe (extra-terrestrial knowledge).

Good practices include techniques of numbering houses in cities, which may vary depending on the country. Alternatively, along a highway, the creation of an additional interchange may be the basis of economic development of a small town.

C Material rules

In this category one can find the rules related to the acquisition and visualization of data. Among these, one can discuss three types of rules:

• those related to data quality control [14, 4, 5],
• and those related to the mutation of object’s geometric types (i.e. area to point) and topological relations according to scales (disjoint to meet). For more details, refer to [10, 11]. For instance, the rule mutating an area into a point according to scale (as its centroid), can be written as follows (in which $O$ is an area-type geographic object, $2Dmap$ a function transforming the object at the scale $\sigma$ and $\varepsilon$’s some thresholds):

$$\forall O \in OG, \forall \sigma \in Scale, G_{\text{type}}(O) = \text{area}, O_\sigma = 2Dmap(O, \sigma): 
\left( (\varepsilon_3^2 < \text{Area}(\ O_\sigma) < (\varepsilon_3^2) \right) 
\Rightarrow \{ \text{Type}(O_\sigma) = \text{point}; O_\sigma = \text{Centroid}(O) \}.$$  

And then a smaller object can disappear:

$$\forall O \in OG, \forall \sigma \in Scale, G_{\text{type}}(O) = \text{area}, O_\sigma = 2Dmap(O, \sigma):$$

$$\text{Area}(\ O_\sigma) \leq (\varepsilon_3^2) 
\Rightarrow O_\sigma = \emptyset.$$

D Rules and plurality of places

Four cases are to be analyzed.

1 / From the previous examples, we can see that most of the rules refer to a unique place. But a rule such as “in England we drive on the left” also applies in other countries.

2 / However the rules related to the flow of people, goods and animals are characterized by two places, a so-called origins and destination. Consider for example those governing the movements of migratory birds. Accordingly the grammar of the rules must allow this scenario by considering three cases:

- bipolar flows (the most common),
- diverging flows where only the source is known, for example emigration,
- and the converging flow where only the destination is known, as for immigration.

3 / A third case is that of the rules corresponding to clusters of areas according to certain criteria, such as for example research of homogeneous areas in geomarketing or the definition of electoral boundaries. It is noteworthy that identifiers to these newly-created areas need to be assigned.

4 / And finally another example is about the importation of a good practice from one location to another location.

E Rules and of shareholders’ logic

One of the difficulties is the fact that among the urban actors, some have different “logics”. With regard to industry creation, an environmentalist or an industrialist may have different ideas on the possible implications of this or that choice. Similarly, some groups may have different priorities: before an empty space, athletes imagine a stadium, pupil’s parents a school and a promoter a building, etc.

From the formal point of view, these aspects will occur in multi-actor and multi-criteria decision support systems.

V. TOWARD GEOGRAPHIC RULES MODELING

Now that various examples were analyzed, it is possible to extract elements of modeling. Firstly, general considerations will be given, and then a computer model will be proposed. According to [12], they can be initially modeled by IF-THEN-fact or IF-THEN-Action. Some WHEN or WHERE clauses can respectively be applied for temporal and spatial aspects.

As we saw earlier, new concepts have emerged and it is necessary to clarify what is meant by superseded rules, metarules, jurisdiction, etc. Finally two tables will make it possible to synthesize the characteristics of rules and our level of knowledge and their formalization.

A Superseded rules

Indeed, certain rules can be superseded locally. In other words, it will be necessary to take account of this aspect not only in designing the rules, but also in the inference mechanism. This can be written such as: WHERE SOME LAND IF CONDITION32 THEN REPLACE RULE#25 BY RULE#28.

B Metarules

A metarule is a rule which arises from other rules for example in a regulatory framework. For example, all local urban plans must be in compliance with some higher level regulations, which thus appear as metarules. In other words, a metarule defines a set of rules that will be valid only when you will refer to this metarule. In addition It can define new concepts, new
legal mechanisms, or even of new decision-making bodies; a metarule can therefore come enriching an ontology with this new terminology; this point although current is extremely complex and will not be treated.

**C Jurisdiction**

One can call jurisdiction, the territory of application of a rule, a metarule and even the entire knowledge base. Therefore, the gazetteer will only deal with place names within this jurisdiction or through it (rivers, roads, etc.). In some cases, it would be advisable to include close external information such as the names of the neighboring territories.

**D Geographic rules and objects**

As seen previously, the geographic rules commonly involve geographical objects (e.g., buildings in flood zone) and also geographic objects; see for instance vegetation (Figure 4) or habitats of animals.

But in addition, geographic objects may be deducted from rules. Take the example in maritime laws that distinguish territorial waters and international waters. The principle is based on the distance of 200 nautical miles except for particular cases.

The big problem is that the variables depth and width are not known explicitly, but from a 3D geometric reasoning from the morphology of the waterway. Furthermore, if the river has narrow meanders, it must ensure that long barges can pass.

Ultimately, we deal with producing rules according to Dietz’ terminology [6]. Generalizing the previous examples, we need to integrate those items deducted geometrically from geographic rules, namely new objects, new types of objects, new attribute values, or even new relationships between two objects and new spaces that can intervene e.g. as jurisdiction.

Thus, once known the rule for the determination of these objects, we can create a new class from this geographic rule. In other words, the consequence part of the rule will enrich the ontology by a creation of new class derived from geometric reasoning, and sometimes even enrich the gazetteer.

**E Summary**

At the level of formalization, rules of physical geography can be encapsulated into programs from a procedural way we need a mechanism to integrate them into a declarative model. With regard to the laws, they are known at a time $t$, but may change over time in the form of statutory instruments for which the translation into declarative forms may be difficult or too simplistic.

In addition to the previous features, need to clarify the nature of the implication ($\Rightarrow$). In fact, it has several meanings:

- in the case of physical phenomena, it corresponds to physical laws or causal chains (cause and effect); the implication is therefore automatic, often within a delay;
- However, if the physical law is only known empirically, the kind $a = f(b, c, d, ...)$, formula where $a$ is the value of an attribute of a geographic object, then the rule will affect the value calculated with a margin of error;
- Rules of legislative type are human laws; generally if the law is not enforced, sanctions may appear, thus involving an ELSE clause in the computer rule.
- In association rules (sometimes referred to as frequent association) from data mining, the semantics of the $\Rightarrow$ sign should be modulated according to the value of confidence related to this association rule;
- Concerning good practice rules, the implication will be judged in desirable manner;
- Finally, if the rule involves fuzzy objects, the semantics of the $\Rightarrow$ sign will be modulated according to the values of fuzzy membership degrees.

Now that the main elements of geographic semantics have been identified, it seems possible to propose a first model.

**F Outline of model**

From the analyzed examples, first there is that there is a many-to-many relationship between the rules and the names of places, and another between the rules and the types of geographical objects.

In addition, some sites have the ability to emit the metarule (countries) which will apply on inner places and will be a normative framework for located rules. Figure 7 depicts this model. However, regarding rules, things are a little more complicated because they are encoded in a language that remains to be defined, for example from an extension of RuleML [3]. Indeed, for the design of the code of these rules, knowledge must include from not only geographic objects and the relationships between them, but also ontologies (especially for the types of objects), the names of places and mathematical models as shown in Figure 8.

From these observations, to describe a set of geographic rules, in this new language, two levels will be necessary, the rules themselves and their sets. As a first step to simplify the problem, it will not handle metarules and their consequences.

![Figure 7. Modeling of geographic rules.](image-url)

As explained in Figure 8, three parts must be analyzed. In the "antecedent" part, it will be necessary to give the jurisdiction of the rule. This could be a place defined by a polygon with its coordinates, a toponym, or a Boolean combination of toponyms.

It will be important to include the elements common to all rules included in this set, i.e., the name of the set, language,
jurisdiction, ontology and gazetteer. It could include references to other sets of rules provided that the language, ontology and gazetteer are compatible. Then will be given the rules themselves. This jurisdiction must be included within the jurisdiction of the rule set. Then a list of relevant geographical objects and possibly Boolean conditions will follow.

The part “implication” should indicate whether this rule is imperative, fuzzy, frequent or good type practice.

But first and foremost, it is necessary to continue this analysis in order to introduce other cases and thus enrich the semantics of the geographic rules. Then, it will be possible to propose a robust, consistent and effective formalism for representing the geographic rules and enable them. It will also need to define the precise specifications of the actions to be undertaken for the treatment of the modulated implications.

REFERENCE


DOI reference number: 10.18293/DMS2016-015