



NRC Publications Archive Archives des publications du CNRC

Geospatial-Enabled RuleML in a Study on Querying Respiratory Disease Information

Gao, Sheng; Boley, Harold; Mioc, Darka; Anton, Francois; Yi, Xiaolun

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. /
La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version
acceptée du manuscrit ou la version de l'éditeur.

Publisher's version / Version de l'éditeur:

*Proceedings of the 3rd International Symposium on Rules, Applications and
Interoperability, 2009, 2009-11-07*

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=83af6777-7dec-4c31-8571-721d8f939560>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=83af6777-7dec-4c31-8571-721d8f939560>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the
first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la
première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez
pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



Geospatial-Enabled RuleML in a Study on Querying Respiratory Disease Information

Sheng Gao¹, Harold Boley², Darka Mioc^{1,3}, Francois Anton⁴, Xiaolun Yi⁵

¹ GGE, University of New Brunswick, Fredericton, NB, Canada

² Institute for Information Technology, NRC, Fredericton, NB, Canada

³ National Space Institute, Technical University of Denmark, Denmark

⁴ Department of Informatics and Mathematical Modelling, Technical University of Denmark

⁵ Service New Brunswick, Fredericton, NB, Canada

Abstract. A spatial component for health data can support spatial analysis and visualization in the investigation of health phenomena. Therefore, the utilization of spatial information in a Semantic Web environment will enhance the ability to query and to represent health data. In this paper, a semantic health data query and representation framework is proposed through the formalization of spatial information. We include the geometric representation in RuleML deduction, and apply ontologies and rules for querying and representing health information. Corresponding geospatial built-ins were implemented as an extension to OO jDREW. Case studies were carried out using geospatial-enabled RuleML queries for respiratory disease information. The paper thus demonstrates the use of RuleML for geospatial-semantic querying and representing of health information.

1 Introduction

Geospatial location provides a solution to link multiple sources in the same area. The spatial component of health data can show the geographical distribution of disease outbreaks, hospitals, air quality, and census. Basic geometric information of location is recorded in spatial data collections, using spatial reference and coordinate arrays. Utilizing spatial information allows the spatial analysis and visualization of health data. For example, with the geometric information of the Georges L. Dumont Hospital in Moncton and the New Brunswick Route 15, the neighboring spatial relationship between them can be deduced. The Semantic Web aims to improve machine understanding of Web-based information and its effective management. By employing Semantic Web (e.g., Web rule) techniques, part of the meaning of the information can be captured by machines, thus enabling more precise information queries and interoperation. To enhance the ability to query health information, its spatial component can also be represented and deduced by rules.

The Semantic Web environment, in which data are given well-defined meaning, can facilitate health data query and knowledge discovery. Similar to the non-spatial attributes of data, the spatial attributes can also be represented in the Semantic Web. The use of spatial information in Semantic Web can support dynamic spatial

relationship discovery for health data, and furthermore, new concepts and new instances can be generated. For example, from the locations of infectious disease outbreaks, we can determine the sensitive areas that are within a certain distance from the disease outbreak locations. Because of the advantages in supporting the representation of a spatial component, we endeavor to include spatial information in the Semantic Web environment to enhance the ability to query and to represent health data. This paper builds on and extends the eHealthGeo results in Gao et al., [1], and includes the geometric representation in RuleML to enhance information reasoning and inference.

2 Semantic Web and Geospatial Semantics

Semantics-level interoperability among heterogeneous information sources and systems can be achieved by the Semantic Web. According to Sheth and Ramakrishnan [2], three kinds of important applications of the Semantic Web are (1) semantic integration, (2) semantic search and contextual browsing, and (3) semantic analytics and knowledge discovery. Ontologies, as shared specifications of conceptualizations [3], constitute an important notion in the Semantic Web. Many XML-based languages, such as RDF(S) and OWL, have been developed for the representation of ontologies. Description Logic (DL) is usually used to represent ontologies. When concepts are defined using ontologies, three types of relation can be distinguished: taxonomic, functional, and partonomic [4]. With the meaning and relations of concepts defined by ontologies, semantic data classification, integration, and deduction can be implemented. One limitation of DL is that it is impossible to represent relationships between a composite property and another (possibly composite) property in the ontology representation; however, the use of rules can establish more complex relationships between properties [5]. Rules encode machine-interpretable conditional knowledge (“if ... then ...”) for automatic reasoning [6]. Rules can describe concepts by using the relation of instances through different property paths. Many different kinds of approaches in combining ontologies and rules have been surveyed (see [7]). RuleML [8] is the de facto open-language standard for Web rules.

Spatial relations can exist between two spatial objects (concepts or instances), and exploring them can advance information query and discovery. Three types of major spatial relations between spatial objects are topological, direction, and metrical relations [9]. Topological relations formalize the notion of neighborhood; directional relations require the existence of a vector space; and metric relations are measuring distances. Topological relations are invariant under continuous translations while directional and metric relations may change during these translations. A well-known method by which to formalize topological relationship between spatial objects in two-dimensional space is the Nine Intersection Model (9IM), developed by Egenhofer, that considers boundaries, interiors, and complements intersection of two spatial objects [10]. The further improved model, the Dimensionally Extended Nine Intersection Model (DE-9IM), considers the 9IM of two spatial objects with the dimensions of -1 (no intersection), 0, 1, or 2 [11, 12]. The commonly known

topological predicates described by the DE-9IM include overlaps, touches, within, contains, crosses, intersects, equals, and disjoint.

With possible spatial relations existing in the data, several studies have been done on the capture of geospatial semantics for facilitating data integration, query, and discovery. Kieler [13] discussed the feasibility of identifying semantic relations between different ontologies by exploring the geometric characteristics of the instances. To represent spatial relations, the explicit storage or dynamic computation of spatial relations is possible. Explicating all the possible spatial relations between every two spatial objects is usually not necessary. While the weakness of dynamic computation is that it is time-consuming, the weakness of explicit storage requires significant storage space and involves reliability issues because of the imprecise nature of relations [14]. Klien and Lutz [15] illustrated the definition of geospatial concepts based on spatial relations and automatic annotation of geospatial data using a reference dataset. The annotation process uses DL in reasoning and focuses on the concept level. Smart et al. [16] distinguished multi-representations, implicit spatial relations, and spatial integrity of geospatial data, claiming that rule expression for geo-ontologies needs to consider spatial reasoning rules and spatial integrity rules. Kammersell and Dean [17] proposed GeoSWRL, which is a set of geospatial SWRL built-ins. GeoSWRL allows users to include spatial relation operators in queries; however, spatial data representation and processing abilities are not fully integrated in the GeoSWRL system.

In addition, spatial operations can generate new spatial objects from existing spatial objects, such as spatial intersection and spatial union. Because rules are able to describe relations through complex property paths, it would be feasible to represent spatial operations and spatial relations of geospatial objects as rules in knowledge deduction. Cartographic principles can also be applied as rules in the deduction. In this paper, we not only enable geometric representation support for RuleML reasoning, but also apply ontologies and rules in health information reasoning, query, and representation. The respiratory disease information queries are used as examples.

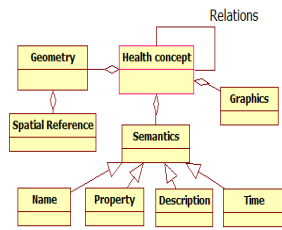


Fig. 1. Metamodel of health concepts.

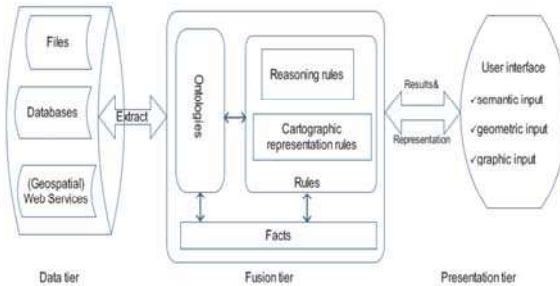


Fig. 2. Health data query and representation framework.

3 Framework for Health Information Query and Representation

Health concepts can be described with semantic, geometric, and (carto)graphic components, as shown in Figure 1. The semantic component deals with the definition of the concepts. The geometric component provides shapes to locate the concepts.

The graphic component solves the issues of how to represent these concepts through maps. For example, in the case of a hospital, the semantic component can describe its name and attributes, the geometric component can describe its polygon shape, and the graphic component can describe its map style. Moreover, relations, including non-spatial and spatial relations, exist between health concepts.

3.1 Framework

Semantic health data queries need to find data with corresponding semantic and geometric attributes. Cartographic representation of the query results allows users to visualize health information. Figure 2 describes the framework for the semantic health information query and representation, including a data tier, a fusion tier, and a presentation tier. (1) Data tier. The health data are obtained from various organizations through files, databases, or (Geospatial) Web Services. Following the ontology implementation, data are extracted to the knowledge base as facts. (2) Fusion tier. The fusion tier contains ontologies, facts, and rules. It queries and fuses semantic, spatial, and cartographic information for representing health data homogeneously. The ontologies are the representation of health concepts and their relationships in the semantic, geometric, and graphic dimensions. Facts are generated from various health data and existing knowledge about health. Rules, supported by ontologies and facts, deduce health information and present the information to users. Two types of rules are considered: reasoning rules and cartographic representation rules. (3) Presentation tier. The user interface allows the input of semantic, geometric, and graphic criteria to find health information.

3.2 Ontologies and Rules in Health Data Fusion

A. Ontologies

Ontologies can be utilized to connect various concepts (e.g., subconcepts and superconcepts). Depending on the requirements, different application ontologies exist in health applications. To facilitate health data exchange and query, a global ontology can improve interoperability. Three types of ontologies are important in querying and representing health data: health domain ontologies, geometric ontologies, and cartographic ontologies.

Health domain ontologies are used for the definition of health information models, concepts, and terminologies. Many standards exist in this field, such as Health Level 7 (HL7), SNOMED-CT, and International Classification of Diseases (ICD-9). Geometric ontologies should be able to describe basic geometry types, such as point and polygon. The European Petroleum Survey Group (EPSG, <http://www.epsg.org/>) coordinate system codes are widely used in the exchange of geospatial data over the Internet. Cartographic ontologies deal with the styles in representing information. For instance, the symbol of hospitals can be represented as point graphics to show the location of hospitals. With the existence of health domain ontologies, geometric ontologies, and cartographic ontologies, the application ontology definition can easily link to these semantic, geometric, and cartographic elements.

B. Rules

Based on the ontologies and facts, rules can define and deduce new information. Besides non-spatial attribute rules, spatial rules can also be applied in this framework. Although the definition of geometric ontologies follows the same methodology for non-spatial ontologies, the inference of geometric relations is different. The utilization of geometries can incorporate the spatial analysis and cartographic representation abilities in rules. Two types of rules are distinguished: reasoning rules and cartographic representation rules.

Reasoning rules cover semantic matching, spatial relation operators, spatial operations, and cartographic comparison of data. (1) Semantic matching rules deal with the domain knowledge for understanding health data. For instance, the manifestation of several symptoms could determine that a patient may have caught a disease. (2) Spatial relation rules are used to determine the topological, directional, and metric relations between geospatial components. For instance, rules can be used to evaluate the direction and distance from the location of an emergency to hospitals. (3) Spatial operation rules can generate new concepts and instances from existing health data. For example, spatial union can combine data from neighboring regions to assist in the comparison of disease outbreaks. (4) Cartographic comparison rules are able to fuse different cartographic representations into a homogeneous form.

Cartographic representation rules focus on the distribution of information to users more efficiently and effectively. Map scale is of great significance in the geometric representation of a concept. For example, a hospital will be shown as a polygon in large scale representation and as a point in small scale representation. Cartographic rules include concept-based rules, attribute-based rules, scale-based rules, priority-based rules, and cartographic generalization rules. (1) Concept-based rules determine graphic styles based on the health concept semantics. For example, standard symbols exist in representing the concepts in national or provincial cartographic design. (2) Attribute-based rules classify health concepts based on their attributes. For example, pie charts can show the age distribution of people in each health region. (3) Scale-based rules are essential in determining what information is represented based on scales. A concept can be stored with multi-representation in the data, and scale can be used to select the optimal representation. (4) Priority-based rules emphasize high priority information. (5) Cartographic generalization (simplification, exaggeration, and displacement) rules allow the dynamic generalization of spatial information.

4 Design and Implementation

4.1 Geospatial Support for RuleML Deduction

OO jDREW is an open source RuleML engine which is used in this study because it supports RuleML's Naf Hornlog sublanguage and backward/forward reasoning [18]. RuleML's POSL presentation syntax is employed in the following. To use spatial information in the reasoning process, the representation of spatial information in the RuleML engine is needed. Therefore, a geometric ontology is designed to support basic geometry types: point, linestring, polygon, multipoint, multilinestring,

multipolygon, and multimix, as shown in Figure 3. A polygon can have an out boundary and many inner holes (inner boundaries). Multipoint, multilinestring, and multipolygon can have one or more points, linestrings, and polygons respectively. Multimix contains collections of points, linestrings, and polygons. Figure 4 lists examples of how to represent each geometry type. Coordinate reference systems are specified with EPSG codes, and coordinates are recorded in the order of (x1,y1,x2,y2,...). With the specification of geometries, the spatial operation (union, buffer, convexhull, difference, distance, intersection) and spatial relation operators (touches, contains, within, crosses, equals, overlaps, intersects, covers, coveredby, disjoint, iswithindistance) can be incorporated into rules.

Based on this design, a geometry type was added and a parser implemented for parsing geometries in OO jDREW. For the geospatial operations and spatial relation operations, the JTS Topology Suite is used in this study. The JTS is an open source Java API for two-dimensional spatial predicates and functions, using the DE-9IM model [19]. Several geospatial built-ins, such as the `gpred_intersects`, `gpred_within` and `gfunc_intersection` built-ins, were created using the JTS library. The `gpred_intersects` built-in checks whether two geometries intersect or not; the `gpred_within` built-in checks whether a geometry is inside another geometry; the `gfunc_intersection` built-in computes the intersection of two geometries.

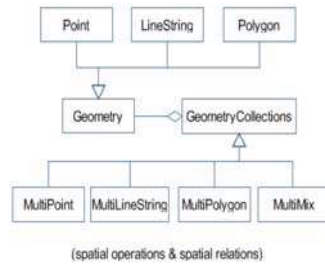


Fig. 3. Geometry type designed for RuleML.

Point:	geo[EPSG4326, point[0.0,0.0]]:Geometry
LineString:	geo[EPSG4326, linestring[0.0,0.0,1.0,1.0]]:Geometry
Polygon:	geo[EPSG4326, polygon[outboundary[-1.0,-1.0,1.0,1.0,1.0,1.0,-1.0,-1.0,-1.0], innerboundary[-0.5,-0.5,-0.5,0.5,0.5,0.5,0.5,-0.5,-0.5,-0.5]]]:Geometry
MultiPoint:	geo[EPSG4326, multipoint[point[0.0,0.0],point[1.0,1.0]]]:Geometry
Multi- LineString:	geo[EPSG4326, multilinestring[linestring[0.0,0.0,1.0,1.0], linestring[2.0,2.0,5.0,5.0,3.5,4.6]]]:Geometry
Multi- Polygon:	geo[EPSG4326, multipolygon[polygon[outboundary[0.0,0.0,0.0,1.0,1.0,1.0,0.0,0.0, .0,0.0]], polygon[outboundary[-1.0,-1.0,1.0,1.0,1.0,-1.0,-1.0,-1.0]]]]]:Geometry
MultiMix:	geo[EPSG4326, multimix[point[0.0,0.0], polygon[outboundary[-1,-1, -1,1,1,-1,-1,1]]]]:Geometry

Fig. 4. Examples of geometry representation.

4.2 Data Sources and Ontology Definition

The health data used in this study were collected from different organizations, such as New Brunswick Lung Association, Service New Brunswick, Statistic Canada census, and Statistic Canada community health survey. Respiratory disease data are used as examples in this study. Following the disease taxonomy of respiratory diseases in the International Classification of Diseases (ICD-9), we created an ontology for respiratory diseases. Respiratory disease data are from hospital patient incidents, which record the time, postcode, disease diagnosis category, age, and gender. Different data could be collected in various spatial boundaries. Taking this study as an example, the disease rate data from the Statistics Canada community health survey were collected at *Health region* and the income data from Statistics Canada census were collected at *Census division*. From these data, the application ontologies of our case study were generated. We also created entities, such as *Health event*, *Hospital*, *Health region*, *Census division*, *Postcode*, *Disease rate*, and *Income*.

Health event can describe a variety of cases, such as patient incidents, health training services, etc. The following properties (POSL: “->”) associated with health events are shown here: the involved participants’ age and gender, the admit date, the disease category diagnosis, and the postcode. Example with a variable (POSL: “?”):
health_event (disease->?:Influenza_with_pneumonia; age->88:Integer; gender->Female;
postcode->E1C; admitdate->date[2000:Integer,1:Integer,1:Integer]).

Hospital introduces general information about hospitals, with attributes: name, address, city, province, telephone, and geometry. Example:

hospital (name->Dr_Everett_Chalmers_Hospital; address->700_Priestman_St; province->NB;
city->Fredericton; telephone->5064525400; totalbeds->384:Integer; geometry->...).

Health region and *Census division* are two kinds of administrative boundaries. They have name, area, perimeter, and geometry attributes. Example:

health_region (name->Health_region_1; area->10455463176.5:Real;
perimeter-> 844278.079968:Real; geometry->geo[EPSG4326,
multipolygon[polygon[outboundary[...],...]]:Geometry).

Postcode shows the central location of the three-digital postcodes. Example:

pcode3 (name->E1A;
geometry->geo[EPSG4326,point[-64.7078903603,46.0967513316]]:Geometry).

Disease rate and *Income* show the value associated with the geometry name, statistic method, and year. Example:

disease_rate (disease->?:Asthma; geometryname->Health_region_1; statistics->average;
year->2003:Integer; rate->0.104:Real).

income (geometryname->Saint_John_County; statistics->average; year->2003:Integer;
incomevalue->32748.56028:Real).

4.3 Scenarios

Case 1. With the collected health events, it is possible to find disease cases fulfilling semantic and geometric requirements. Since disease cases include outbreak locations using postcodes, geospatial semantic query of diseases can discover whether the location of a postcode is inside any spatial boundary. The following *disease_locator* rule queries a patient’s age, gender, and postcode within a certain health region, disease category, age type, and period:

```
disease_locator (healthregionname->?name; disease->?disease:Respiratory_diseases;
    startdate->?startdate; enddate->?enddate; age->?age;
    age->?age:Integer; gender->?gender; postcode->?postcode) :-
    health_event (disease->?disease:Respiratory_diseases; age->?age:Integer;gender->?gender;
        postcode->?postcode; admitdate->?date),
    age (age->?age; age->?age:Integer),
    earlier (?date, ?enddate), later (?date, ?startdate),
    health_region (name->?name; geometry->?hrgeometry:Geometry!),
    pcode3 (name->?postcode; geometry->?pcgeometry:Geometry!),
    gpred_within (?pcgeometry:Geometry, ?hrgeometry:Geometry).
```

The *disease_locator* rule conjoins several subqueries for the semantic query of disease cases. The *earlier* and *later* queries search disease cases in which the admit date is between the start date and end date. The *age* query is used to determine to which age group a certain age belongs. The *gpred_within* built-in query is used to locate postcodes in health regions.

Case 2. Since data collected from different organizations may use different kinds of spatial boundaries, the ability to integrate those data is useful. New concepts and instances will be generated in the integration process. The below `disease_income_correlator` rule figures out the intersection between disease rate and income. For example, a user would like to know those spatial areas where the asthma disease rate is higher than 0.1 and the average income is above \$30,000 in 2008.

```
disease_income_correlator (disease->?disease:Respiratory_diseases; year->?year:Integer;
    minincome->?minincome:Real;minrate->?minrate:Real; geometry->?geometry:Geometry):-
    disease_rate(disease->?disease:Respiratory_diseases; year->?year:Integer;
        geometryname->?dgeometryname; rate->?rate:Real!),
    income (geometryname->?igeometryname;
        year->?year:Integer;incomevalue->?incomevalue:Real!),
    health_region (name->?dgeometryname;geometry->?hrgeometry:Geometry!),
    census_division (name->?igeometryname;geometry->?cdgeometry:Geometry!),
    greaterThan (?rate:Real,minrate:Real), greaterThan (?incomevalue:Real,minincome:Real),
    gfunc_intersection (?geometry:Geometry,?hrgeometry:Geometry,?cdgeometry:Geometry).
```

Case 3. To provide better representation of the information to users in the query process, it is beneficial to allow users to define queries with semantic, geometric, and graphic requirements. For example, a user wants to get the asthma rate in 2008 (semantic) in a spatial boundary `geometry1` (geometric) with a graduated color ramp1 (graphic). Firstly, the user can define geometric and graphic requirements. The graphics here use graduated color with two categories. One category ranges from 0.0 to 0.2 in green; the other category ranges from 0.2 to 1 in red:

```
geometries (geometryname->geometry1;
    geometry->geo[EPSG4326, polygon[outboundary[...]]:Geometry).
graduated_colors (name->ramp1; startvalue->0.0:Real; endvalue->0.2:Real; color->0x00FF00).
graduated_colors (name->ramp1; startvalue->0.2:Real; endvalue->1:Real; color->0xFF0000).
```

Then, the user can use the `disease_rate_finder` rule to query disease rates. This rule deduces the graphics for disease rate instances within specified geospatial boundaries.

```
disease_rate_finder (disease->?disease:Respiratory_diseases; geometryname->?geometryname;
    rampname->?rampname; year->?year:Integer;
    geometryname->?healthregionname; color->?color):-
    disease_rate (disease->?disease:Respiratory_diseases; geometryname->?healthregionname;
        year->?year:Integer; rate->?rate:Real!),
    health_region (name->?healthregionname; geometry->?hrgeometry:Geometry!),
    geometries (geometryname->?geometryname; geometry->?geometry:Geometry),
    graduated_colors (name->?rampname;startvalue->?startvalue:Real;
        endvalue->?endvalue:Real; color->?color),
    greaterThanOrEqual (?rate:Real,?startvalue:Real), lessThan (?rate:Real,?endvalue:Real),
    gpred_intersects (?geometry:Geometry,?hrgeometry:Geometry).
```

Case 4. Depending on the scale of representation, the cartographic information represented to users could be different. For example, between the scale of 1:1,000 to 1:1, hospitals are shown as polygons. Between the scale of 1:1,000,000 and 1:1,000, hospitals are shown as points. With the scale smaller than 1:1,000,000, hospitals disappear. In this case, we can add a minimum scale and maximum scale in the hospital entity for the cartographic representation purpose. The following sample fact shows one geometric representation for the multi-representation of a hospital:

```
hospital (name->Dr_Everett_Chalmers_Hospital;address->700_Priestman_St;
    city->Fredericton;province->NB; telephone->5064525400; totalbeds->384:Integer;
```

```
minscale->0.001:Real; maxscale->1:Real; geometry->geo[EPSG4326,
  polygon[outboundary[66.65654990041024, 45.93896756130009,...]]:Geometry).
```

With a scale input by users, this rule finds the optimal representation of hospitals:

```
hospital_locator (name->?name; scale->?scale:Real; geometry->?geometry:Geometry;
  totalbeds->?totalbeds:Integer):-
  hospital (name->?name; geometry->?geomery:Geometry; minscale->?minscale:Real;
    maxscale->?maxscale:Real; totalbeds->?totalbeds:Integer!),
  lessThan (?scale:Real,?maxscale:Real), greaterThanOrEqual (?scale:Real,?minscale:Real).
```

Complex queries can then be supported by combining the available predicates exemplified in the above cases. For example, users can define the spatial area of interest (using customized geometries of Case 3). Then they may like to know where high disease rate and low income values exist within the area of interest (using the disease_income_correlator of Case 2). After that, users can get the information about hospitals within the previously determined high disease rate and low income areas in a certain map scale (using the hospital_locator of Case 4 and gpred_within of Case 1). All these steps can be chained into complex rules to formalize user queries.

5 Discussion and Conclusions

Our health data query and representation framework provides a solution for health experts to express knowledge as ontologies and rules (regarding semantic, geometric, and graphic dimensions) in health information integration and representation. The use of rule techniques enables health experts to exchange reasoning and representation rules on the Web. Much research has been done on semantic health information integration and query using non-geospatial information in the reasoning. However, fewer investigations utilize geometric information for dynamic spatial reasoning in this process. This research builds an integrated system that supports geospatial-enabled semantic health information retrieval. A basic geometric ontology is designed for the spatial component representation. Spatial operations and spatial relations are expressed in RuleML for knowledge representation and deduction. Basic geometries, spatial operations, and spatial relation operators for RuleML are enabled through the extension of the OO jDREW engine. This implementation thus facilitates semantic health data integration and query with the use of both non-spatial and spatial operations and relations. Complex queries and reasoning processes can be implemented to allow the use of semantic, geometric, and graphic dimensions.

The current system implementation uses the interface of OO jDREW in the query process. More customized user interfaces in the presentation tier will be implemented to facilitate health information query and representation. Moreover, as dynamic spatial reasoning and computation has demanding time and memory requirements, the balance between caching computed results and dynamic spatial computation need to be optimized for efficient health information querying. In addition, the ontology designed in this study is based on the data collected. Its implementation supports the transformation of various health data to facts in the knowledge base. To further improve data integration and query, upper-level or domain-level ontologies need to be investigated. Various health and geospatial standards can be taken into consideration, such as the HL7 ontology and Open Geospatial Consortium (OGC) standards.

With the rapid growth of health data, the semantic query of health information becomes increasingly important for health practitioners in understanding health phenomena. The support of spatial operations and spatial relation operators by rule systems is useful for health data integration, query, and representation. In this study, an integrated semantic system has been built to support geospatial-enabled query and reasoning of health information. With the use of RuleML, we have enabled geometry types, spatial operation rules, and spatial relation rules for health information query. The case scenarios in this study demonstrate the benefits of including a geospatial component in semantic health data query, permitting the fusion of various kinds of data in the semantic, geometric, and graphic dimensions. This research fosters the use of ontologies and rules in representing these dimensions of public health information. It facilitates the deduction of information collected by different health organizations. Our future work will be devoted to the exploration of ontologies and rules for further semantic integration, query, and representation of health information.

References

1. Gao, S., Mioc, D., Boley, H., Anton, F., Yi, X.: A RuleML Study on Integrating Geographical and Health Information. LNCS, 5321, 174--181 (2008)
2. Sheth, A.P., Ramakrishnan, C.: Semantic (Web) Technology In Action. IEEE Data Engineering Bulletin. 26, 4, 40--48 (2003)
3. Gruber, T.R.: A Translation Approach to Portable Ontology Specifications. International Knowledge Acquisition. 5, 2, 199--220 (1993)
4. Luscher, P., Burghardt, D., Weibel, R.: Ontology-Driven Enrichment of Spatial Databases. In: 10th ICA Workshop on Generalisation and Multiple Representation, Moscow (2007)
5. Antoniou, G., Damasio, C.V., Grosz, B., Horrocks, I., Kifer, M., Maluszynski, J., Patel-Schneider, P.F.: Combining Rules and Ontologies: A survey. (2005)
6. Boley, H.: Are Your Rules Online? Four Web Rule Essentials. In: Proc. Int'l RuleML Symposium on Rule Interchange and Applications. LNCS, 4824, 7--24 (2007)
7. Bruijn, J.D.: ONTORULE: ONTOlogies meet business RULEs, State-of-the-art survey of issues. <http://ontorule-project.eu/deliverables-and-resources?func=fileinfo&id=1>
8. The Rule Markup Initiative, <http://www.ruleml.org/>
9. Rashid, A., Shariff, B.M., Egenhofer, M.J., Mark, D.M.: Natural-Language Spatial Relations between Linear and Areal Objects. Int. J. Geogr. Inf. Sci., 12, 215--245 (1998)
10. Egenhofer, M. J.: Reasoning about Binary Topological Relations. LNCS, 143--160 (1991)
11. Clementini, E., Felice, P.: A Comparison of Methods for Representing Topological Relationships. Inf. Sci., 80, 1--34 (1994)
12. Clementini, E., Felice, P.: A Model for Representing Topological Relationships between Complex Geometric Features in Spatial Databases. Inf. Sci., 90, 121-136 (1996)
13. Kieler, B.: Derivation of Semantic Relationships between Different Ontologies with the Help of Geometry. In: Workshop at AGILE 2008, Girona, Spain (2008)
14. Jones, C.B., Abdelmoty, A.I., Fu, G.: Maintaining Ontologies for Geographical Information Retrieval on the Web. LNCS, 2888, 934--951, (2003)
15. Klien, E., Lutz, M.: The Role of Spatial Relations in Automating the Semantic Annotation of Geodata. LNCS, 3693, 133--148 (2005)
16. Smart, P.D., Abdelmoty, A.I., El-Geresy, B.A., Jones, C.B.: A framework for combining rules and geo-ontologies. In: First International Conference on Web RR Systems. (2007)
17. Kammersell, W., Dean, M.: Conceptual Search: Incorporating Geospatial Data into Semantic Queries. Terra Cognita - Directions to the Geospatial Semantic Web (2006)
18. OOjDREW, <http://www.jdrew.org/ojdrew/>
19. JTS Topology Suite, <http://www.vividsolutions.com/jts/jtshome.htm>