Construction of Knowledge Graph Based on Geographic Ontology

Chunfang Guo a, Tingting Xu b, * and Liu Liu c

School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China

a chfguo@bjtu.edu.cn, *, b Corresponding author email: sdxutingting@163.com, c 15125391@bjtu.edu.cn

Abstract. Here are massive archival data in the geography of surveying and mapping industry, which provides important guarantee for the national economy and social development. Efficient storage and retrieval systems for archival data must be built, in order to make effective use of these data. In this paper, OWL language is used to establish the ontology relation model of computer language, and then the ontology reasoning mechanism is studied on this basis. The reasoning model is educed with open source project Jena inference engine. Then the knowledge map of regional geographic ontology is constructed with the powerful ontology graphical representation ability of graph database. Finally, the search system is developed based on the java language, achieving the required retrieval requirements.

1. Introduction

In the organization and expression ability of Geographic information system in the geographical space knowledge is weak, which owns rich data source but lack of knowledge. Most of spatial knowledge base based on the geometric ontology is loosely distributed, and its way of storage, organization and expression do not well distinguish geographic data and geographic knowledge, resulting defects of the knowledge expression and retrieval in the knowledge system. China's research on the knowledge map is still relatively backward, facing many challenges in the field of ontological exploration. In this paper, Jena reasoning machine is used successfully for deducing more information about the relationship between the ontology constructing the knowledge map of the example area and developed a set of retrieval tools through the modeling of regional spatial relations and hierarchical relations. It provides a certain reference to geography and enterprise optimization of retrieval structure in improving the retrieval efficiency.

2. Summary of citing research

In 1995, when Egenhofer and Mark began to study geo-information science, ontology was first introduced into geographic information science [1]. In the research schedules given by UCGIS in 2002, the geography ontology is ranked among the top ten long-term research challenges and has become the focus of research in the field of geographic information. For the study of geographical ontology, domestic scholars have done some related research and achieved some results. Wuhan University Cui has published the geography ontology papers; the paper has in-depth study on the ontology of the
geographic information system semantic integration and interoperability [2]. In 2006, the idea of linking data (calling linked data) is proposed, calling for the promotion and improvement of relevant technical standards such as URI (uniform resource identifier), RDF (resource description framework), OWL (Web ontology and to prepare a semantic network research boom, knowledge map technology is the existing semantic network technology sublimation. Sillanp complete the conflict research literature visualization, through 40 high-quality journals in 1300 papers analysis in 2009 and agree that conflict research can be divided into four major areas and the knowledge map provides new content and application perspective [3]. Janssens complete the visualization of the library and information science knowledge structure in 2013 [4]. The specific use of five representative journals including nearly 1,000 papers composed of six categories of library materials and information science.

3. Formal expression of geographic ontologysummary of exiting research

3.1. The Concept, Classification and Ontology of Geographical Objects
Geographical concept types are extremely complex. It not only includes physical geography, but also human geography [5]. Its complexity to the construction of the geographical ontology has brought great difficulties. The classification of geographical concepts can be a good solution to the complex concept of geographical concepts. According to the relevant information, this paper divides the geographical environment into physical geography and human geography.

OWL has strong ontology attribute expression ability. It can define the ontology and its attribute relation. This paper uses OWL language to represent geographic ontology based on the classification of geospatial. The classification of concept in geological ontology can be represented by OWL class [6]. A variety of conceptual hierarchy in the geographical ontology can be represented by subClassof. The geographicObject is the base class of the whole concept, including NatureGeography and AnthropoGeography.

The human geography is defined as the class AnthropoGeography, which has the characteristics of disagreement with natural geography, and is a direct subclass of the base class GeographicObject, as with natural geography.

3.2. The Spatial Relationship of Geographical Ontology and Its Expression
The most important feature of the geographical ontology is that it has spatial feature. The spatial relationship between each entity in space is called spatial relation. The relationship between spatial entities not only includes logical and numerical relations, but also includes topological relations, distance relations and spatial relations.

Topological spatial relations are mainly concerned with the connection, inclusion, adjacency and phase separation of geographic entities [7]. The relationship between spatial topological relations is as follows:

N-intersection topological relations: This kind of topological analysis method divides the spatial object into several constituent parts, and then judges the spatial relation between the objects through the intersection relation among the objects.

RCC model theory: The relationship in the RCC model has reflexivity and symmetry. For example, we use C (x, y) to denote the adjacency between region x and region y, and we can use the existence relation of the point in the region and the relationship to define the topological relationship between the two regions x, y.

Representation of topological relations based on hierarchy (HBM): This method uses the hierarchical structure of the tree diagram to represent the topology between the spatial objects. It is divided into disjoint, touch, overlap, in, and equal.

The method (HBM) has a strong ability to describe topological spatial relations, providing better features for topological description models [8].
In this section, HBM hierarchical topological relation classification is used to describe the topological spatial model of the geographic ontology. The root attribute of the topological relation is established as al.

The other topological relations attribute is derived according to the tree hierarchy relationship established above. As the Table I shows:

| Topological Relations | Features         |
|-----------------------|------------------|
| disjoint              | Symmetric        |
| equal                 | Symmetric Transitive |
| touch                 | Symmetric        |
| overlap               | Symmetric        |
| in                    | Transitive       |

From the above figure we can clearly see that the "in" attribute is subordinate to "unequal" and is one of the sub-attributes of "unequal", and its domain and range are both GeographicObject and TransitiveProperty represents the in-pass attribute.

When dealing with spatial azimuthal relationship, the conic and azimuthal relationship determinations are two of the most widely used descriptive models [9]. In this paper, the relationship model will focus on them.

The main idea of the cone model is to divide the area around the space target into multiple regions according to the direction. The spatial relationship between the two objects is described by the intersection of the target object and the divided area.

As it shown in the FIGURE 1, the four-direction conical model is centered at 0:00, and the spatial target and its surrounding area are divided into four equal parts according to the north-east direction axis. Among them, N, E, S and W denote north, east, South and west of the four directions of the region, when the target object and the current reference object orientation, only four regions and the target object can be intersected, as shown in Fig.1, the center of the object A O points are divided into four directions: the target object B falls in the E region, B is the east of A, that is, dir (A, B) = E,

![Figure 1. Tetragonal conical model](image)

The eight-direction model is similar to the basic idea of the four-direction model. It is a reference object center O as a reference, the target object and its surrounding area are divided into four or eight aliquots, and then the target object and these areas to determine the relationship between the intersection.

The triangulation model, starting at a point in the spatial reference object, makes two rays along a certain direction to form a triangular region, using the intersection of the region and the target object to represent the azimuthal relationship between them.

MBR model is generally used as a filter to determine the topological relationships between objects [11]. When it is described as a spatial object azimuthal relationship, it includes nine directional relationship model, that is, the Chinese character "well" model.
According to the MBR model, the directionalRelation is the root attribute of the azimuth relation in the ontology expression attribute, and the other topological attributes are derived from the directionalRelation.

The above-mentioned azimuthal relationship OWL expression is a formal expression of "NE (Northeast)", where subPropertyOf means that NE belongs to directionalRelation, domain and range are GeographicObject, which means that nw belongs to GeographicObject and its value range is also for GeographicObject.

3.3. The Hierarchical Relationship of Geographical Ontology and its Expression

The relationship between geography ontology is mainly based on spatial relations and hierarchical relationships [12]. Inheritance is a kind of hierarchical relationship, that is, A is a subclass of B or A is-a-kind-of B, in the field of geographical ontology, not only the class has this inheritance relationship, the class attributes also exist Inheritance relationship.

According to the classification of the definition of the administrative region will be divided into national, provincial / municipalities, provincial cities / prefecture-level cities, regional / autonomous prefectures, municipal districts / cities / county-level cities, counties / autonomous counties / flag government, town, Township, village and other levels as shown in the picture shows the administrative division of Jilin Province level diagram.

OWL use isSubRegionOf and hasSubRegion to express the ontology hierarchy in the geographic object, where the domain and range of isSubRegionOf and hasSubRegion are the base classes of the geographic concept [13]. At the same time both reverse each other inverseOf, are also passable property TransitiveProperty.

4. Construction and retrieval realization of knowledge mapping of geography ontology in jilin region

4.1. Writing the Sample Area Ontology Data

It is necessary to import the ontology class and the relation network of each ontology into the graph database In order to construct the knowledge map of the geographic ontology. In this paper, we will use Jena reasoning machine, manually set the reasoning rule base, reason the existing ontology based on spatial relations and hierarchical, and reasoned out more ontology network. This section uses the Cypher query language to input operations, localizing the ontology and ontology network and the reasoning ontology network.

4.2. The Construction of Knowledge Map Based on Graph Database

After the geographic ontology and ontology relations are stored in the database, we can see the connection between the ontologies and other ontologies by intuitive visualization based on the powerful visualization tools of the graph database.

4.3. The Development of an Example Area Retrieval System

Because that this chapter mainly interprets the realization of knowledge map construction and retrieval system from the functional point of view, because the development process is not the focus of this paper, the development process and the required configuration environment are briefly described, and the realization of the search results is analyzed emphatically.

This paper aims to establish a sample area of knowledge map, and use the established knowledge map on the relevant geographic ontology retrieval, the system development environment is as follows:

- Java version jdk1.7, the system for Windows 7 Professional Edition.
- The framework for the development is Spring, SpringMVC, Mybatis and other common development framework java.
- Database using Neo4j non-relational graph database.

The system design frame is as follows:
Persistence layer Neo4j docking with the retrieval system

In addition to establishing a simple synonym collection, this paper also sets up the relationship between the words and the words (or phrases), and forms the knowledge map database according to the tree structure, which is related to each other.

Enter the Songhua River in the Ontology search box, the ontology retrieval system will search the keywords based on the above knowledge map and read the relevant concepts from the Songhua River attributes. The system has achieved the following functions:

1) Implementation of the left sidebar modul.
   Related vocabulary module: knowledge map correlation rules and correlation threshold based on the search vocabulary (custom correlation threshold, if the correlation threshold is higher and relevant vocabulary is more, The correlation of ontology information tend to be small), in this case, the body related to Songhua River Basin is Changchun City, Jilin City, Yalu River
   2) Provide the retrieval module based on release time according to release date of the file
   3) Provide a search engine interface for keywords
   4) Retrieve the main implementation
   According to the keyword definition in the database attributes, return a brief introduction to the keyword

The system in ontology correlation diagram analyze the search words Songhua River and search out the relevant map nodes, respectively, through the relevant (Changchun City, Jilin City) and attribute-related (belong to the River)

Brief analysis of the number of archives in the administrative area of the Songhua River area and the comparison of the number of archives between the administrative regions

5) Implementation of right sidebar
   6) Provide a popular search, the current stage of search keywords to rank and provide users with popular search charts.
   My search will record the user search history information, lists the last two search information, user-friendly search history and the query ontology for comparison analysis.

5. Conclusion
This paper, using the java language, combines with the model OWL ontology and relational model and ontology model, the results of the reasoning and map database, constructs the geographic ontology knowledge map in Jilin area. On this basis, the retrieval system is developed, achieving the required retrieval requirements. This paper works out the final knowledge map based on the establishment and
reasoning of the relationship between the various ontology in the field of geographic information. However, the next important research direction is that how to access to the ontology and relationship information we need accurately, how to set more comprehensive reasoning rules, how to use Jena reasoning machine to reason out more information and how to filter the information we want and clean the information we do not need.

References

[1] Egenhofer and J. Max, Geographic information science, Springer, 2004, pp. 122-123.

[2] W.Cui, "Realization of Semantic Integration and Interoperability of Geographic Information System with Ontolog," University of Wu Han, 2004.

[3] A. Bhatnagar, W. Hogland, and M. Marques, "An overview of the modification methods of activated carbon for its water treatment applications," Chemical Engineering Journal, vol. 219, no. 3, pp. 99–511, 2013.

[4] Janssens and H Koen, Modern Methods for Analysing Archaeological and Historical Glass. 2013, pp. 49-65.

[5] D. Mark, "Geographic information science: Defining the field," Foundations of geographic information science, vol. 1, 2003, pp. 3-18.

[6] J. Euzenat and P. Valtchev, "Similarity-Based Ontology Alignment in OWL-Lite," European Conference on Artificial Intelligence, pp. 333-337, 2004.

[7] A. Buccella, A. Cechich, and P. Fillottrani, "Ontology-driven geographic information integration: A survey of current approaches," Computers & Geosciences, vol. 35, no. 4, pp.10-723, 2009.

[8] N. Cardoso, D. Cruz, and M. Chaves, "Using geographic signatures as query and document scopes in geographic," Advances in Multilingual and Multimodal Information Retrieval, pp. 802-810, 2008.

[9] M. Silva, B. Martins, and M. Chaves, "Adding geographic scopes to web resources," Computers, Environment and Urban Systems, vol. 30, no.4,pp. 378-399,2006.

[10] D. Mark, "Geographic information science," URISA Journal, vol.12, no. 1, pp. 45-54, 2000.

[11] J. Carlson, "Universal clutch puller and installer," Foundations of geographic information science, pp. 3-18, 2003.

[12] B. Smith and D. Mark, "Ontology with human subjects testing," American Journal of Economics and Sociology, vol. 58, no. 2, pp. 245-312, 1998.

[13] I. Abdelmoty, P. Smart, and C. Jones, "A critical evaluation of ontology languages for geographic information retrieval on the Internet," Journal of Visual Languages & Computing, vol. 16, no. 4, pp. 331-358, 2005.