Construction of Geo-Ontology Knowledge Base about Spatial Relations

Xue Zhang [1], Jun Xu[2]

State Key Laboratory of Resources Environment Information System,

Institute of Geographic Science and Natural Resources Research, Chinese Academy of Science

11A Datun Road, Beijing, China

[1]Zhangxue@lreis.ac.cn

[2]xujun@lreis.ac.cn

***Abstract*— The spatial relation analysis, query and reasoning in current geographic information systems usually generalize geographic objects into geometric points, lines and polygons. However, in the real world and human's cognition geographic objects are not simply geometric objects but spatially distributed objects with geographic semantics. If the geographic entities belong to different types, we may use different words to describe their spatial relationship although their shapes and geometric relationships are exactly the same. Aiming at above phenomenon, this paper analyzes what kinds of semantic information are involved in spatial relationship describes and queries. Based on the semantic analysis of geographic relations, an ontological knowledge base is established to store the knowledge related to spatial relations between geographic objects. The knowledge base is implemented with Protégé and OWL, and finally is connected to the spatial relation query system.**

***Keywords*— Spatial relations; geographic ontology; knowledge base; geographic semantic; OWL**

I. INTRODUCTION

Spatial relation is one of an important aspect of the geographic spatial information, so from cognition point of view to understand spatial relations and the formal expression of the spatial relations computer models plays a key effect in promoting the application of geographic information science. The spatial relation analysis, query and reasoning in current geographic information systems usually generalize geographic objects into geometric points, lines and polygons. However, in the real world and human's cognition geographic objects are not simply geometric objects but spatially distributed objects with geographic semantics. Therefore, in geographic information science, people need to pay more attention on the feature of the spatial relations, and on the semantic of the geographic objects. The general spatial relation query systems do not consider the semantic information; therefore the process and results of their queries about spatial relations are mechanical.

With the development of research on spatial relations, people take more and more factors into account, such as geographic semantic, human's cognition, and context-contingent. Some scholars have used these aspects to assist the spatial relation analysis, query and reasoning. Yao and Thill [1] researched how far is far in different context-contingent; Mark [2], [3], [4] focused on natural language understanding of the spatial relations between lines and regions; Jones [5], [6] researched on build geographical ontology for intelligent spatial search on the web. In China, Xu J [7], [8] researched on natural language understanding of the spatial relations between linear objects; Liu Y [9] focused on representation and reasoning of spatial relations in geographical space; BIAN Fu-ling [10] researched on build location ontology for geographic knowledge base; JING Dong-sheng [11] focused on geo-spatial information semantic expression and service based on ontology.

In the field of ontological and semantic research, one of a powerful tool is geo-ontology knowledge base which can take the geographic semantic into account in spatial knowledge representation and reasoning. In this paper, based on the semantic analysis of geographic relations, and human's cognition, an ontological knowledge base is established to store the knowledge related to spatial relations between geographic objects. The combination between the geo-ontology knowledge base and spatial relation query can reflect human's cognition better, and the knowledge base is one of the important parts in the spatial relation query system. To apply human's cognition in spatial relation query, this paper studies the design and structure of a geo-ontology knowledge base which stores the knowledge related to semantic information of geographic entities. This paper builds a knowledge base mainly on the base of analysis of geographic semantic information implicated in the description of spatial relations. The remainder of the paper is organized as follows. The expression of geographic semantic related to spatial relations is described in Section 2. Section 3 designs the geo-ontology knowledge base on the analysis of Section 2, and connects it to the spatial relation query system. Conclusions are given in Section 4.

II. THE EXPRESSION OF GEOGRAPHIC SEMANTIC RELATIVE TO SPATIAL RELATIONS

Spatial relations are the core content of GIS, and they play a major role in special data model, spatial query, spatial analysis, spatial reasoning and cartographic generalization. At present, most of people are studying the special relations created by space objects' geometrical features. These special relations are generally classified into four kinds: spatial distance, spatial orientation, spatial topological relations and similar relations. In geographic information system, the spatial entities are usually generalized to points, lines and polygons, so the quantitative computation and expression of special relations become more easily. However, all the spatial geographic ontology have specific geographic feature and semantic, and only the expression includes the spatial information and geographic semantic at the same time can meet people's cognition. In this part, we will from geographic entity types and human's cognitions to analyse the geographic semantic related to spatial relations query and description. There are 4 types of semantic information related to spatial relations. Each paragraph as follow represents one of the conditions.

For the same spatial relations, when the geographic entities are different, the words which used to describe their spatial relations may be different. For example, when the spatial relations are same, according to the semantic features between two geographic ontology types, two rivers are described by "flow into". In most cases, the word "flow into" is unidirectional, the reason is tributary can only flow into the main river rather than backflow. If the geographic entities are two roads, we will describe their spatial relation by words such as "intersects", "goes to" and "merges into".

When we use the same word to describe the spatial relations of different types of geographic entities, we might mean different kinds of spatial relations. For example, two rivers are mutually perpendicular. In this case, the rivers are disjoint. People will not query "one river is or isn't as perpendicular as the other river"; when two roads are mutually perpendicular, they are disjoint or intersectant and the angle is also in accordance with people's conception of perpendicular; if a road is perpendicular to a river, the perpendicular is disjoint.

Usually, we express the distance use "far" or "near", but what's the standard? In different contexts, everyone's understanding is different. If the purpose is not same, people's perceptions of distance are different. For instance, a sportsman thinks that 1000 meters is near, but an old man who goes shopping by walking will think the way is far. Other factors, such as language, culture, age and gender, can also influence the understanding and expression of distance.

In the daily lives, people always use up and down, left and right, east and west, south and north to describe the direction between two geographic entities. Because of the difference of spatial cognition, everyone has a reference system to distinguish direction, so the understanding of orientation is different. For example, when people read the sentence "there is a traffic light to the west of the national conference center", most of people may think the light in the sentence contains the traffic light at the northwest of the national conference center. Compared to the previous sentence, "at the east of the traffic light is national conference center" is not consistent with people's spatial cognition. In people's cognition, reference features must be huge, stable and easy to distinguish relative to target features.

III. CONSTRUCTION AND SOLUTION OF THE GEO-ONTOLOGY KNOWLEDGE BASE

In order to solve the problem which related to geographic semantic and spatial relations, this paper will establish a geo-ontology knowledge base about spatial relations. The geo-ontology knowledge base doesn't pursue the large and complete pattern, but it needs to be typical which can be used in spatial relation query. The main target is to establish the relationship between entities and attribute which are involved in spatial relations. People can not only understand the relationship between entities and attribute through the geo-ontology knowledge base, but also use it as the basis of reasoning. More significantly, the geo-ontology knowledge base can solve the problem in spatial relations query caused by the phenomena mentioned in Section 2.

A. Methods and Tools of the Construction of Ontology

It is a complicated work of building ontology, and distinct application fields need different field ontology, but they all follow the five basic principles which proposed by Gruber [12]. In the actual development process, there are three common methods of building ontology. They are Top-Down, Bottom-Up, and Middle-out. Each has its strong point. The geo-ontology hierarchy implemented in the paper is clear, and subordinate closely linked. So, according to the basic principles, it's reasonable to select the Top-Down method to build the knowledge base. The Top-Down method first defines the total frame structure of the knowledge base, and then forms sub hierarchies layer by layer.

In order to solve the conditions in Section 2, after the research analysis, we select OWL as modelling language, and open-source software Protégé as modelling tool. The reason is that owl can not only provide user with amounts of readable documentations, but also process the information of documentations and clearly express the words meaning and relations. Protégé is an open source ontology editor, and developed by Stanford University. It's compiled by Java, and possesses friendly interface style, users can easily learn to use. It has tree hierarchical directories to display ontology structure. Users can edit class and attribute by clicking the corresponding project, and can visually design ontology model.

B. Design of the Geo-Ontology Knowledge Base

Usually using ontology to building knowledge base will involve implementing five basic modelling meta-languages: they are class or concept, relation, function, axiom and instance. Among them, the most complex implementations are the definition of concepts and classes, and the determination of their relations. This part mainly introduces the process of the construction of the class and property.

1) *The Construction of Classes:* This paper takes Beijing as an example. According to different geographical ontology type contained in different spatial relationships, we mainly divided the geo-ontology into linear entity and polygon entity. The former contains river and road, and road is classified into downtown loop and trunk road, and river is classified into

main river and tributary; the later contains lake and building, and building is classified into colleges-university, and administrative division.

Ordering the class in hierarchy, we can use the class model in Protégé to create class and subclass. Fig.1 is the class hierarchy structure of classes.

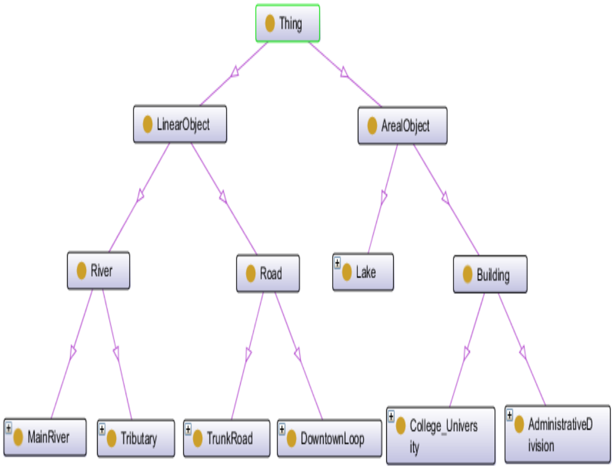

Fig. 1 The class hierarchy structure

2) *The Construction of Property:* After finishing the basic classes, we use object properties model and data properties model to build the properties of classes. There are object properties, data properties, function properties, transmit properties and inverse properties. In Protégé, defining the "Domain" and "Range" of the properties can express constrain of concrete concept classes. The property and the class that can use the property are connected by domain. And the scope of the property is determined by range. TABLE I lists part of the domain and range of the properties in the knowledge base.

TABLE II
DOMAIN AND RANGE OF THE PROPERTIES IN THE KNOWLEDGE BASE

Property		Domain	Range
Object properties	flow into	River	Lake; River
	near	Road	Building
	perpendicular1	River	River
	perpendicular2	Road	River
	starts in	River	Administrative Division
	inside	River	Administrative Division
Data properties	ID	All classes	int
	Area	Building; Lake	float
	Length	River; Road	float

For the spatial relation words that in the tableIII expressed by natural language, we know different people have different expressions and understandings of the natural language. For example, a spatial relation word "flow into", in order to expression the same meaning, some people may select "inflow" to express, others may think "afflux" more suitable. However, they all want to describe the same spatial relation that a river flows into another river or a river flows into a lake. Consider the situation, this paper using equivalent object properties which in the Object Properties Module to build the similarity property words of the spatial relation words. The method to some extent can solve the fuzziness caused by the natural language. For example, we may use the different words to describe the same relationship between two rivers or the same relationship between a river and a lake, for example, "Which water body the Kunyu river flows into?" or "Where the Kunyu river inflows?" In fact, the words "flow into" and "inflow" contain the same geographic semantic. Fig.2 shows their geographic semantic with the property restriction of the River class and relations with River Class and Lake Class in Protégé. In the Fig.2 we can see the two words have the same Domain and Ranges. The Domain is River, and the Ranges are River and Lake, so from the legend in the right of the Fig.2 we can see the two words show the same relations between River Class and Lake Class.

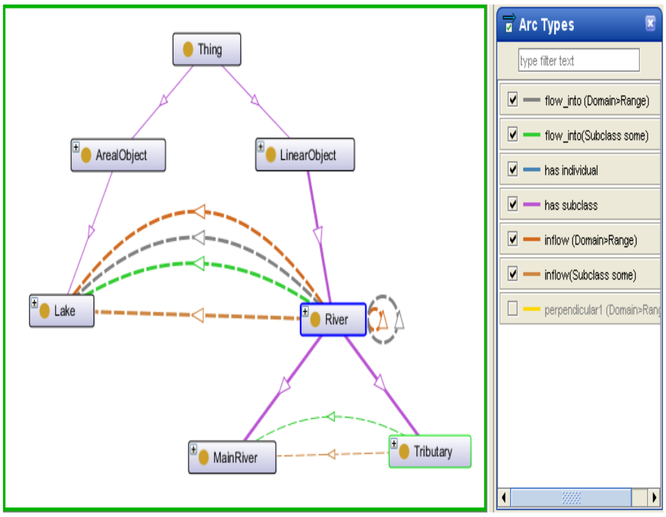

Fig. 2 Different words show the same spatial relation between the River class and the Lake class

3) *The Construction of Individuals:* According to the previous construct classes, properties and the permissible value of the classes and properties, in the Protégé using the Individuals model to add the specific individuals. For example, the Administrative Division class includes Haidian district, Chaoyang district, Dongcheng district, Xicheng district and so on. These individuals inherit the properties and relations of the classes. The ontology in this paper builds 242 individuals of the total. Among them, the Building class includes 63 individuals, the Lake class contains 16, the River class includes 12, and the Road class contains 151 individuals. After the above three steps, the basic framework of the geographic ontology of the Beijing city is constructed.

C. Solutions of Conditions in Section 2

So, the geo-ontology knowledge base which aims at different geographic ontology type and different spatial relation words, and sets different spatial relations between different geographic objects, can solve the condition of same spatial relation for different geographic ontology types. And for the other condition with the same spatial relations words and different types of geographic entities, we should set more sub relation of the spatial relation to distinguish different kinds of situations, such as perpendicular1 and perpendicular2 showed in Fig.2. For each sub relation, more detailed information about geographic entity type is specified.

D. Integrated with the Query System

The knowledge base will be integrated with the spatial relation query system to realize the specific query. The query system mainly includes three parts: the parsing model, the geo-ontology knowledge base, and the spatial relation query model. The parsing participle model used to analysis the input natural language query sentences. The geo-ontology knowledge base used to provide the relations between the geographic objects, and the properties of the objects. In addition, the knowledge base is combined with the query system by OWL port. The spatial relation query model calculates the indices which are used to quantifiably represent the spatial description words, and call spatial operators. Spatial operators are functions which fulfil the spatial relation queries.

Flowing is the steps of the query when the knowledge base is integrated with the query system. First, after natural language parsing the input the query sentence, we can get continuous phrase; at the same time, call OWL file through the OWL port. Second, matches the analysis phrase with the OWL file. If in the geo-ontology knowledge base we can find the concrete concept, we will also get the farther class and sub class of the concrete concept, and get the relations with other classes. Finally, according to the matched spatial relation property, calls the function of the same spatial relation word in the program, and use the spatial relations operators in the program to proceeds the spatial relations calculation between the geographic objects, and get the query result. Fig.3 shows the flow chart of the query.

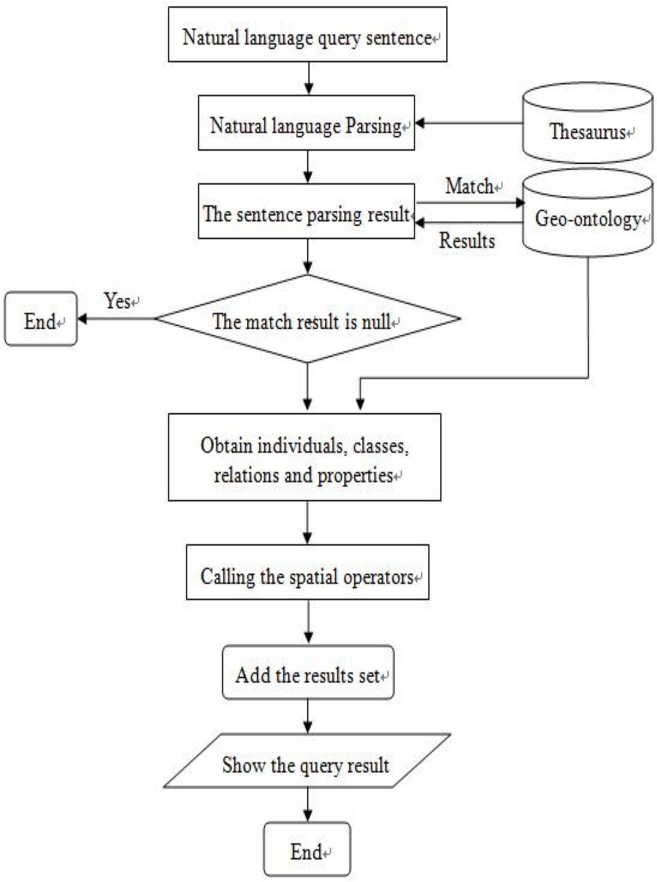

Fig. 3 The flow chart of the query

Following, taking the query sentences “Which place the Kunyu River flows into?” or “Where the Kunyu River inflows?” as example, we can elaborate the condition of using the different spatial words to describe the same spatial relations by applying the knowledge base in the query system. First, go through the natural language analysis, and obtain the geographic entities and spatial relation to be queried in the sentence, and at the same time, in the background the OWL port calls the OWL file. Second, matching the result of the participle with the OWL file, we can obtain “Kunyu River” is mapping with the “Kunyu River” in the River class of the knowledge base, and flow into or inflow is mapping with the flow into or inflow in the spatialrelations of the Object Properties. In Protégé “flow into” and “inflow” are the equivalent object properties, they have the same Domain and Ranges, and the Domain of the “flows into” and “inflows” is River, the Ranges of the “flows into” and “inflows” is River and Lake. Now we can receive two equivalence queries that one is Kunyu River flows or inflows into a river, the other is Kunyu River flows into or inflows a lake. We finally fixed the specific individual and classes. Third, we can map the spatial relation word “flows into” or “inflows” in the OWL file with the spatial query operator in the query system, and call the spatial query operator to calculation. Finally, show the result on the map.

IV. CONCLUSIONS

This paper presents a geo-ontology knowledge base which is constructed to assist spatial relation query. In particular, the paper focuses on the problem of solving spatial relation query in different context, which mainly performance in three conditions in Section 2. The knowledge base is integrated to the spatial relation query system, and it is used in spatial relation query. But the knowledge base is not complete, and needs to include more spatial relations and more geographic object types, such as points, to serve the query of spatial relations between points and lines or points and regions.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under grant number 40701151.

REFERENCES

[1] X. B. Yao, J. C. Thill, "How far is too far- a statistical approach to proximity modeling," *Transactions in GIS*, vol.9, pp.157~178, Nov.2005.

[2] D. M. Mark, D. Comas, M. J. Egenhofer, *Evaluating and refining computational models of spatial relations through cross-linguistic human-subjects testing*, in Frank A and Kuhn W. Spatial Information Theory – A theoretical Basis for GIS, International Conference COSIT', Austria, Lecture Notes in Computer Science, Berlin: Springer-Verlag, 1995, paper 988, p.553-568.

[3] D. M. Mark, M. J. Egenhofer, "Calibrating the meaning of spatial predicates from natural language: Line-regions relations," *Proceedings, Spatial Data Handing*, vol.1, pp.538~568, Nov.1994.

[4] D. M. Mark, M. J. Egenhofer, "Topology of prototypical spatial relations between lines and regions in English and Spanish," *Proceedings, Auto Carton 12.Charlotte, North Carolina*, pp.538~568, Nov.1995.

[5] G. H. Fu, C. B. Jones and A. I. Abdelmoty, "Building a Geographical Ontology for Intelligent Spatial Search on the Web," in *Proceedings of the IASTED International Conference on Databases and Applications,* 2005, p.167-172.

[6] A. I. Abdelmoty, P. Smart, C. B. Jones, "Building Place Ontologies for the Semantic Web: Issues and Approaches," in *GIR '07: Proceedings of the 4th ACM workshop on Geographical information retrieval*, 2007, p.7-12.

[7] J. Xu, "Formalize natural-language spatial relations between linear objects with topologic and metric properties," *International Journal of Geographical Information Science*, vol.21, pp.377~395, Nov.2007.

[8] J. Xu, "Natural-language representation and query of liner geographic objects in GIS," Ph.D.Eng.thesis, Department of Geography, State University of New York at Buffalo, 2005.

[9] Y. Liu, Y. X. Gong, J. Zhang, Y. Gao, "Representation and Reasoning of Spatial Relations in Geographical Space," *Geography and Geo-Information Science*, vol.23, pp.1-7, Nov.2007.

[10] F. L. Bian, K. Mei, "Building Location Ontology for Geographic Knowledge Base," *Geomatics World*, vol.4, pp.27-32, Nov.2008.

[11] D. S. Jing, "A Research on Geo-Spatial Information Semantic Expression and Service Based on Ontology," Ph.D.Chn.thesis, Institute Of Remote Sensing Applications, CAS, Beijing, 2005.

[12] T. R. Gruber, "Towards Principles for the Design of Ontologies Used for Knowledge Sharing," *International Journal of Human Computer Studies*, vol.43, pp.907~928, Nov.1995.