The research on reasoning of tourism information and knowledge based on semantic Web

Zhiqiang Liu*, Yila Su, Fei Wang, Huimin Li

College of Information Engineering, Inner Mongolia University of Technology, Hohhot, 010051, China

Received 1 October 2014, www.cmnt.lv

Abstract

As a kind of modelling tool of describing knowledge framework based on semantics and knowledge, ontology provides a standardized description of concepts and lays a foundation for the sharing of knowledge. This paper constructs a tourist information knowledge base body via ontology tools, and meanwhile makes a series of reasoning based on the constructed knowledge base. Furthermore, one can manually construct inference rules to probe into the field of knowledge and complete the key objectives of the Semantic Web—“machine-understandable.” For the uncertainty which cannot be described by descriptive logic, this paper recommends the introduction of Bayesian networks, which can not only easily construct conditional probability tables, but can conduct a rapid reasoning of probability and finally provides people with a reasonable travel plan.

Keywords: semantic web, ontology, tourism, reasoning

1 Introduction

World Wide Web has become the main channel to get information, and has deeply affected every aspect of human social life; people on the Web can read news, search for information, buy and sell goods and services. However, the current World Wide Web is people-oriented rather than machine-oriented [1]. Internet founder Tim Berners-Lee proposed a model of the next generation Internet Semantic Web, also known as the Semantic Web (Semantic Web). In the vision of Tim Berners-Lee, the Semantic Web is an extension of the Internet now, Semantic Web information with a defined semantics, enabling computers and people to work together better [2].

Since the introduction of the Semantic Web, the Semantic Web has attracted academia and industry. Many companies in the emerging Semantic Web, experts and scholars have tried a variety of organizations to improve the quality of the network information resources. For example, the traditional tourism information system is developed on the basis of database system. These systems are powerful, covering tourism, food, lodging, transportation, guide, shopping trips and other services, but with the dramatic increase of tourist service information and demand, inquiry, based on the traditional tourist information, cannot meet the current tourists’ choices. With the current new developments of Semantic Web technologies, we propose a Semantic Web-based tourism information services to build the body of knowledge, to increase machine-understandable semantic content, semantic reasoning to achieve a query. Semantic Web ontology as the fourth seven-layer architecture is a 21st century frontier research.

2 Semantic Web

2.1 SEMANTIC WEB OVERVIEW

Semantic Web is currently a hot research Internet technology. In order to promote the development of the Semantic Web and increase the importance of the Semantic Web research, many foreign universities, research institutes, large companies have set up a special project team to promote the development of this technology. World Wide Web Consortium W3C (World Wide Web Consortium) set up a special working group and a variety of related technical standardization [3].

Semantic Web is the development trend of the next generation of the World Wide Web, which provides a common framework to share and reuse data beyond applications, enterprises and communities border. By means of Semantic Web technology, people could weave a large-scale computer easily to understand and process knowledge networks, thus facilitating reuse and integration of data, so that the future of the World Wide Web can provide better applications and services.

2.2 SEMANTIC WEB ARCHITECTURE

Seven of the Semantic Web architecture shown in Figure 1 [4].

The first layer: Unicode and URI. The Semantic Web architecture, the layer is the basis for the semantic network, which is responsible for processing resources Unicode encoding. URI is responsible for the identifier of the resource.

The second layer: XML + NS + xmlschema, as the grammar layer. This layer is responsible for representing data from the grammatical content and structure of language through the use of standard forms of network
information, data structures and content separation.

Third layer: RDF + rdf schema, as the data layer. RDF model into an object-oriented type system can be expanded to RDF, providing an appropriate domain and range, as well as class and subclass levels.

The fourth layer: The main body layer (Ontology Layer) as a semantic layer. On this layer, users can not only define the concept but also define the relationship between the concepts of wealth.

Fifth to seven: Logic, Proof, Trust, Logic Layer provides intelligent reasoning rules; Proof Layer supports the exchange of evidence of communication between agents; digital signatures and Trust Layer are to ensure the exchange of information security issues and design.

Seven-story structure of the Semantic Web XML, RDF and Ontology three, mainly for representing information semantics, is the core and key. The following mainly describes related content of Ontology

![Architecture of Semantic Web](image)

FIGURE 1 Architecture of Semantic Web

2.3 RELATED THEORY OF ONTOLOGY

The concept of the ontology was originally derived from the philosophy, which means the objective existence of a systematic explanation and description of the objective reality of an abstract nature. Later, with the development of artificial intelligence, it was given a new definition by artificial intelligence community.

In 1993, Gruber gives the ontology a most popular definition, namely "ontology is an explicit conceptual model specification." Later, Borst gave ontology of another definition on this basis: "Ontology is a shared conceptual model of the formal specification." Studer etc. conducted an in-depth research on the two definitions, that "ontology is a shared conceptual model of clear formal specification." This includes four layers of meaning: conceptualization, explicitness, formalization and share. The ultimate goal of ontology is "to accurately represent that implicit (or explicit) information."

The language to describe ontology was proposed by W3C in 2004. It has three increasingly expressive sub-languages to suit different occasions: OWL Lite, OWL DL and OWL Full. Considering the application, from the perspectives of expression and reasoning as well as the judgment of the reasoning, the choice of OWL DL ontology representation language as this passage can provide a satisfactory subset of the language.

3 Building ontology

3.1 BUILD TOOLS AND PLATFORMS

It is necessary to build ontology on a good platform before they can proceed smoothly. Protégé, an ontology modeling tool developed by Stanford University, is also currently the world's most popular ontology modeling tools.

First, install JDK and configure environment variables. Second, download and install the Protégé ontology tool developed by the Java language, and as a visual development platform.

Third, download and install the insert graphic, which can graphically display the parent class and subclass body diagram.

Fourth, download and install Mysql and the client tools for the body stored in the database.

Fifth, use RacerPro inference reasoning in the reasoning process.

Sixth, hand-constructed semantic web rule that comes Protégé SWRL Rules.

Seventh, this article will construct Bayesian network topology through the Netica and the uncertainty of knowledge reasoning.

3.2 ONTOLOGY CONSTRUCTION OF THE TOURIST INFORMATION

3.2.1 Create class relations attributes and domains

Tourism related classes that Ontology created has Season class, Destination class, Accommodation class, a series of activities related and so on. Semantically speaking, there are four kinds of basic relations: part-of, kind-of, instance-of, and attribute-of. In this paper we used a more attribute-of relations, such as has Accommodation, has Activity, has Part, has Rating, locatedIn etc., which means that the concept of a concept is another attribute.

Meanwhile, the object attributes can be customized according to the actual situation. There are two cases: domain and range. For example, definition domain of hasRating is the Accommodation, range is AccommodationRating. There are four levels to choose from (OneStarRating, TwoStarRating, ThreeStarRating, FourStarRating), detailed in the Table 1.

<table>
<thead>
<tr>
<th>Property Name</th>
<th>Domain</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasRating</td>
<td>Accommodation</td>
<td>AccommodationRating</td>
</tr>
<tr>
<td>hasActivity</td>
<td>Destination</td>
<td>Activity</td>
</tr>
<tr>
<td>hasPark</td>
<td>Destination</td>
<td>Park</td>
</tr>
<tr>
<td>hasGrassland</td>
<td>Destination</td>
<td>Grassland</td>
</tr>
<tr>
<td>hasMuseum</td>
<td>Destination</td>
<td>Museums</td>
</tr>
<tr>
<td>belonging</td>
<td>Accommodation</td>
<td>Destination</td>
</tr>
<tr>
<td>locatedIn</td>
<td>Accommodation</td>
<td>Destination</td>
</tr>
<tr>
<td>hasGoodSeason</td>
<td>Activity, Destination</td>
<td>Season</td>
</tr>
<tr>
<td>isGoodSeasonTo</td>
<td>Season, Destination</td>
<td>Activity</td>
</tr>
</tbody>
</table>

3.2.2 Create an instance of ontology

Select the appropriate class in OWL / RDF format, create an individual entity and assign the value for it. Take
tourism ontology as an example, its related attributes such as tourist season, tourist destination, tourist accommodations and activities, you can create such an instance, that instance to the appropriate assignment, such as the tourist season in spring; destination RuralArea; activities Sightseeing; Accommodation Hotel. When ontology is defined and instantiated, they could be put together to make a knowledge base. We will build the tourism ontology as a specific field of knowledge, give full play to the conceptual model of the ontology explicit and formal description of properties. See Table 2.

3.2.3 Tourism ontology class diagram obtained

Open the tourist information ontology project file which is conducted, click the tab OWL Viz tab (first you need to download and install the plug-graphviz-2.8, visual graphics software), click the yellow button on the top left of the first one, let the mouse close to the display show class, click after a blank area on the right shows the hierarchy of the relevant ontology relationship between parent and child classes at a glance.

<table>
<thead>
<tr>
<th>Master Class</th>
<th>Subclass</th>
<th>Corresponding instance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodation</td>
<td>BedAndBreakfast</td>
<td>Yurts, nomads people</td>
</tr>
<tr>
<td></td>
<td>BudgetAccommodation</td>
<td>Grand Castle Hotel, Shuya hotels, Yinghua Hotel</td>
</tr>
<tr>
<td></td>
<td>LuxuryHotel</td>
<td>Three star: Olympic Hotel, Baotou Hotel; Four Star: Jinyi Holiday Inn Taiwei hotels, Tianhe International Hotel, Shangri-La Hotel, Zhaojun Hotel</td>
</tr>
<tr>
<td>Destination</td>
<td>RuralArea</td>
<td>Paul co little town (holy water beam), the ancient road village board (ski), Takekawa (Xilamuren grasslands), Yau Yau board (Mu Tu Senlin Park)</td>
</tr>
<tr>
<td></td>
<td>UrbanArea</td>
<td>Muslims in Saihan District, New Town, Yuquan District</td>
</tr>
<tr>
<td>Activity</td>
<td>Sports</td>
<td>Horseback riding, wrestling, archery</td>
</tr>
<tr>
<td></td>
<td>Adventure</td>
<td>Hiking, skiing, sandboarding</td>
</tr>
<tr>
<td></td>
<td>Relaxing</td>
<td>Solarium, camera Photography</td>
</tr>
<tr>
<td></td>
<td>SightSeeing</td>
<td>Princess House Park, five reservoirs, Mu Tu Senlin Park; big Zhao Temple, General Government Office, Zhaojun; Nanhu Wetland; Sand Bay; Xilamuren grasslands, Sailantala grasslands, holy water beam</td>
</tr>
<tr>
<td>AccommodationRating</td>
<td></td>
<td>One star, two star, three star, four-star</td>
</tr>
<tr>
<td>Season</td>
<td></td>
<td>Spring, summer, autumn, winter</td>
</tr>
</tbody>
</table>

3.3 SAVE THE ONTOLOGY IN THE DATABASE

Save the ontology in the database, here we use Mysql 5.0, first establish a database after installation protégé Db, save the ontology in the database with MyEclipse6.5. The program is saved as Onto2Database.java, the key code as follows:

```java
// creating a database connection
IDBConnection con = new DBConnection(strURL, strUser, strPassWord, strDB);
// using the database connection parameters to create a model making machines
ModelMaker maker = ModelFactory.createModelRDBMaker(con);

Program has performed: “this OWL file has been already stored in the database ontology has persisted to the database ...”

Display 5 database, protégé db is built for the user, while the other four as the system default.

A database only needs to be created one time, but you must start mysql session each time, first select it and then use the command and then use this database, then you can create a new data in the database table. Ontology knowledge was stored in the database, once created, multiple use, and call to the database every time. Meanwhile, you can create tables in its database and conduct a series of queries, delete, insert, update, and other operations; the Construction Grammar is similar with SQL syntax so that the operation is more convenient.

3.4 INQUIRY AND REASONING OF ONTOLOGY KNOWLEDGE LIBRARY

Descriptive logic is a decidable subset of the first-order logic, which has strong communication skills. The important feature of descriptive logic is strong skills and decidability. It can ensure the inference algorithm can always stop and return the correct results. Rule language is achieved by describing logic in this article [5]. Query operation under the Protégé environment is achieved by QueryTab components. For example, there is a skiing sport location at the sight spot, as shown below:

![FIGURE 2 The querying interface of traveling ontology](image)

Click the Find button to query, we can see Guluban village on the right, and display what we located in the suburbs.

3.4.1 Jess-based reasoning

Jess (java expert shell system) is CLIPS inference engine based on the Java language. When a user uses class / slot to describe ontology, Protégé only use Jess for reasoning.
JessTab is plugin of Protégé that allows users to integrate with Jess and Protégé. Jess integrates with Protégé, then you can complete four tasks: change SWRL rules into Jess Jess rules or directly create rules based on ontology; express knowledge of individual's OWL related as Jess facts; running Jess rule-based Reasoning and facts engine Reasoning; updated OWL knowledge based on the new facts that has reasoned [6]. The working process is shown in Figure 3.

![FIGURE 3 the Jess reasoning process of the ontology knowledge base](image)

### 3.4.2 BUILD SWRL RULES OF TOURISM ONTOLOGY

In building domain ontology, the relationships between class and class, class and subclass levels and among distinct attachment are quite obvious. If some kinds of object attributes are given a series of restrictions, the different classes of objects becomes mutual implication between attributes nested relationship; if a specific inference rules are builded, the relationship implied can be inferred.

Consider general SWRL rules: A (? X) ∧ B (? X) → C (? X)

If you analyze the above rules form the perspective of the first-order predicate, you can get the following semantics: if the variable x is an instance of both class A and class B, the variable x is also an instance of class C.

**TABLE 3 Part of the structure rules of traveling ontology**

<table>
<thead>
<tr>
<th>Name</th>
<th>Rule Expressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule-1</td>
<td>hasAccommodation(?x, ?y) ∧ hasRating(?y, ?z) → hasRating(?x, ?z)</td>
</tr>
<tr>
<td>Rule-2</td>
<td>locatedIn(?x, ?y) ∧ belonging(?y, ?z) → belonging(?x, ?z)</td>
</tr>
<tr>
<td>Rule-3</td>
<td>hasMuseum(?x, ?y) ∧ hasMuseum(?x, ?z) ∧ locatedIn(?y, ?d) → locatedIn(?z, ?d)</td>
</tr>
<tr>
<td>Rule-4</td>
<td>hasDesert(?x, ?z) ∧ locatedIn(?x, ?d) → locatedIn(?z, ?d)</td>
</tr>
<tr>
<td>Rule-5</td>
<td>hasGrassland(?x, ?y) ∧ hasActivity(?y, ?z) → hasActivity(?x, ?z)</td>
</tr>
</tbody>
</table>

### 4 Uncertainty reasoning

In reality, the information and knowledge people get, often contain a lot of inaccuracy, incompleteness, inconsistency. That is, the information obtained is not definite and clear. Uncertainty Reasoning refers to building this knowledge and evidence on uncertainty reasoning, which is based on Bayes' rule.
The topology tool of Building the Bayesian network is Netica. And it is developed by the world-wide Norsys as application software. Netica is a powerful diagram. It is easy to use and complete. The drawn out of the network node map is very intuitive, by entering the individual probabilities to represent the association between each variable, based on the input of the individual probabilities equations of the form, or to learn from the data file obtained [10].

4.1 THE UNCERTAINTY OF THE TOURIST INFORMATION AND KNOWLEDGE

As we all know, when travellers go to visit some places, they should know not only what scenic spots to visit, but also what season is the best choice to travel. In this example, Attractions is not selected before. The default probability of season node is \( P(\text{Spring}) = P(\text{Summer}) = P(\text{Autumn}) = P(\text{Winter}) = 0.25 \). The causal node associated and compiled, probability values change accordingly. In addition, they will determine the value of tourists travel plans, which is the manner of climatic conditions, geological conditions, and visit scenic spots and so.

The knowledge also can be learned from the body. The treatment of uncertainty in the evidence does not have any description, so it is uncertain whether to visit Inner Mongolia. Therefore, people build data based on node value to the investigation or an accepted fact, and as a priori probability for input, on the other hand, the value is also affected by seasonal effects. Similarly, the season also input value in the form of a priori probability.

4.2 BAYESIAN NETWORK INITIALIZATION

Reading ontology probability, the parent node does not exist. A priori probabilities can be its CPT table input. For the parent node exists, if only one parent, it can be directly used as the conditional probability of the corresponding CPTs. If you have more than one parent node, then enter the corresponding CPTs, but it needs to be based on the edge of statistical independence assumptions and conditional probabilities to calculate its joint conditional probability, as shown in Table 4.

Bayesian network reasoning compiled diagram form in Figure 4. Visit Mongolia this season and two nodes can be directly or indirectly as a reason, compile, execute the following results.

4.3 THE INFERENCE OF BAYESIAN NETWORK

There are three cases which is causal reasoning, diagnostic reasoning, prediction and diagnosis of combining reasoning in Bayesian network inference. First, causal reasoning is reason to push the conclusion that the top-down reasoning.

Click summer which is Season value. The updated calculation shows that visiting Inner Mongolia is more probability, the probability is 0.8, and they decide to go to the maximum probability prairie landscape. The result of the conditional probability is \( P(\text{Grassland}) = 0.480 \). The probability of grassland adjacent to Hohhot is 0.446, to Baotou is 0.554 (because of the summer's most popular attractions also include Desert which is located in the desert near Baotou, namely Sand Bay). That is known by the ontology library that prairie named Saihan Tara is located in the steppe territory Baotou, followed by a Desert named Sand Bay; Grassland located Hohhot territory. In addition, the results from reasoning can also be seen near a prairie accommodation situation in which most of the two-star hotel star, followed by a star and three-star.

TABLE 4 CPT table of some relevant nodes

Bayesian network reasoning compiled diagram form in Figure 4. Visit Mongolia this season and two nodes can be directly or indirectly as a reason, compile, execute the following results.

FIGURE 4 Part nodes of Bayesian topology of traveling ontology

FIGURE 5 Bayesian topology for casual reasoning

FIGURE 6 Bayesian topology for diagnosis reasoning
Second, diagnostic reasoning and causal reasoning are on the contrary; it is the reason for the conclusion to infer, reasoning is from the bottom up.

When we click FourStar Rating which is Accommodation Rating star status value, corresponding probability is that the maximum probability season is winter, \( P(\text{Winter}) = 0.370 \), and the choice of the scenic spots can only be snow-capped mountains. Probability is shown in the following figure, Travellers do not choose to visit Inner Mongolia in the winter is relatively large probability, perhaps because the weather is cold and dry. But as long as one chooses to go to Inner Mongolia, the maximum probability is going Snow Mountain, the value is 0.395, and the probability of attractions located in Hohhot is maximum.

![Bayesian topology for mixed reasoning](image)

Third, reasoning combined predicts and diagnose. In this example, collect every bit of evidence. Select scenic spots near three-star hotel and travel plans for the fall season. The Bayesian inference that choose to visit and not consistent with the probability of visiting Mongolia. If the trip, then visit the museum visitors choose to Hohhot maximum probability, its value is 0.545. Museum is located in the new city with generals Government Office (Jiangjunyashu), Yuquan District Zhaojun (Zhaojun Museum) and Jokhang Temple (DaZhaoSi); Second, they can also be pleasing to play Hohhot park, there are located (locatedIn) Muslims in the Princess House Park (GongzhuangPark), Muslims Yau Yau board Mu Tu Senlin Park (WusutuWoodPark), and scenic spots in Daqing Water Park (WaterPark), that is fifty-one reservoir.

Probabilistic description logic is based on Bayesian networks will make the results more satisfying through continuous improvement. Then we should make the appropriate efforts in the following aspects, by changing the existing OWL ontology language to enable them to represent uncertainty, or the Bayesian network to compensate for their deficiencies in knowledge representation.

5 Conclusions and recommendations

It is an inevitable trend of future network development which the construction is based on the Semantic Web System. We know that the Semantic Web ontology is the most critical technology. It makes it possible to share and reuse the information on the Web. Especially, when there are several uncertain events in real life, it is more probable to reflect it’s unique through Bayesian event calculus method. And, for discrete events, we can diagnose the previous event status according to the uncertain non-complete data through Bayesian posterior probabilities, such as diagnostic reasoning.

We know that the Semantic Web ontology is the most critical technology. It makes it possible to share and reuse the information on the Web. However, the current ontology editing tools are less developed. If we want to make ontology editing tools more widely used, we should go further in the interface, functional improvement, reasoning support and integrate with other development tools. With the further develop of the Semantic Web technology research and the gradual expansion of the application fields, we have every reason to believe people will receive more systematic, scientific, intelligent service in life, work, business, medical, learning and entertainment. If we want to get more humane reasoning, intelligent information and knowledge through inference, we can gradually be improved by trying to learn and explore. For example, if reasoning and solving methods can be integrated together or either of which, Hybrid Bayesian networks, dynamic Bayesian networks, fuzzy logic, constraint satisfaction problems, neural network and so on, is combined together, the result would be more reasonable.

Acknowledgments

This work is partially supported by National Natural Science Foundation of China (61363052), Natural Science Foundation of Inner Mongolia Autonomous Region (2012MS0904), College and University Scientific Research Project of Inner Mongolia Autonomous Region (NJZZ12046) and Inner.

References

[6] Li Jie, Ding Ying 2007 Semantic gateway key technical overview [J] Computer Engineering and Design. 28(2) 66-90
Authors

Zhiqiang Liu, November 24, 1972, Inner Mongolia, China.
Current position, grades: vice professor at Computer Department, Inner Mongolia University of Technology.
University studies: MSc degrees in computer science from Beijing University of Aeronautics and Astronautics in 2003.
Scientific interest: artificial intelligence, embedded technology and its application, the Internet of things, knowledge information systems.
Publications: over 10 papers.
Experience: 2000.7-Now, Inner Mongolia University of Technology, as a vice professor.

Yila Su, March 17, 1964, Inner Mongolia, China.
Current position, grades: professor at Computer Department, Inner Mongolia University of Technology.
University studies: PhD degree in computer science from International WIC Institute, Beijing University of Technology in 2009.
Scientific interest: artificial intelligence, knowledge discovery and data mining, Web intelligence, intelligent agents, knowledge information systems.
Publications: over 50 papers.

Fei Wang, October 25, 1989, Inner Mongolia, China.
Current position, grades: graduate student, Inner Mongolia University of Technology.
University studies: BSc degree in computer science from Chongqing Three Gorges University in 2012.
Scientific interest: artificial intelligence, embedded technology and its application, the Internet of things, knowledge information systems.

Huimin Li, August 7, 1989, Inner Mongolia, China.
Current position, grades: graduate student, Inner Mongolia University of Technology.
University studies: BSc degree in computer science from Chongqing University of Post and Telegram in 2011.
Scientific interest: artificial intelligence, embedded technology and its application, the Internet of things, knowledge information systems.