Towards the Real-World Semantic Web—Web Search based on Spatial and Temporal Metadata

Jun-ichi Akahani†, Kaoru Hiramatsu, Akira Sugiyama, Yutaka Yanagisawa, and Tetsuji Satoh

Abstract

To enrich human communication in ubiquitous environments, NTT Communication Science Laboratories have been conducting research on the Real-World Semantic Web. When you deal with real-world information in ubiquitous environments, spatial information and temporal information play important roles. However, it is difficult to deal with spatial or temporal search conditions in current Web search engines. Therefore, we propose a Web search framework based on spatial and temporal metadata.

1. Introduction

A huge amount of information has been accumulated on the World Wide Web and it is still growing day by day. Web information is now indispensable to our daily activities. On the other hand, various sensors in the real world have been connected to each other via networks and ubiquitous environments will soon be available. The integration of Web information and sensory information will enable us to enrich the communication among humans. For example, a system could monitor a human conversation and provide Web information related to the conversation topic. With the aim of enriching human communication in ubiquitous environments, NTT Communication Science Laboratories have been conducting research on the Real-World Semantic Web.

We propose a Web search framework based on spatial and temporal metadata. This framework is based on the Semantic Web [1]. In the Semantic Web, each Web page is annotated with metadata based on semantic structures, called ontologies. An ontology represents the meaning of terms in vocabularies and the relationships among those terms. To deal with spatial and temporal search conditions, we define ontologies for time and space. Based on this framework, we have constructed a prototype system that allows users to make queries in natural language. After briefly reviewing the Real-World Semantic Web, we present our Web search framework and describe the prototype system.

† NTT Communication Science Laboratories
Soraku-gun Kyoto, 619-0237 Japan
E-mail: akahani@cslab.kecl.ntt.co.jp

2. Real-World Semantic Web

The concept of the Real-World Semantic Web is illustrated in Fig. 1. Various sensors, such as GPS (Global Positioning System) devices attached to users and video cameras located in the environment, connected to each other via networks provide "ubiquitous contents." The core technologies for the Real-World Semantic Web are query processing for the Semantic Web and semantic integration of ubiquitous contents. We have been developing technologies for Web searches based on spatial and temporal metadata, approximate query reformulation for multiple ontologies [2], [3], techniques for understanding conversations based on ontologies, and a personal repository for inter-personal communication. In this article, we focus on Web searches based on spatial and temporal metadata.

3. Spatial and temporal metadata and ontologies

Imagine you have arrived at a trip destination late at night and want to have a late dinner. You might desire information about "the nearest restaurant that is cur-

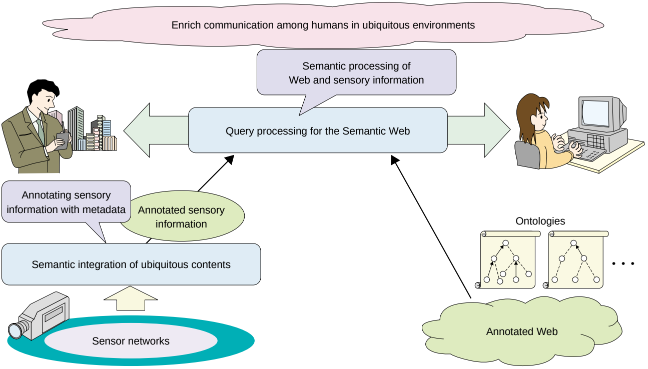

Fig. 1. Real-World Semantic Web.

rently open". This involves both spatial search conditions (*e.g.*, "nearest to the current location") and temporal search conditions (*e.g.*, "currently operating").

The current time and location can be obtained by sensors such as a watch and a GPS. However, keyword-based Web search engines cannot handle these kinds of sensory information. A Web search based on temporal conditions requires the temporal inference that the current time is within the operating hours described in the Web pages. Furthermore, if a restaurant's Web page gives its regular holiday, then we need the temporal inference that the current date does not match it.

There are various semantic structures for time such as time intervals (*e.g.*, "from 5 pm to 11 pm"), day of the week, date, and month. Therefore, a Web search based on temporal conditions requires ontologies for time that intuitively correspond to dictionaries for time. Similarly, a Web search based on spatial conditions requires ontologies for space, such as street address and latitude and longitude.

Our Web search framework is based on spatial and temporal metadata. In this framework, each Web page is annotated with metadata based on ontologies for time and space. Figure 2 shows an example. The left-hand side indicates the Semantic Web layers for ontology, metadata, and contents.

The ontology layer provides a framework for describing the semantic structure. Ontologies are described using classes and properties. For example, "Restaurant" and "Location" are classes, and "locatedAt" and "openHours" are properties. Classes are characterized by their properties. There are several built-in properties, such as "subClassOf" for describing class hierarchies. For example, Fig. 2 shows the following definitions. (We indicate the "subClassOf" property by a solid arrow for easy visualization.)

- "Japanese Restaurant" and "French Restaurant" are subclasses of the "Restaurant" class.
- "Restaurant" has the properties "locatedAt," "openHours," and "regularHoliday."
- The values of "locatedAt" must belong to the "Location" class. (The "Location" class is called the "range" of the "locatedAt" property.) Similarly, the ranges of "openHours" and "regularHoliday" are the classes "TimeInterval" and "DayOfWeek," respectively.

The Metadata layer provides a framework for describing metadata with which contents are annotated based on ontologies. For example, the Web page for "Gion" (Gion is a district in the city of Kyoto, but

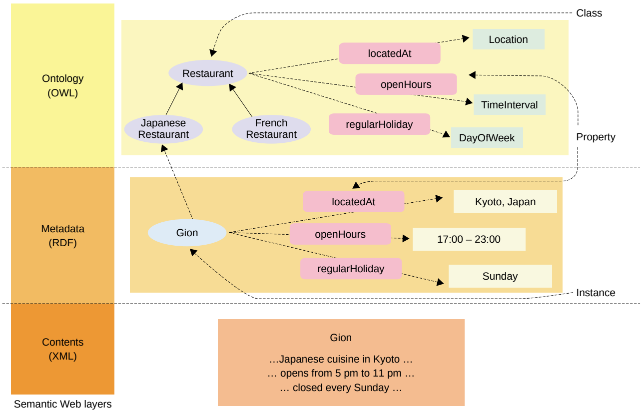

Fig. 2. Example of spatial and temporal metadata and ontologies.

here the context is the name of a restaurant) has the metadata shown in Fig. 2. This metadata indicates the following.

- "Gion" is an instance of "Japanese Restaurant."
- "Gion" is "locatedAt" Kyoto, its opening hours are from 17:00 to 23:00, and its regular holiday is Sunday.

In our framework, users can make queries based on ontologies. For example, one can make a query about a restaurant that is currently open as follows.

- Find an instance of "Restaurant" class such that the current date does not match its regular holiday and the current time is inside the time interval of its opening hours.

If the current time is 10 p.m. on Tuesday, then "Gion" will be returned as a candidate answer to this query when we use the ontology and metadata.

4. Web search system based on spatial and temporal metadata

The system architecture of our prototype is shown in Fig. 3. This system lets users make queries in natural language, such as "find the nearest restaurant that is currently open." The query analysis module extracts spatial and temporal search conditions from the user input and formulates a query based on ontologies. We adapted the query analysis module of the question answering system SAIQA (System for Advanced Interactive Question Answering) developed at NTT Communication Science Laboratories. For the formulated query, Web pages are retrieved based on spatial and temporal metadata, as described above. Search results are displayed on a map and in tree and list views.

Figure 4 illustrates the user interface of the Web search system. In this example, a user inputs the query "find bus stops near Kinkakuji" in Japanese. The query analysis module formulates a query for the Web pages of "Kinkakuji" and bus stops near this famous temple. The search results are displayed on the map. The upper left icon is an anchor of the Web page for "Kinkakuji" and represents its location on the map. Similarly, the center icon is an anchor of the Web page for the nearest bus stop ("Kinkakuji-Mae"). The search results are also displayed in 2D/3D views and list view.

If we incorporate a route finding system into this

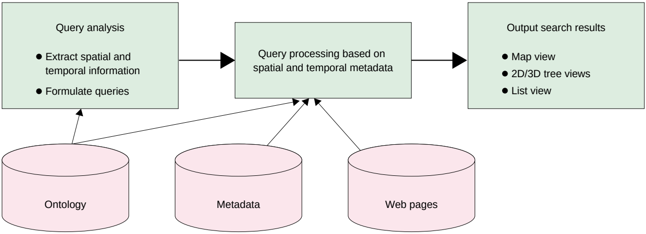

Fig. 3. Architecture for Web search system based on spatial and temporal metadata.

Fig. 4. Example of Web search based on spatial and temporal metadata.

system, we can provide various services related to regions. However, the ontologies for those services may differ. We can overcome this problem by using approximate query reformulation for multiple ontologies [2], [3].

5. Conclusion

Spatial information and temporal information play important roles in dealing with real-world information in ubiquitous environments. Our Web search framework based on spatial and temporal metadata enables users to make queries based on ontologies for time and space. Our prototype system based on this framework lets users make queries in natural language. The current version presumes explicit text input from users in natural language, but we have also been developing technologies for understanding conversations based on ontologies. Together with these technologies, we plan to construct a system that monitors human conversation and provides Web information related to the conversation topic.

The Real-World Semantic Web requires semantic integration of ubiquitous contents. As a first step, we have been developing an efficient method to handle moving object trajectories [4]. This method enables users to find people who are viewing similar exhibits based on their trajectories at an event site. People with similar interests will be able to enjoy much richer and deeper communication by using an ad-hoc electronic bulletin board and accessing Web information related to the exhibits.

References

[1] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," Scientific American, Vol. 284, No. 5, pp. 34-43, 2001.

[2] J. Akahani, K. Hiramatsu, and T. Satoh, "Approximate Query Reformulation for Multiple Ontologies in the Semantic Web," NTT Technical Review, Vol. 1, No. 2, pp. 83-87, 2003.

[3] J. Akahani, K. Hiramatsu, and T. Satoh, "Approximate Query Reformulation based on Hierarchical Ontology Mapping," International Workshop on Semantic Web Foundations and Application Technologies (SWFAT), Nara, Japan, pp. 43-46, Mar. 2003.

[4] Y. Yanagisawa, J. Akahani, and T. Satoh, "Shape-based Similarity Query for Trajectory of Mobile Objects," 4th International Conference on Mobile Data Management (MDM-2003), Melbourne, Australia, LNCS 2574, pp. 63-77, Springer-Verlag, Jan. 2003.

Jun-ichi Akahani

Senior Research Scientist, Supervisor, Social Communication Laboratory, NTT Communication Science Laboratories.

He received the B.E. and M.E. degrees in applied mathematics and physics from Kyoto University in 1983 and 1985, respectively. In 1985, he joined the Electrical Communication Laboratories (NTT). From 1989 to 1990, he was a Visiting Scholar at the Robotics Laboratory, Stanford University. He has been conducting research in the fields of artificial intelligence and the Semantic Web, and serves as a member of academic committees including the Organizing Committee for the International Semantic Web Conference 2004 (ISWC-2004). He is a member of IEEE, IPSJ (the Information Processing Society of Japan), and the Japanese Society of Artificial Intelligence.

Kaoru Hiramatsu

Member, Social Communication Laboratory, NTT Communication Science Laboratories.

He received the B.S. degree in electrical engineering and the M.S. degree in Computer Science from Keio University in 1994 and 1996, respectively, and the Ph.D. in Informatics from Kyoto University in 2002. In 1996, he joined NTT Communication Science Laboratories and has been working on Digital Cities and the Semantic Web. Since 2003, he has been a Visiting Research Scientist at the Maryland Information and Network Dynamics Laboratory, University of Maryland. He is a member of the Association for Computing Machinery, IPSJ, and the Japanese Society of Artificial Intelligence.

Akira Sugiyama

Member, Social Communication Laboratory, NTT Communication Science Laboratories.

He received the B.S. and M.S. degrees in computer science from Tokyo Institute of Technology in 1994 and 1996, respectively. In 1996, he joined the NTT Communication Science Laboratories and has been working on the Spoken Dialogue System and the Semantic Web. He is a member of IPSJ and the Association for Natural Language Processing.

Yutaka Yanagisawa

Member, Social Communication Laboratory, NTT Communication Science Laboratories.

He received the B.S., M.S., and Ph.D. degrees in information systems engineering from Osaka University in 1994, 1996, and 1998, respectively. In 1998, he joined NTT Basic Research Laboratories. In 1999, he moved to NTT Communication Science Laboratories and has been working on moving object databases. He is a member of IPSJ, the Japanese Society of Artificial Intelligence, the Japan Society for Software Science and Technology, and the Virtual Reality Society of Japan.

Tetsuji Satoh

Executive Manager, Social Communication Laboratory, NTT Communication Science Laboratories.

He received the B.E. degree in electrical engineering from Yamanashi University in 1980 and the Ph.D. in Electrical Engineering from Osaka University in 1994. In 1980, he joined the Electrical Communication Laboratories (NTT). He has been conducting research in the fields of Database and the Semantic Web, and serves as a member of academic committees including the Information and System Society Steering Committee of the Institute of Electronics, Information and Communication Engineers (IEICE). Since 2002, he has been a Visiting Professor at Osaka University. He is a member of IEICE and IPSJ.