Topic map for Topic Maps case examples

Motomu Naito

Knowledge Synergy Inc., 3-747-4-203 Kusunokidai Tokorozawa,
Saitama 359-0037, Japan

motom@green.ocn.ne.jp
http://www.knowledge-syergy.com

Abstract. When developing topic maps and their applications, key challenges are how to pick up the main subjects in targeted domains and how to systematize those subjects. This paper introduces a topic map development about topic map case examples. It also introduces what kinds of subjects were extracted and how the identifiers of those subjects were given and how those subjects were classified in the first version. Then the difficulties which were emerged during the development are discussed. In order to promote sharing of the case examples and make good use of them, I provide some consideration and future works.

Keywords: Topic map development, subject, subject classification, subject systematization, Topic Maps case example

1 Introduction

Potentiality and practicability of Topic Maps attract many people increasingly. More and more Topic Maps case examples have been developing by many people. But it is difficult for us to search out the case examples which we really want to find. Many presentation documents have published from such as Topic Maps 2007, 2008, TMRA, AToMS conference web site. Mostly those web sites only enumerate the abstracts of presentation and have links to the presentation documents. If those presentations can be navigated and accessed according to specific subjects and viewpoints, convenience, availability and usefulness of those web sites will increase significantly.

Many people often ask what kinds of case examples exist and where those case examples can be found. In order to answer those questions, to share the case examples in Topic Maps community and with new comers, and to find expected case examples easily, I have been developing a topic map for Topic Maps case examples. At the start of the development, 67 presentations at Topic Maps 2007, TMRA 2007 and AToMS 2007 were targeted. According to Steven R. Newcomb [5], Topic Maps activity started to try to make master index for many documents. Similarly this try can be said as the effort to make master index for Topic Maps related presentations, activities, products, etc. throughout cyberspace.

In this paper, I address the challenges how to classify Topic Maps case examples and how to make easy to find target case examples. In the domain, I consider what main subjects are and how to organize those subjects. Concerning the topic map about Topic Maps case examples, the remainder of this paper is the following. In section 2, analogical topic maps are given. In section 3, developing process and the result until now are described. In section 4, problems are studied. Finally conclusion and future work are showed in section 5.

2   Related topic maps

Topic map for conference proceeding was provided at XML 2001 conference in Orlando, USA. And it got into the news of those days. That topic map had topic types such as author, presentation title, country, keywords, etc. and it could be navigated from the viewpoint of those topic types. The topic map did not include information of multiple conferences but only included information of that conference.

TOPICMAPS.COMMUNITY's website [9] is the website for Topic Maps related information. And itself developed based on topic map. It is possible to navigate and access the presentation documents used in conferences such as Topic Maps 2007 and 2008 in Norway. It also has links to other Topic Maps conferences such as TMRA and AToMS. It doesn't have the master index across a number of those conferences.

Fuzzzy.com's website [10] has much information about Topic Maps Portal and Topic Maps Online Application. It has short descriptions of those portals and applications and has links to those sites. It has many tags from some viewpoints but those are not categorize and organized.

All of the above cases don't have sites crossing index. In this paper I will describe one trial to make master index, in other words topic map, across to multiple website.

3 Topic Maps case examples topic map and its web application

In this section, making process, the topic map and its application are described.

3.1 Making process

The first version of the topic map for Topic Maps case examples were created by the following processes.

1. Data collection and analysis

2. Ontology making

3. Topic map making

4. Application development

Data collection and analysis In the first stage, targeted data was 67 presentations at three conferences, Topic Maps 2007, TMRA 2007 and AToMS 2007. Those data was collected on EXCEL with CSV format by hand. The collected items (subjects) can be roughly divided into two groups, the fact data group and the data group picked out from presentation document based on my understanding. The items in fact group are Events, Sessions, Presentations, Persons, Countries, Organizations, and so on. And the items in picked out group are Activities, Products, Purposes, Industrial domains, Target information/knowledge, Providing services, Activity entities, Users, and so on. These items would be strong candidates of the subject for the topic map. The reason those items were chosen was those items seemed to become good view point to navigate the topic map and to find wanted case examples.

Ontology making Ontology was made according to the collected items (subjects) and relations between them. Fig. 1 shows the ontology diagram. In Fig. 1, topic types are represented by squares and association types are represented by lines.
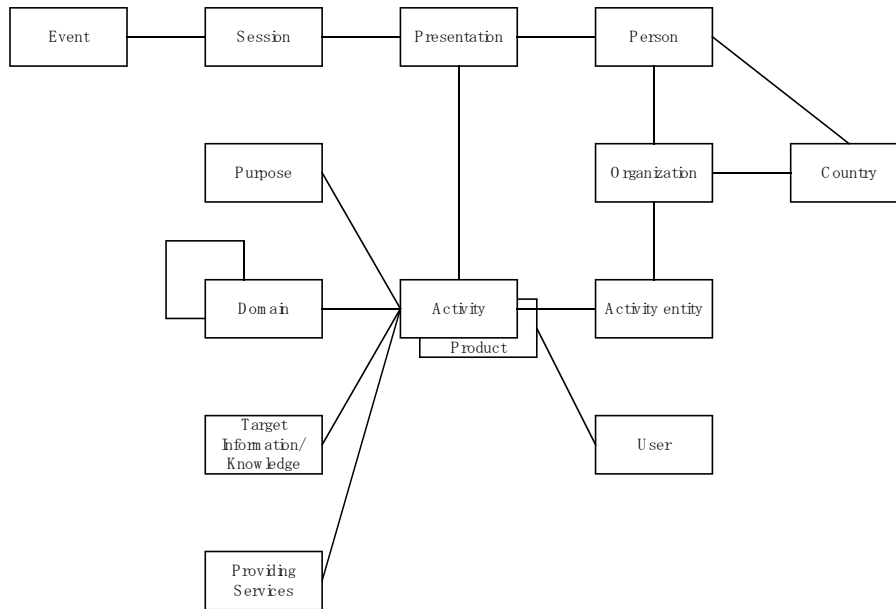
Fig. 1. Ontology diagram of Topic Maps case examples topic map

In this trial, it was differentiated the Presentation topic type and the Content topic type of presentation (it was represented Activity topic type or Products topic type.) Each instance of the Presentation topic type identifies the each presentation. And it has close relations between topic types such as Event, Session, Person, Country and Organization. Each instance of the Content topic type identifies the each Activity or Product mentioned by the presentation and it has close relations between topic types such as Purpose, Industrial domain, Target information/knowledge, Service providing, Activity entity, Users.

Topic map making Using the collected data as input and based on above ontology, the Topic Maps case examples topic map was generated. Specifically, the topic map was generated using the DB2TM module included in OKS (Ontopia Knowledge Suite)™ [6]. The details of the topic map are described in section 3.2.

Application development In order to display and navigate the topic map, a web application was developed. The web application was developed according to the topic map and using Navigator Framework function of OKS. The details of the web application are described in section 3.3.

3.2   Topic Maps case examples topic map

In order to use DB2TM module, the ontology definition file and the XML configuration file were made. Topic types, Association types, Association Role types and Occurrence types were defined in the ontology definition file with LTM [7] format. The mapping rules from collected data with CSV format into the ontology definition file were described in the XML configuration file. And then according to the ontology file and the XML configuration file, Topic Maps case examples topic map was generated by batch process of DB2TM module.

As topic type there are Event, Session, Presentation, Person, Country, Organization, Activity, Product, Purpose, Industrial domain, Targeted information/knowledge, Providing services, Activity entity and User in the topic map. Gathered information is 67 presentations from Topic Maps 2007, TMRA 2007 and AToMS 2007. And structure of the topic map was corresponded the Ontology which was showed in Fig 1. At the moment, version 1.0, the numbers of Topic Maps components were showed in Table 1 and Table 2.

Table 1. The number of types

Type	The number of types
Topic	17
Association	15
Association Role	30
Occurrence	1

Table 2. The number of Instances

Instance	The number of instances
Topic	682
Association	1094
Occurrence	67
Total	1843

3.3   Topic Maps case examples topic map application

The application was developed using OKS Navigator Framework. The Navigator Framework is based on the Java 2 Platform, Enterprise Edition (J2EE), using the Java Servlets and Java Server Pages (JSP) technologies. It is said that OKS makes possible to develop Topic Maps based web applications rapidly and easily. It consists of a set of JSP tag libraries, and a Java API. Applications developed with it can be deployed into any J2EE container. In Topic Maps web applications, we can navigate related topics according to associations in a subject centric way.

Fig. 2. Country's web page

The application consists of about 20 JSP programs at the moment. It makes possible to start to navigate inside the topic map from each topic type described in section 3.2. It can display the instance list of each topic type and the details, (namely associations and occurrences) of each instance topic. It also has the functions such as Sort, Count, Full text search, Graphical representation and so on. Fig. 2 shows Country's web page generated by the application. Fig. 3 shows graphical representation of the topic map.
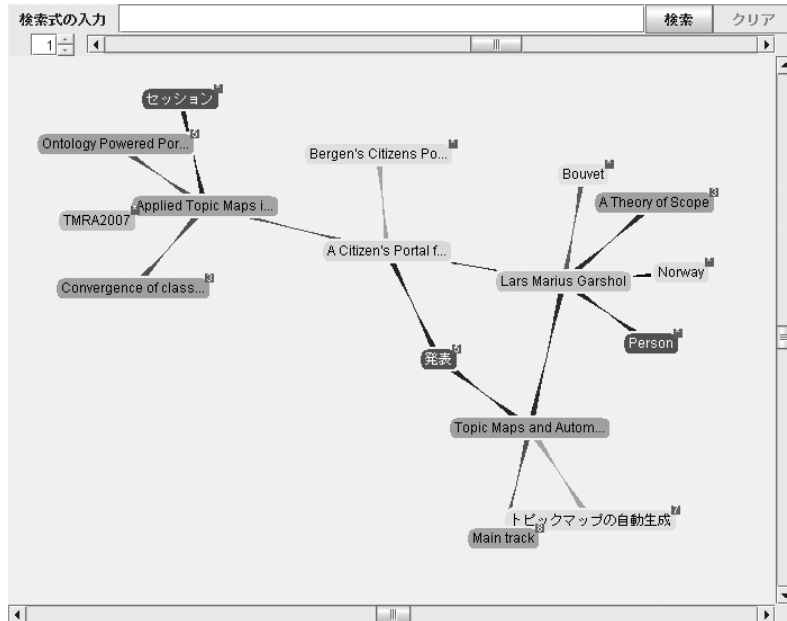
Fig. 3. Graphical representation of the topic map

The application brings many interesting results. Table 3, Table 4, Table 5 and Table 6 show Country basis ranking, Person basis ranking in other words the most frequent speaker ranking, Organization basis ranking and Industrial basis ranking for each. Norway wins the first prize. The most frequest speakers are Lars Marius Garshol, Markus Ueberall, Michihiko Setogawa and Sam Gyum Oh. The most frequent organization is Bouvet. The top industrial domain is Information and Communications. Those are results of tolog [8] query with the functions such as Sort and Count and be picked up from the web application.

Table 3. Country basis ranking

Ranking	Country	The number of presentation
1	Norway	16
2	Germany	15
3	Japan	10
4	USA	7

Table 4. Person basis ranking

Ranking	Person	The number of presentation
1	Lars Marius Garshol	3
1	Markus Ueeberall	3
1	Michihiko Setogawa	3
1	Sam Gyun Oh	3

Table 5. Organization basis ranking

Ranking	Organization	The number of presentation
1	Bouvet	8
2	Hitachi System and Services	3
2	J.-W.-Goethe University	3
2	National Institute of Informatics	3
2	Networked Planet	3
2	Ontopedia	3
2	Sungkyunkwan University	3
2	University Leipzig	3

Table 6. Industrial domain basis ranking

Ranking	Industrial domain	The number of presentation
1	Information and communications	38
2	Education-Learning support	15
3	Government	7

4 Issues and discussion

Some issues became clear through this experience. Those are regarded as general problems in topic map creation.

4.1   Coding scheme of Subject Identifier

The first issue is what kind of coding scheme is suitable for Subject Identifier, especially for fragment of IRI (Internationalized Resource Identifier). It is easy to assign serial number within the limits of collected data. But generally it is not a good way because it lacks the generality, the scalability and the reusability, and it is not intuitive and not friendly for human. If there is already authorized code system such as country code, it is appropriate to use it. The issue is in the case of we can not find those code system, for example presentation identifier, person identifier, etc. I am using conference name + serial number for presentation identifier (example: TMRA2007-1, TMRA2007-2...) and family name + first name for speaker identifier (example: MaicherLutz) at the moment, but these code systems don't seem the best. Those identifiers include some problems such as synonym and homonym problem. I have to seek after a better system from now on.

4.2   Classification scheme

In this work the most difficult part is to build a classification system. If I can categorize some subject according to human's conceptual system, it seems to become easy to navigate intuitively along the classification system. If I know the existence of a suitable classification system, I can use it. If I don't know the existence of that, I have to devise a classification system by myself. I could find a suitable classification system for Industrial domain for this work. I used Japan Standardized Industrial Classification.   I needed only to map the industrial domain of presentations to the classification. Therefore it was relatively easy work. Japan Standardized Industrial Classification is four layered classification. Those layers are L category, M category, S category, T category. The L category is the following:

  A: Agriculture
  B: Forestry
  C: Fisheries
  D: Mining
  E: Construction
  F: Manufacturing
  G: Electricity-Gas-Heat supply and Water
  H: Information and Communications
  I: Transport
  J: Wholesale and Retail trade

K: Finance and Insurance
L: Real Estate
M: Eating and Drinking places- Accommodations
N: Medical - Health Care and Welfare
O: Education-Learning support
P: Compound Services
Q: Services- N.E.C.
R: Government- N.E.C.
S: Industries unable to classify

I could not find a suitable classification system for Activity, Purpose, Targeted information/knowledge, Providing service, etc. Therefore the classifications of those subjects were very difficult. For the work I assigned appropriate words for the subject of the presentations, and then made effort to classify those words. I think those are similar process to KJ Method. I'd like to continue the effort to search for the method of classification as well as to develop the good method to build classification system.

4.3   Appropriate metadata for posting

I picked out suitable words from presentation documents for Purpose, Targeted information/knowledge, Service providing, User, etc. from my point of view. And still I'm taking great pains over building of those classification systems. Like as Industrial domain, if there are good classification systems authors can select and attach suitable category to their activities as metadata. They can publish the activities with the metadata. In result more appropriate classifications become possible. Therefore I think publication with appropriate metadata is very meaningful. I think we need to construct and share common vocabularies for those metadata.

5   Conclusion and Future Work

As the first step of the developing, it became possible to navigate 67 presentations from the three conferences. It also became possible to access those documents easily and I can use the topic map for my Topic Maps activity usefully. I can reply the questions about Topic Maps case examples in my Topic Maps popularization activity.

As future work I am planning the following:

- Review and improve the ontology
- Add for more viewpoints
- Review and improve Identifier coding scheme
- Review and improve Classification system

I think those works lead to improvement the topic map itself. After the topic map and its application make the pass mark, I'm planning to open the topic map through website as well as add more presentation from other conferences, an individual case examples and others one after another. Moreover I wish it becomes possible to discuss classification system and cooperate and merge with other topic maps in the open environment.

References

1. Steve Pepper: Topic Maps for the Three Kingdoms : The Many Applications of Topic Maps, AToMS Asian Topic Maps Summit, 2nd June, 2006. pp 37-64

2. Steve Pepper: The TAO of Topic Maps : Finding the Way in the Age of Infoglut, `http://www.ontopia.net/topicmaps/materials/tao.html.`

3. Steve Pepper: As We REALLY May Think : Memex, Topic Maps, and subject-centric computing, `http://www.ontopedia.net/pepper/slides/AToMS2007.ppt.`

4. Steve Pepper, Sylvia Schwab: Curing the Web's Identity Crisis : Subject Indicators for RDF, `http://www.ontopia.net/topicmaps/materials/identitycrisis.html.`

5. Jack Park, Sam Hunting: XML Topic Maps : Creating and Using Topic Maps for the Web, ISBN0-201-74960-2.

6. Ontopia: Software Products : The Ontopia Knowledge Suite, `http://www.ontopia.net/solutions/products.html.`

7. Ontopia: The Linear Topic Map Notation : Definition and introduction, version 1.3, `http://www.ontopia.net/download/ltm.html.`

8. Ontopia: tolog : Language tutorial, `http://www.ontopia.net/omnigator/docs/query/tutorial.html`

9. TOPICMAPS.COMMUNITY : `http://www.topicmaps.com/tmc/`

10. Fuzzzy.com : Topic Maps Portal or Online Application `http://www.fuzzzy.com/tag/?id=2238`

11. Riichiro Mizoguchi: Science of Intelligence: Ontology Engineering, Ohmsha, ISBN4-274-20017-5.

12. Motomu Naito, Hiroyuki Kato, Takashi Kiriyama, Yushi Komachi, Mi Setogawa, Keiji Nakabayashi, Mitsuo Yoshida: An Introduction to Topic Maps, Tokyo Denki University Press, ISBN4-501-54210-1.

13. Lars Marius Garshol: Metadata? Thesauri? Taxonomies? Topic Maps: Making sense of it all!
    `http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html.`

14. Frederic Andres and Motomu Naito: Dynamic Topic Mapping Using Latent Semantic Indexing"in IEEE International Conference on Information Technology and Applications Sydney 4th-7th July, 2005. pp 220-225

15. Motomu Naito and Frederic Andres: Application Framework Based on Topic Maps, First International Workshop on Topic Maps Research and Applications, TMRA 2005 Leipzig, Germany, October, 2005. pp 42-52

16. Sachit Rajbhandari, Frederic Andres, Motomu Naito, Vilas Wuwongse : Semantic-Augmented Support in Spatial-Temporal Multimedia Blog Management, Second International Conference on Topic Maps Research and Applications, TMRA 2006 Leipzig, Germany, October, 2006. pp 215-226