A Perspective of Fusing Ontology and Metamodeling Architecture in Interconnection Environment

Jin Liu[1,2], Keqing He[1], Fei He[3], Bing Li[1]

[1]The State Key Laboratory of Software Engineering (SKLSE), Wuhan University
43072 Wuhan, P.R.China
[2]Institute of Computing Technology, Chinese Academy of Sciences,
100080 Bejing, P.R.China
[3]Graduate School of Science and Engineering, Waseda University
169-8555 Tokyo, Japan
`xzhliu@whu.edu.cn`

Abstract. To achieve interconnection environment, the proposed approach consists of MMFI for ontology and reference ontology for normalizing semantic space, ontology-induced schema mapping for transition from semantic interconnection to schema mapping, MMFI (ISO 19763) for normalizing interconnection environment, and a semantic matcher with schema mapping broker for dynamically clustering services or resources. In this case, while ontology helps identifying content of structural elements, a metamodeling architecture provides requisite structures for ontology representation. Our approach promotes fusing of ontology and metamodeling architecture and enhances interoperability in interconnection environment. This approach has been implemented in WHCRP prototype as a potential demonstration for MMFI. WHCRP aims to establish a R&R federation for sharing software components. Theoretical analysis and the prototype indicate the effectiveness of this approach.

1 Introduction

The emergence of the Internet pushed computing environment into distributed interconnection environment. In this environment, versatile resources are loosely connected to each other, which makes knowledge diffusion, resource sharing and distributed collaboration possible. People hope that intelligent services can be provided so that service results can be pushed to users according to their requirements in a more accurate, effective and efficient way [1,2]. In this case, the competence of machine inference should be elevated to deal with this challenge. However, several characters of the WWW limit achieving this vision. The most important one is that vast versatile resources are unstructured Web documents that can only be understood by humans, which makes it impossible for machines to draw inferences about them.

An expected interconnection environment should enable interoperation among versatile resources based upon machine inference. As the capability of different information systems to communicate, interoperability usually relates to semantics and structures, because any communication should be semantically correct and machineprocessible information is usually organized in terms of structural schema. Therefore, at least four problems should be solved to achieve the expected interconnection environment: (1) Semantic interconnection; (2) Transition from semantic interconnection to structural schema mapping; (3) Appropriate normalization of semantic space or structural schema; (4) Dynamically clustering services or resources.

Current prevalent approaches to achieve semantic interconnection is to use markup languages to annotate data with semantic tags so that machines can recognize intended meaning and use rules for manipulating semantic information[3]. This manner has a preference for ontology, i.e. a logical theory explaining the intended meaning of a formal vocabulary. Mapping between different schemata is also crucial in interoperability. Schema is mainly used to refer to implemented structure of data, not so much to indicate its intended meaning. Data normally takes a well-defined structure with formal interpretations, so that machines can operate on it with corresponding algorithms. These schemata can be produced from a metamodeling architecture [4]. Recently there is a trend of adopting ontology and metamodeling architecture together to solve problems in interconnection environment. However, even people perceive the implicit relationship between ontology and metamodeling architecture, they hardly illuminate how to establish the collaboration between them.

We gain a growing awareness that ontology and metamodeling architecture complement each other in the expected interconnection environment. Ontology helps to understand the intended meaning of information artifacts as well as content of structural elements in a schema. A metamodeling architecture produces requisite structure for representing information artifacts as well as structure for representing formal ontology. This paper illuminates our efforts to achieve the expected interconnection environment under the circumstance of a prospective demonstration for MMFI, i.e. WHCRP (WuHan Component Repository Platform) [5]. ISO/IEC/SC32/WG2 19763 Information Technology—Framework for Metamodel Interoperability (MMFI) is an ISO project that aims at using ontology and a metamodel framework to realize registering information resources and promote integration, sharing and collaboration of heterogeneous data or services [5]. The intention of WHCRP is to establish a MMFI-based federation for sharing software components in interconnection environment.

To solve the first problem, this paper proposes MMFI for ontology and reference ontology to establish semantic interconnection [5]. As the authors of MMFI for ontology (MMFI Part 3), we design the primitive structures via metamodeling process to represent formal ontologies for registration, and the mechanism of registering & managing these ontologies. Software components registered according to these ontologies are assigned semantic tags, so that the semantic interconnection environment is established. Moreover, the introduction of reference ontology normalizes semantic space and guarantee recognizing semantics in this environment.

This paper also proposes ontology induced schema mapping and MMFI for mapping to solve the second problem. When resources are located via identifying semantic information, the result should be pushed to users with the expected schema. Schema mapping performs the transformation from local schema to target schema. MMFI for mapping (MMFI Part 4) intends to record mapping relationship between structural elements of two schemata that can be described by MMFI for model structure (see 3.1). If such mapping doesn’t exist, ontology induced schema mapping associates ontology term with structural elements of schemata and helps to identify mapping relationship between structural elements from different schemata. We developed a software tool to assist establishing mapping manually.

MMFI as an ISO specification and reference ontology are efforts to solve the third question. This paper proposes semantic matcher and schema mapping broker in the federation based on MMFI to solve the last question.

In this way, our approach promotes fusing of ontology and metamodeling architecture to enhance their adaptability to the expected interconnection environment. The results from theoretical analysis and the prototype of WHCRP demonstrate the effectiveness of our approach to achieve the expected interconnection environment.

2 Ontology and Metamodeling Architecture

Ontology in computer science has been used for information understanding, knowledge sharing and reasoning. Gruber defines ontology as a specification of a conceptualization [6]. Guarino further explained that the intended models of a logical language with a formal vocabulary are constrained by its ontological commitment to a particular conceptualization [7]. Intelligent computing entities in interconnection environments should be semantics-aware to adapt to interconnection environment. Ontology is closely relative to semantics recognition and is pivotal for semantic interconnection.

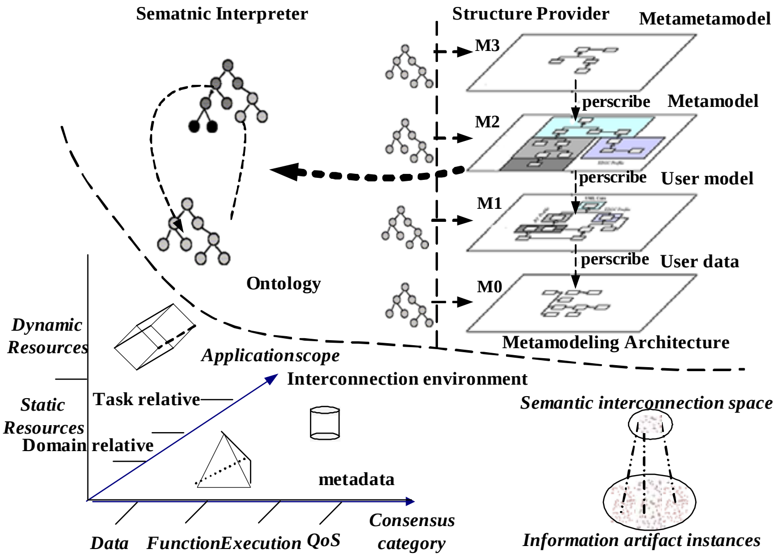

Fig. 1. Fusing of ontology and a metamodeling architecture in interconnection environment

A metamodel is a model that describes allowable model constructs, not so much to do with model content. It can be a conceptual model for a set of models, or a model for a modeling language. Metamodeling is the process of defining metamodels with a model referenced as metametamodel. Generally, metamodeling prescribes metamodels and a metamodel prescribes models. In this way, a hierarchical metamodeling architecture is established and produces structures representing information artifacts. For software development, metamodeling follows this process: A part of modeling knowledge is formalized as explicit modeling abstracts and is absorbed into metamodels. Tools that implement these metamodels assist people in software development. Tolvanen recorded early metamodeling research and summarized metamodeling process [4]. Besides software models, a metamodeling architecture can also produce structures for domain models, business models, data models, etc. MOF metamodeling architecture maintained by OMG (Object Management Group) is a prevalent metamodeling architecture that is consisted of several specifications such as Meta-Object Facility (MOF) and Unified Modeling Language [8].

While ontology helps understanding content of structures from a metamodeling architecture, a metamodeling architecture produces structures to represent formal ontology. Fig. 1 indicates that collaboration of ontology and a metamodeling architecture is beneficial to promoting adaptability of information artifacts in interconnection environment. It is also possible for ontology to indicate domain knowledge that facilitates understand model structures in each level of a metamodeling hierarchy.

3 Semantic Interconnection

We propose MMFI for Ontology for semantic interconnection and reference ontology for semantic normalization. MMFI is an ongoing project that aims at achieving interoperability between metamodel-based applications by registering information artifacts with ontology and a metamodel framework [5].

3.1 MMFI for Ontology

Semantic heterogeneity occurs when the intended meanings of information artifacts is uncertain. This can be mitigated with machine-recognizable semantic tags, e.g. term from an ontological vocabulary that provides a criterion of common understanding. MMFI is designed as a metamodel framework for ontology registration. These ontologies register information artifacts and annotate semantic tags to them. Thus a MMFI-based semantic interconnection environment is established.

MMFI is consisted of Model Core, MMFI for Ontology, MMFI for Mapping and MMFI for Constructs. Model Core inherits MOF metamodels and MDR (ISO/IEC 11179-3) metamodel and specifies a basic structural framework for metamodel and model registration. As a registration framework, Model Core reuses MDR structures that prescribe the general attributes of registry objects. Meanwhile, MMFI Core inherits metaclasses from MOF in a normative way. The other parts in MMFI inherit metaclasses from Model Core and add new metaclasses to meet the registration requirements. As the editor of MMFI for Ontology, we plan to provide primitive structures to represent formal ontologies and the registration mechanism for ontologies.

There are two relationships between MMFI for Ontology and Model Core. While MMFI for Ontology inherits registry framework from Model Core, registered ontologies can support classification for information artifacts. Fig. 2 shows the relationship between Registry Model from MMFI Core and Kernel Model from MMFI for Ontology. For Registry Model, we define metaclass ModelInstance to represent registered models and metaclass ModelConcept to represent Model_Insatnce. A ModelConcpet prescribes ModelInstances. ModelConcept is an abstract concept and its content is described by ModelDomainProfile. ModelSign is a symbol designating a named element in a namespace. Meaings of ModelInstance is specified by ModelSign. ModelSelection specifies the condition of choosing ModelInstance. Ontology from Kernel Model inherits ModelDomainProfile from Registry Model, Ontology_Component inherits ModelComponent and Ontology_Atomic_Construct inherits ModelClassifier. Basic structures that describe formal ontology are similar. Generally, ontology, sentence and symbol are basic structures, where ontology is consisted of sentences and a sentence is consisted of symbols. Taking ODM (Ontology Definition Metamodel) as a reference, we defined structures in MMFI for Ontology, i.e., Ontology, Ontology_Component and Ontology_Atomic_Construct. Mapping among different ontological language depends on related work in ODM [9].

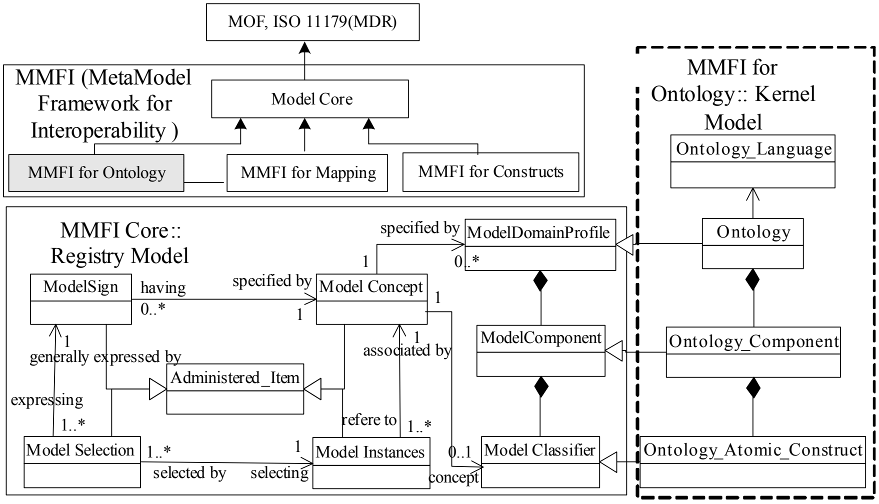

Fig. 2. Registry Model and Kernel Model

3.2 Normalizing Semantic Space

Normalizing semantic space is to correctly reflect semantic information so as to ensure integrality of information and correctness of operation in semantic space. Semantic interconnection abstracts semantics into a uniform image in semantic layer, so that people or service can efficiently and correctly operate resources according to normative semantics. Therefore, we introduce reference ontology to MMFI for Ontology.

As a semantic criterion published by authoritative administrators, reference ontologies specify a normative semantic space in MMFI-based interconnection environment. Local ontologies as the result of reusing main body of existent normative ontologies are allowed in semantic layer. Modification includes adding, deleting ontology terms, or combining normative ontology into a new one. Thus ontology evolution educes local ontology from reference ontology. We define Reference_ Ontology and Local_Ontology in MMFI, where Reference_Ontology can be reused by Local_Ontology to satisfy individual requirements. Accordingly, Reference_ Ontology_Component can be reused by Local _Ontology_Component and Reference_ Ontology_Atomic_Component can be reused by Local_Ontology_Atomic _Component. Thus we can identify semantic association between Local_Ontology and Reference_Ontology or association among different Local_ Ontology based on the same Reference_Ontology. Fig. 3 shows ontology instances that follow Kernel Model, i.e. reference ontology SC_RO and local ontology A in conceptual graph.

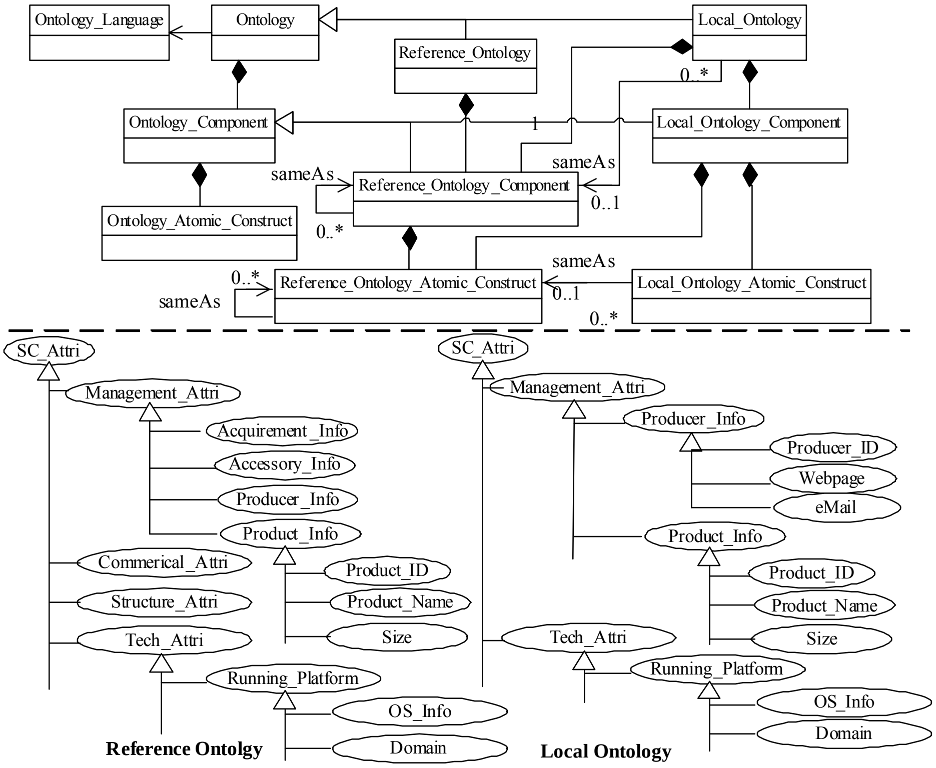

Fig. 3. Overview of Kernel Model and ontology instances

Ontology A inherits concepts from SC_RO and adds new concepts. In order to describe the evolution information of Local_Ontology, we further define Evolution Information Model and an Evolution Rule Model in Fig. 4. The complex ontology evolution is composed of single steps. Evolution_Info records evolving steps of local ontology and Evolution_Info_Item records each step. Rules implemented in the evolvement are defined in Evolution Rule Model. Evolution Rule Model depends on Kernel Model and Evolution Information Model. According to our research, ontology evolution can be reduced to two categories. One is that evolution only changes ontology's nomenclature. The other is to convert an existent ontology into a new one by adding properties to it. Thus we design Renaming_Rule and Composition_Rule in

Evolution Rule Model. Evolution Rule Model provides a framework to describe the rules that produce Local Ontologies without semantic collision with Reference Ontology. Thus semantic matching could be performed to recognize intended meanings.

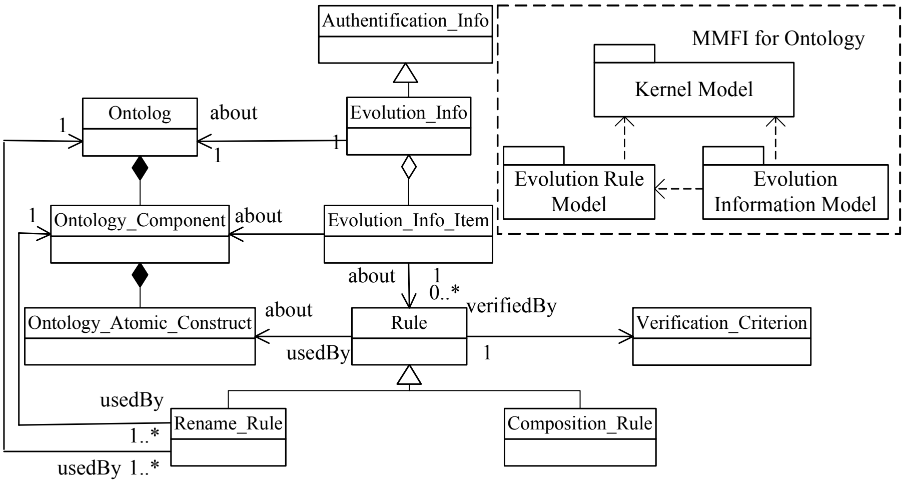

Fig. 4. Framework for Ontology Evolution

4 Ontology Induced Schema Mapping

In MMFI-based interconnection environment, applications congregate resources or services to achieve relevant tasks and push service result to clients according to their requirements. The basic service provided by MMFI-based federation of software components (SC for short) could be distributed query service (see 5.1). Distributed registry applications may record registered information artifacts with local data schema. If the intended meanings of registered SCs was identified via semantic match, the result of query would be pushed to clients with expected data schema. Thus a mapping between local and target schema may be required. MMFI for mapping intend to record mapping between schemata described by MMFI for Construct. If a schema mapping is not immediately available, it should be established in time to achieve interoperability. The relationship between structural elements from two schemata can be identified according to their content. Generally, there are two ways to identify content of schema's structure elements. One is to estimate content according to structural similarity. However, this approach helps to, but not be sufficient to determine semantic similarity. For two properties with the same structure, their contents may be quite dissimilar. Another approach is to characterize content information with terms from domain ontology. Following this approach, we propose ontology-induced schema mapping in interconnection environment.

In our approach, structural elements of schemata are associated with ontology terms, which helps to identify mapping relationship between structural elements from schemata. Conceptualization makes it possible to capture content information. Since it is impossible to determine all contents with a fixed descriptor set, we attach a partial term representation to structural elements. OWL (Web Ontology Language) based on DL (Description Logic) can represent content expressions and describe general concept properties [10]. The content expression can be described as a scalable vector:

```
Content IR_Name
unit:
     conceptualization:  term;
      structural_Element: element;
        restriction:      {optional}
endunit
    ............
unit:
     conceptualization:  term;
      structural_Element: element;
        restriction:      {optional}
endunit
End
```

Or `Content(IR_Name)` $= (<C_1, E_1, R_1>, \ldots, <C_n, E_n, R_n>)$, where C_i ($1 \leq i \leq n$) is the content dimension specifing IR content, E_i ($1 \leq i \leq n$) the structural dimension w.r.t content dimension, R_i ($1 \leq i \leq n$) the optional restriction on E_i. $<C_1, \ldots, C_n>$ is term sequence selected from ontologies to capture content of structural elements.

To stand out our work, we adopt XML to represent and storage information artifacts, i.e. XML document for information artifact instance and XML schema for structural schema. We emphasize that ontology terms different from structural element name in that ontology is conceptualization of the world, whereas name may not. By associating ontology terms to structural elements, their content can be explicit.

To implement ontology-induced schema mapping, we have developed a software tool to assistant establishing mapping manually (see 5.2). Mapping exists as XSLT (Extensible Stylesheet Language Transformation) document. It is possible to automatically establish mapping between schemata if their structure is simple. In this case, reasoning on content is crucial for ontology-induced schema mapping. In practice, mapping can be supported by MMFI-based Registry facilities. While there facilities adopt local schema for registered artifacts, content expressions can attach to schemata.

5 Case of Fusion: WHCRP

5.1 Federation of R&R for SC

Since reusing software components (SC) is an effective approach to implement large-scale software production, an intelligent software infrastructure would be established for managing and sharing versatile SCs scattered on the Web. As a prospective demonstration for MMFI, WHCRP (WuHan Component Repository Platform) is an ongoing project in State Key Lab of Software Engineering, Wuhan University.

As a part of MMFI based SC federation, the WHCRP prototype reflects the characters of our approach, i.e. establishing the expected interconnection environment and fusion of ontology and a metamodeling architecture in this environment. In this federation, ontology helps to understand the intended meaning of SCs as well as content of structural elements from schemata. A metamodeling architecture produces requisite structure, i.e. schemata for registered SC as well as structure for formal ontology. Fig. 5 shows our solution of SC R&R (registry and repository) federation based on MMFI. There are three layers in Fig. 5. MMFI layer are Specifications for the interconnection and interoperability among R&R facilities. Ontology & Metamodel layer includes SC attribute ontologies, SC registry metamodel and SC repository mapping metamodel. Registry model layer contains registry model structure. A registry models can be developed for particular purpose. The number of registry items may be different from each other. Moreover, registry models can adopt well-known schemata such as ebXML and UDDI [11]. Relying on ontologies, registry metamodel and repository mapping model, an interconnection environment for SC resources can be established.

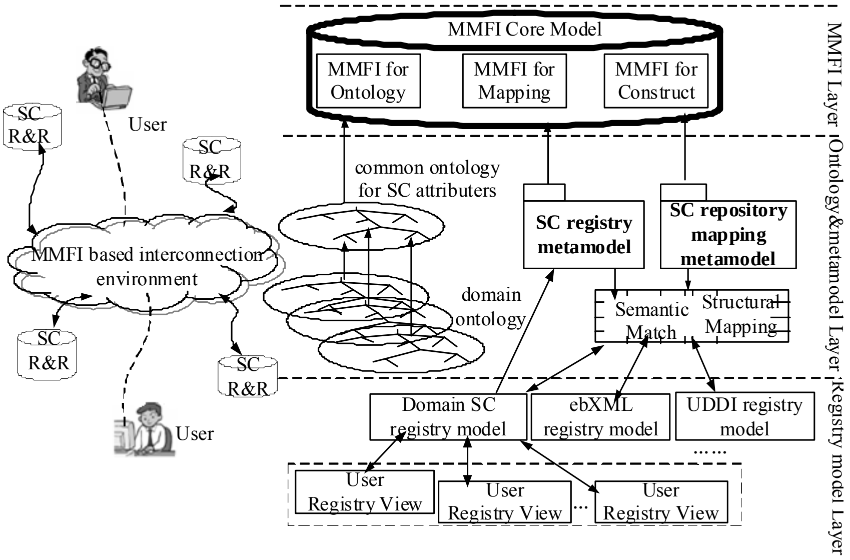

Fig. 5. Federation of R&R facilities for SC

5.2 Scenario

At present, the basic service of WHCRP is distributed query with registered ontology. Fig. 6 presents service of WHCRP with a typical scenario of querying SCs. An authority in the federation is responsible for issuing a reference ontology SC_RO w.r.t SC and registering it at an ontology registry center based on MMFI for Ontology. Users might develop their own local ontology induced from SC_RO, e.g., A, B and C. AS SC_RO variants, A, B and C are also registered at this registry center based on MMFI for Ontology. MMFI for Ontology provides requisite structures for representing formal ontologies. Site 2 and 3 are SC R&R facilities that conform to MMFI, where site 2 takes the well-known ebXML and site 3 takes UDDI as registry schema to register SC attribute information respectively. Meanwhile, site 2 adopts local ontology A and site 3 adopts local ontology B to register SC attributes. Site 3 adopts local ontology C to register SC attributes and records them according to local schema S. We represent all these with XML. In Fig. 6, the querying process comprises the following phases.

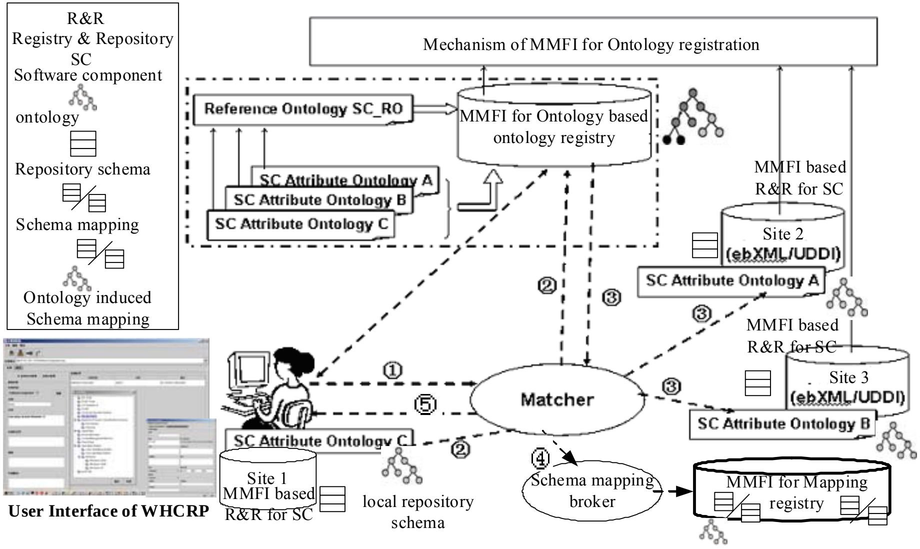

Fig. 6. Typical service scenario of WHCRP

1. A user adopts local ontology C to describe SC features and inputs it as query to the GUI interface at MMFI-based site 1. Site 1 executes local query and sends this request to Semantic Matcher.
2. Semantic Matcher inquires the ontology registry center and understands that C is local ontology induced form reference ontology SC_RO.
3. Semantic Matcher forwards query to other sites and finds that ontologies at site 2 and site 3 are SC_RO variants. It compares A with B and C, and judge B and C are semantically approximate if similarity degree between C and A is acceptable. Site 2 and 3 execute local query and fetch results with local schemata.

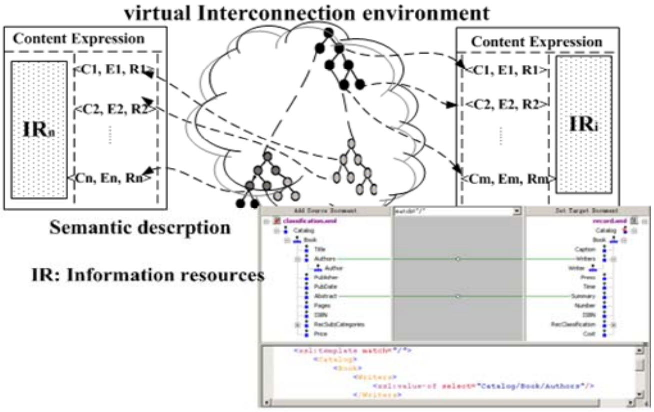

Fig. 7. Ontology-induced schema mapping and a tool prototype

4. Semantic Matcher informs Schema-Mapping Broker to inquire corresponding mapping in registry center of MMFI for mapping and transforms local registered

information into schema at site 1. If no such mapping exists, Schema-Mapping Broker executes ontology-induced schema mapping automatically.

5. The Semantic Matcher pushes the results to site 1. Site 1 unites the data sets and returns the final result to the user.

Reference ontology SC_RO and local ontology A can be found in Fig. 3. Further we take a simulation to simplify this scenario and replace UDDI and ebXML with two terse schemata, i.e. M and N. In this case, no mapping relationship between S and M exists. Since structures of S and N are quite simple, ontology-induced schema mapping can be executed automatically. Fig. 7 indicates that mapping relationship can be established manually with the assistance of our tool if schemata are complicated.

6 Related Work

Zhuge and the China Knowledge Grid Research Group have made many efforts to explore the operating principles in the future interconnection environment [12]. They proposed theory of normative semantics and corresponding algorithm. They also developed an experimental platform for knowledge grid that includes a multi-dimension knowledge model and knowledge operation language [12-14].

OMG is developing ODM to support ontology applications with metamodel. ODM specifies the mapping among metamodels of popular ontological languages such as OWL, RDF(Resource Description Framework) and SCL (Simple Common Logic) [9].

Evermann and Wand extended Bunge's Ontology and proposes a method to restrict syntax of a modeling language [15]. This is also an approach to interpret structure with ontology. Unlike us, this method captures relationships among domain elements via constraints on the language metamodel, thus restricting the set of statements about the domain that can be generated with the language.

Sheth et al aim to extend specifications of Semantic Web techniques to promote dynamic clustering and to achieve more dynamism and scalability [16]. These techniques include Business Process Execution Language for Web Services, Web Service Description Language and Simple Object Access Protocol.

7 Conclusions

The interconnection environment is a trend for future computing environment. To achieve this, we propose four problems as the characters of the expected interconnection environment. MMFI for ontology and reference ontology are proposed to establish semantic interconnection. Ontology induced schema mapping and MMFI for mapping are proposed to achieve transition from semantic interconnection to structural schema mapping. The approach has been implemented in WHCRP prototype, which aims to establish a SC R&R federation and provides access to distributed SC resources and services.

The characters of our approach are: (1) established MMFI-based infrastructure for interconnection environment and developed WHCRP as potential demonstration for

MMFI (ISO 19763); (2) proposed reference ontology and local ontology for semantic normalization; (3) proposed ontology-induced schema mapping that identify registered content of structural elements with ontology terms to achieve schema mapping. This approach promotes fusing ontology and metamodeling architecture and enhances adaptability of information artifacts in the expected interconnection environment. The results from theoretical analysis and WHCRP indicate that our approach is effective.

The future work includes: (1) establish a sound registry center of reference ontology to support normative semantic interconnection environment; (2) promote automatic schema mapping with complex schema; (3) optimize query reformulation in the SC federation to improve query efficiency; (4) extend WHCRP into practice.

Acknowledgements

This research was supported by ISO/IEC/SC32/WG2 19763 under the Grants ISO 1.32.22.01.03.00; the NFSC under the Grants No. 60373086; the Natural Science Foundation of Hubei under Grants No.2005ABA240 and No.2005ABA123. The authors thank Y. F. He and C. Wang for their help and cooperation.

References

1. H. Zhuge: The Future Interconnection Environment. IEEE Computer.38 (4) (2005)27-33
2. H. Zhuge: Clustering Soft-Devices in Semantic Grid. IEEE Computing in Science and Engineering, 4 (6) (2002) 60-62
3. T. Berners-Lee, et al.: The Semantic Web. Sci-Amer. 284(5) (2001) 34–43
4. J. P. Tolvanen. Incremental Method Engineering with Modeling Tools: Theoretical Principles and Empirical Evidence[Ph.D. thesis], University of Jyväskylä, Finland, 1998
5. Information Technology--Framework for Metamodel Interoperability (ISO/IEC 19763). ISO/IEC JTC1 SC32 WG2 Project. (2005). http://metadata-standards.org/19763/
6. T. Gruber: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal of Human and Computer Studies. 43(5/6) (1995) 907-928
7. Guarino N.: Formal Ontology and Information Systems. In: Guarino N.(eds.): Formal Ontology in Information Systems: Proc. of the 1st Inter. Conf., IOS Press, (1998) 3-15
8. Meta-Object Facility (MOF 2.0). (2004). http://www.omg.org/cgi-bin/doc?ptc/2004-10-15
9. OMG Ontology Definition Metamodel RFP. (2003) http://www.omg.org
10. The Description Logic Handbook: Theory, Implementation, and Applications. In: F. Baader, D. et al. (eds.). Cambridge University Press, (2003)
11. Universal Description, Discovery and Integration (UDDI 3.0). (2005) http://www.uddi.org
12. H. Zhuge: China's E-Science Knowledge Grid Environment. IEEE Intelligent Systems, 19(1) (2004) 13-17
13. H. Zhuge: A Knowledge Grid Model and Platform for Global Knowledge Sharing. Expert Systems with Applications, 22 (4) (2002) 313-320
14. H. Zhuge: Resource Space Grid: Model, Method and Platform. Concurrency and Computation: Practice and Experience, 16 (14) (2004) 1385-1413
15. J. Evermann, Y. Wand: Toward Formalizing Domain Modeling Semantics in Language Syntax. IEEE Transactions on Software Engineering. 31(1) (2005) 21-37.
16. K. Sivashanmugam, K. Verma, A.P. Sheth: Discovery of Web Services in a Federated Registry Environment. Proc of the IEEE Int. Conf. on Web Services. (2004) 270-278