Volume 4, Issue 6, June 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Bootstrapping Ontology in Medical Case Studies

K.Panimalar , V.Gomathi
Assistant Professor CSE dept
SMVEC,Puducherry
India

M.Priyadharshini
B.Tech,CSE dept
SMVEC,Puducherry
India

M.Abiya
B.Tech,Final Year, IT dept
SMVEC,Puducherry
India

***Abstract** -Maintaining ontology remains a daunting task. Bootstrapping the ontology, which aims at automatically generating concepts and their relations in a given domain, is a promising technique for ontology evolution. The approach enables the automatic construction of an ontology that can assist, classify, and retrieve relevant files or services. Since the task of updating ontology remains difficult, our approach can be valuable in practice. The proposal was first put forward for Search Engines, where the ontology had to be updated each time a new WSDL was added in the UDDI. The future work proposed was to apply the Bootstrapping Technique for Ontology in other domains. For our demonstration, we consider the medical case studies done by interns. Whenever they create reports of a patient, the details are entered manually. Hence the work load is quite difficult. Analyzing the possibility of the disorder or problem is done completely by the interns. Here we propose to bootstrap the Ontology of each patient with the help of a knowledge base. This helps to reduce the time to identify the case and prepares reports appropriately. Additionally, this result can be used as an evidence for Dempster-Shafer Theory (DST). We are aware of the fact that DST is not very successful in finalizing a disorder when there is a conflict in diagnosis. The Dempster-Shafer Theory (DST) is a mathematical theory of evidence. It allows one to combine evidence from different sources and arrive at a degree of belief that takes into account all available evidence.*

***Keywords** -- Ontology, Bootstrapping, Web Services Description Language, Dempster-Shafer Theory, Free Text Document*

I. INTRODUCTION

In philosophy, ontology is the study of the kinds of things that exist. The term ontology has largely come to mean one of two related things. In general, ontology is a representation vocabulary, often specialized to some domain or subject matter. More precisely, it is not the vocabulary as such that qualifies as ontology, but the conceptualizations that the terms in the vocabulary are intended to capture. Thus, translating the terms in ontology from one language to another, for example from English to French, does not change the ontology conceptually.

In computer science and information science, ontology formally represents knowledge as a set of concepts within a domain, and the relationships between pairs of concepts. It can be used to model a domain and support reasoning about entities. Ontology are the structural frameworks for organizing information and are used in artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise bookmarking, and information architecture as a form of knowledge representation about the world or some part of it. The creation of domain ontology is also fundamental to the definition and use of an enterprise architecture framework
.

II. ONTOLOGY IN WEB SERVICES

Maintaining ontology remains a daunting task. Bootstrapping the ontology, which aims at automatically generating concepts and their relations in a given domain, is a promising technique for ontology evolution. The approach enables the automatic construction of an ontology that can assist, classify, and retrieve relevant files or services. Since the task of updating ontology remains difficult, our approach can be valuable in practice. The proposal was first put forward for Search Engines, where the ontology had to be updated each time a new WSDL was added in the UDDI. The future work proposed was to apply the Bootstrapping Technique for Ontology in other domains. For our demonstration, we consider the medical case studies done by interns. Whenever they create reports of a patient, the details are entered manually. Hence the work load is quite difficult. Analyzing the possibility of the disorder or problem is done completely by the interns. The Existing System is heavily based on the existing ontology or is domain specific but the Proposed System automatically creates the ontology. We need to provide a knowledge base for bootstrapping the ontology. In case of Web Services, a free-text document was created resembling the WSDLs.

Bootstrapping requires analyzing the web service using 3 parameters:

- Term Frequency / Inverse Document Frequency
- Web Context Extraction Method
- Free Text Description Verification Method

i. **Term Frequency / Inverse Document Frequency (TF / IDF):**
does an internal analysis of a web service, that is, what concept describes the WSDL document.

ii. **Web Context Extraction Method:**
describes the WSDL document from an external point of view, that is , what most common concept represents the answers to the web search queries based on the WSDL document.

iii. **Free Text Description Verification Method:**
is used to resolve the inconsistencies with the ontology constructed after the applying the above two methods. The first and second methods (TD/IDF and Web Context Extraction Method) are based on WSDL document content. Whereas, the third method (Free Text Description Verification Method) is based on the Free Text Descriptor which is similar to a WSDL. The Free Text Descriptor is defined explicitly.

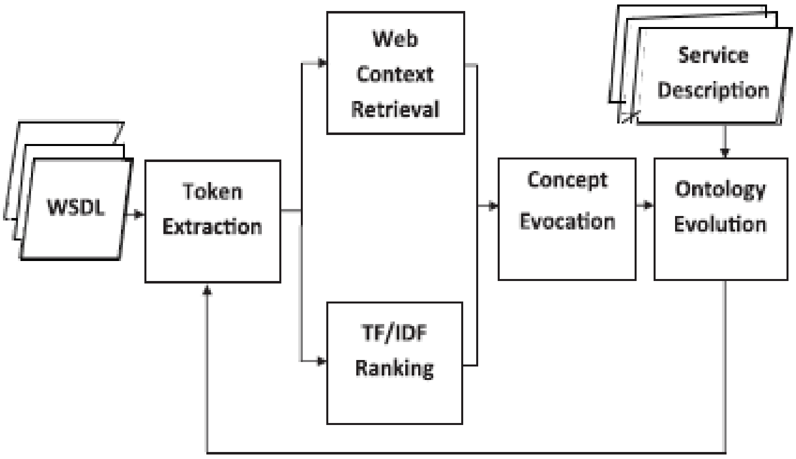

Fig. 1.Webservice ontology bootstrapping process.

III. RELATED STUDY

The lack of semantics in Web Services Description Language (WSDL) prevents automatic discovery and hence automatic invocation and composition. In our work, we are interested in extending existing approaches for the description of Semantic Web Services. This is a recent implementation of automatic ontology generation in web services. It was first demonstrated in search engines. Swoogle is a crawler-based indexing and retrieval system for the Semantic Web documents - i.e., RDF (Resource Description Framework) or OWL (Web Ontology Language) documents. It analyzes the documents it discovered to compute useful metadata properties and relationships between them. The documents are also indexed by using an information retrieval system which can use either character N-Gram or URIs as terms to find documents matching a user's query or to compute the similarity among a set of documents. One of the interesting properties computed for each Semantic Web document is its rank - a measure of the document's importance on the Semantic Web.

IV. ALGORITHMS

In [1], the author proposes a bootstrapping algorithm using the 3 methods discussed above for web services.

1: For each Web service
2: Extract tokens from WSDL
3: TF/IDF_{result} = Apply TF/IDF algorithm to $\mathcal{D}_{wsdl}$
4: $WebContext_{result}$ = Apply Web Context algorithm to $\mathcal{D}_{wsdl}$
5: $PossibleCon_i = TF/IDF_{result} \cap WebContext_{result}$
6: If ($PossibleCon_i \subseteq \mathcal{D}_{desc}$)
7: $Con_i = TF/IDF_{result} \cap WebContext_{result}$
8: $PossibleRel_i = TF/IDF_{result} \cup WebContext_{result}$
9: For each concept pair Con_i, Con_j
10: If ($Con_i \subseteq Con_j$)
11: Con_i subclass Con_j
12: Else
13: $Re(Con_i, Con_j) = PossibleRel_i \cap PossibleRel_j$

Fig. 2.Bootstrapping OntologyAlgorithm for Web Services.

The first step includes extracting the tokens from the WSDL for each web service (line 2). The next step includes applying the TF/IDF and the Web Context to extract the result of each algorithm (lines 3-4). The possible concept, *Possible* Con_i, is based on the intersection of tokens of the results of both algorithms (line 5). If the *Possible* Con_i tokens appear in the document descriptor, D_{desc}, then $PossibleCon_i$ is defined as concept, Con_i. The model evolves only when there is a match between all three methods. If$Con_i = \varnothing$, the web service will not classify a concept or arelation. The union of all token results is saved as $PossibleRel_i$ for concept relation evaluation (lines 6-8). Each pair of concepts, Con_i and Con_j, is analyzed for whether the token descriptors are contained in one another. If yes, a subclass relation is defined. Otherwise the concept relation can be defined by the intersection of the possible relation descriptors, $PossibleRel_i$and $PossibleRel_j$, and is named according to all the descriptors in the intersection (lines 9-13).

V. IMPLEMENTATIONS

Here we propose to bootstrap the Ontology of each patient with the help of a knowledge base. The knowledge base gives the relation between disorders, signs and symptoms, how the disease was contracted, pathogen responsible, treatment approach and prevention methods. The interns have to identify the primary signs and symptoms and the system will guide through the possibilities of disorders. Hence the interns can do the labs only for those disorders. This helps to reduce the time to identify the case and prepares reports appropriately.

Bootstrapping requires analyzing the web service using 3 parameters:

- TF / IDF - Term Frequency / Inverse Document Frequency
- Web Context Extraction Method
- Free Text Description Verification Method

Similarly, Bootstrapping in Medical Case Studies can be done using 3 parameters:

- Analyze the primary and secondary symptoms
- Analyze the contact of the symptom and thus the disorder
- Detect pathogen present using lab results

When the primary and the secondary symptoms are analyzed, the possible disorders will be suggested. This can be refined further by analyzing the contact of the symptoms or the allergy. Thus the labs to be done can be chosen accordingly. That is unnecessary lab tests can be avoided. Hence leading to monetary benefits and saving time. Finally based on the symptoms, contact and pathogen, the disorder is diagnosed. This is updated to the case (patient's) profile. Also, the medication, side effects and vaccine (if available) is automatically added to the case study. Thus the ontology of the medical case (patient) is bootstrapped reducing human intervention as much as possible.

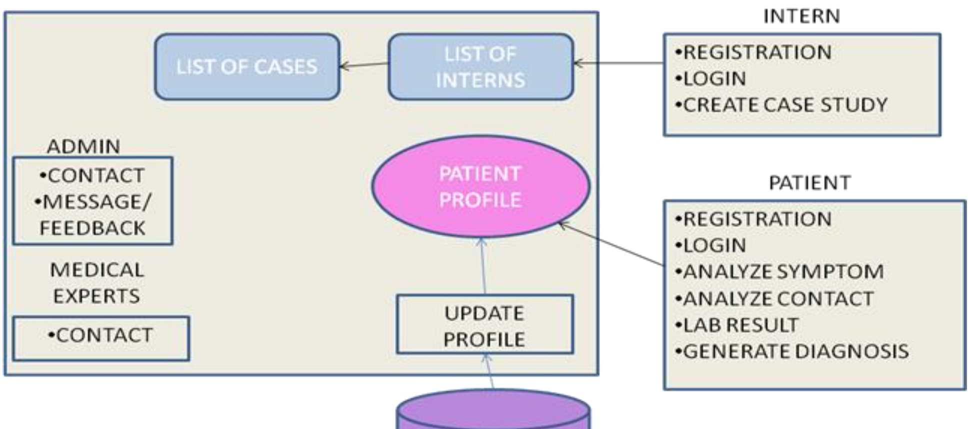

Fig. 3.Bootstrapping Ontologyfor Medical Case Study.

A. INTERNS:-

The interns are provided with a profile when registered. A usual login is used to keep their profile discreet. They can maintain a list of their case studies. When the intern registers a new case, the basic information is first stored in the patient details relation. The first step is to analyze the primary and secondary symptoms. The list of possible disorders is suggested for the intern. In order to refine this list, the contact of the symptoms or the allergy is analyzed. Now the labs can be done, based on the refined list. This reduces time and unnecessary labs. Based on the 3 parameters, the disorder is diagnosed and it is updated in the ontology of the patient. Also the medication, side effects and vaccine are automatically added to the patient ontology from the knowledge base.

B. MEDICAL EXPERTS:-

The Interns can contact the Medical Experts for help in the cases. The Medical Experts are the Attendings or the Residents in top hospitals. They may be experts in various fields of medicine. Thus the extremely rare cases can be studied with better understanding. This is provided via e-mail or phone.

C. ADMIN:-

The Interns can contact the Admin for feedback and suggestions. This helps to improve the system for the benefit of the Interns. Thus this system respects innovation from every perspective.

D. PATIENTS:-

These are the medical case study patients, the Interns are interested in. They are initially analyzed for their primary and secondary symptoms and then the contact of the symptom or the allergy. They are taken for the required labs. Based on this, the disorder is diagnosed. The patient profile is automatically built based on the diagnosis. Hence the complete report of the case will consist of the primary and secondary symptoms, their contact of the symptoms or allergy, their lab results, diagnosed disorder, medication, side effects and vaccine. Hence Bootstrapping Technique for Ontology Construction and Evolution helps to reduce human intervention as much as possible.

VI. CONCLUSION

The Existing System is heavily based on the existing ontology or is domain specific but the Proposed System automatically creates the ontology and evolves them automatically from the beginning. Implementing the proposed system in the Semantic Web would reduce human intervention and thus makes maintenance easier. This Bootstrapping technique can be applied to other application domains like: Analysis of Law Documents and Medical Case Studies. Here we propose to bootstrap the Ontology of each patient with the help of a knowledge base. The knowledge base gives the relation between disorders, signs and symptoms, how the disease was contracted, pathogen responsible, treatment approach and prevention methods. The interns have to identify the primary signs and symptoms and the system will guide through the possibilities of disorders. Hence the interns can do the labs only for those disorders. This helps to reduce the time to identify the case and prepares reports appropriately. This result can be used as an evidence for Dempster-Shafer Theory (DST). We are aware of the fact that DST is not very successful in finalizing a disorder when there is a conflict in diagnosis. The Dempster-Shafer Theory (DST) is a mathematical theory of evidence. It allows one to combine evidence from different sources and arrive at a degree of belief that takes into account all available evidence. Additionally, this Bootstrapping Technique can be utilized for constructing and maintaining Ontology in other Domains.

REFERENCES

[1] Aviv Segev and Quan Z. Sheng, Bootstrapping Ontologies for Web Services, IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 5, NO. 1, 2012.

[2] D. Kim, S. Lee, J. Shim, J. Chun, Z. Lee, and H. Park, “PracticalOntology Systems for Enterprise Application,” Proc. 10th AsianComputing Science Conf. (ASIAN ’05), 2005.

[3] M. Ehrig, S. Staab, and Y. Sure, “Bootstrapping OntologyAlignment Methods with APFEL,” Proc. Fourth Int’l Semantic Web Conf. (ISWC ’05), 2005.

[4] G. Zhang, A. Troy, and K. Bourgoin, “Bootstrapping OntologyLearning for Information Retrieval Using Formal ConceptAnalysis and Information Anchors,” Proc. 14th Int’l Conf.Conceptual Structures (ICCS ’06), 2006.

[5] S. Castano, S. Espinosa, A. Ferrara, V. Karkaletsis, A. Kaya, S.Melzer, R. Moller, S. Montanelli, and G. Petasis, OntologyDynamics with Multimedia Information: The BOEMIE Evolution Methodology,” Proc. Int’l Workshop Ontology Dynamics (IWOD ’07),held with the Fourth European Semantic Web Conf. (ESWC ’07), 2007.

[6] L. Ding, T. Finin, A. Joshi, R. Pan, R. Cost, Y. Peng, P. Reddivari,V. Doshi, and J. Sachs, “Swoogle: A Search and Metadata Enginefor the Semantic Web,” Proc. 13th ACM Conf. Information andKnowledge Management (CIKM ’04), 2004.

[7] A. Patil, S. Oundhakar, A. Sheth, and K. Verma, “METEOR-S WebService Annotation Framework,” Proc. 13th Int’l World Wide WebConf. (WWW ’04), 2004.

[8] Y. Chabeb, S. Tata, and D. Belad, “Toward an Integrated Ontologyfor Web Services,” Proc. Fourth Int’l Conf. Internet and WebApplications and Services (ICIW ’09), 2009.

[9] Z. Duo, J. Li, and X. Bin, “Web Service Annotation Using Ontology Mapping,” Proc. IEEE Int’l Workshop Service-Oriented System Eng. (SOSE ’05), 2005.

