MODELLING ONTOLOGICAL AGENTS WITH GAIA METHODOLOGY

María Auxilio Medina N.
Universidad Tecnológica de la Mixteca, Huajuapan de León, Oaxaca, México

Alfredo Sánchez, Ma. Lourdes Fernández
Universidad De Las Américas -Puebla,
Cholula, Puebla, México

Keywords: Ontologies, digital libraries, agents

Abstract: Multi-agent systems have been successfully applied in information retrieval tasks, especially in environments whose sources of information are distributed and highly heterogeneous. They can be perceived as an alternative to face problems that traditional search engines are not able to solve yet. On the other hand, ontologies have shown their efficiency to management different sources of information. We present the model of some software agents that use ontologies to improve information retrieval tasks in a set of federated digital libraries. Gaia methodology is used for this purpose and the paper highlights some of its main advantages. It also shows that this methodology can be easily used in similar environments to avoid ad hoc construction of agent-based systems.

1 INTRODUCTION

The use of the World Wide Web as a space widely used to publish documents and the expansion of Internet are relevant factors that have contributed to the increase of sources of information. In this document we are interested in digital libraries mainly because they are invaluable repositories of reliable and structured information.

In order to retrieve relevant resources from several digital libraries, users require knowing their allocation as well as features of their interfaces to express their information requests appropriately. Once search mechanisms of each digital library are activated, users are in charged of collecting the results of each source and afterwards they choose a set of sources of information. Although a wide range of technological tools have supported digital libraries operation, nowadays this task is a very time consuming one and it is often that it is delegated to software or human agents.

This problem is not new and it has been analyzed in areas such as distributed systems, interoperability systems, federated information systems and multi-agent systems. Due to the problem has to deal with different levels of semantic heterogeneity; ontologies emerge naturally as appropriate tools [Pomerantz & Silverstein 2001].

This paper proposes the use of ontologies by a agent-based system to retrieve information from a set of federated digital libraries. Ontologies are used to support the disambiguation of keywords in user queries as well as to describe the repositories or sources of information. In order to avoid ad hoc development of agent-based software, the design of these agents is based on Gaia methodology.

The paper is organized as follows: Section 2 briefly describes previous work with the purpose of establish the antecedents of ontological agents. Section 3 presents the design of ontological agents using Gaia methodology. Related work is presented in Section 4. Conclusions and future works are presented in Section 5.

Auxilio Medina N. M., Sánchez A. and Lourdes Fernández M. (2004).
MODELLING ONTOLOGICAL AGENTS WITH GAIA METHODOLOGY.
In *Proceedings of the Sixth International Conference on Enterprise Information Systems*, pages 556-560
DOI: 10.5220/0002645205560560

2 PREVIOUS WORK

In order to understand the design of ontological agents, it is necessary to know the context in which these agents are going to be integrated. This is the Virtual Reference System [Sánchez et al. 2001]. In the further, the term Vref is used to make reference to this system.

VRef is a virtual space in which users post their information needs or queries expressed in natural language and a staff of reference librarians as well as a set of software agents called *reference agents* are in charged of searching relevant digital or physical resources from the UDLAP[1] digital library. A detailed description of this kind of agents can be found in [Medina et al. 2002] and [Medina et al. 2003]. One of the most relevant characteristics of Vref is that it has a knowledge base that is maintained and enriched with the successful interactions of users. The use of this knowledge base allows a reference librarian make use of the work generated by other reference librarians or by reference agents. Likewise, a reference agent can also have access with the same objective. It is also in charged of implement a mechanism to associate a similarity measure between queries.

Vref has been operating for almost a year and a half. At present, we are considering to use it as the interface to have access to a set of federated digital libraries. In this paper, we took the definition of a federated system proposed by [Busse et al 1999], this is a kind of system which offers read and write operations to data management such as any traditional database system. It is commonly built through an integration technique of schema. It is formed by structured components that can be accessed by query languages.

The particular federation of digital libraries we are thinking in makes use of the Open Archives Initiative Protocol for Metadata Harvesting, (abbreviated OAI-PMH). It can be perceived as a basis for supporting interoperability in digital libraries. The OAI-PMH protocol includes a model consisting of two main parts: data providers, which expose metadata, and service providers, which harvest and process data automatically. Although this protocol is relatively new, (less than four years old), it has been incorporated into the development of many important research projects. [Van de Sompel & Lagoze 2001] attribute its growing popularity to the availability of tools to build OAI-PMH repositories and harvesters.

The proposal of retrieving information from a set of digital libraries federated by the OAI-PMH protocol is summarized in the next general algorithm:

1. A reference agent receives the query and gets the keywords.

2. An ontological agent uses a general ontology of natural language to present to the user some alternatives of new words or referred words in order to eliminate as much ambiguity as possible from the original keywords.

3. Once a user chose the appropriate alternative, the ontological agent identifies a set of digital libraries which according with a set of well defined ontologies potentially have relevant resources to satisfy the query.

4. The ontological agent sends the query to a mobile agent, which is in charged of visiting the sources of information and of the retrieval process of relevant resources. A mobile agent is used for each digital library identified at step 3.

5. The partial results are sent to the reference agent and they are presented to the user in a transparent way at the Vref interface.

3 DESIGN OF ONTOLOGICAL AGENTS BASED ON GAIA METHODOLOGY

Gaia is a top-down methodology for agent-oriented analysis and design. It is based on the concept of roles. A multi-agent system is understood as a computational organization formed by the interaction of various roles. Gaia methodology takes into account the macro-level and the micro-level aspects of agent-based systems [Wooldridge et al. 2000].

At the requirements statement, the identification of requirements enables to model and to specify in a conceptual level of detail ontological agents. It

[1] UDLAP is the short name of Universidad De Las Amèricas Puebla

describes the semantics of a system without concern with implementation details. We adopted the same requirements of a federated system: scalability and adaptability. Analysis stage can be summarized in the next following tasks: identification of roles, identification and documentation of protocols and elaboration of the roles model. The roles model is presented by the next set of tables.

Table 1: Role to process a query

Description	It is in charged of getting the keywords from a query expressed in natural language.
Protocols and activities:	EliminateStopWords
Permissions:	Read a query
Responsibilities:	
Liveness:	keywords were extracted from the query
Safety:	suggest new words just when ambiguity is detected

Table 2: Role to formulate a query

Description:	This role involves ensuring to eliminate the much ambiguity as possible from the keywords of the queries.
Protocols and activities:	ShowNewWords, FormANewQuery
Permissions:	suggest other words
Responsibilities:	
Liveness:	keywords were extracted from the query
Safety:	suggest new words just when ambiguity is detected

Table 3: Role to describe the sources of information

Description:	This role is related with the description of a source of information.
Protocols and activities:	RepresentSchema
Permissions:	represent a source of information
Responsibilities:	
Liveness:	the source can be described
Safety:	have a representation of a source of information

Table 4: Role to visit different sources of information

Description:	This involves to have access to a source of information
Protocols and activities:	RetrieveResources, SendResources
Permissions:	search sources
Responsibilities:	
Liveness:	there is a set of sources
Safety:	have access to a source

Table 5: Role to collect results

Description:	This role has the objective of collecting the results of each source of information
Protocols and activities:	JoinResources, FiltrateResources
Permissions:	AcceptsResults
Responsibilities:	
Liveness:	there are available resources
Safety:	establish a minimum set of resources

We have identified four protocol definitions: KeyWordsAgree, IdentificationRelevanteResource, RetrieveResources, PresentResources. They are represented in the templates of Figure 3. In this templates, the top cell has the name of the role, middle cells are used for initiator and responder. The bottom cell has processing attributed. Inputs and outputs are briefly described on the right from top to bottom respectively.

From the protocols presented above represent the services identified by each type of agent. It is worth to mention that was during this stage that we have decided to define two types of ontological agents we have termed: NLP-Agent and SourceDescriptor Agent. The service associated to the first type of ontological agent is related with the purpose of eliminating as much ambiguity as possible from keywords of user query; instead, the second type of ontological agent has the function of representing the description of a source of information. The direct graph of Figure 2 represents the acquaintance model for the agent types.

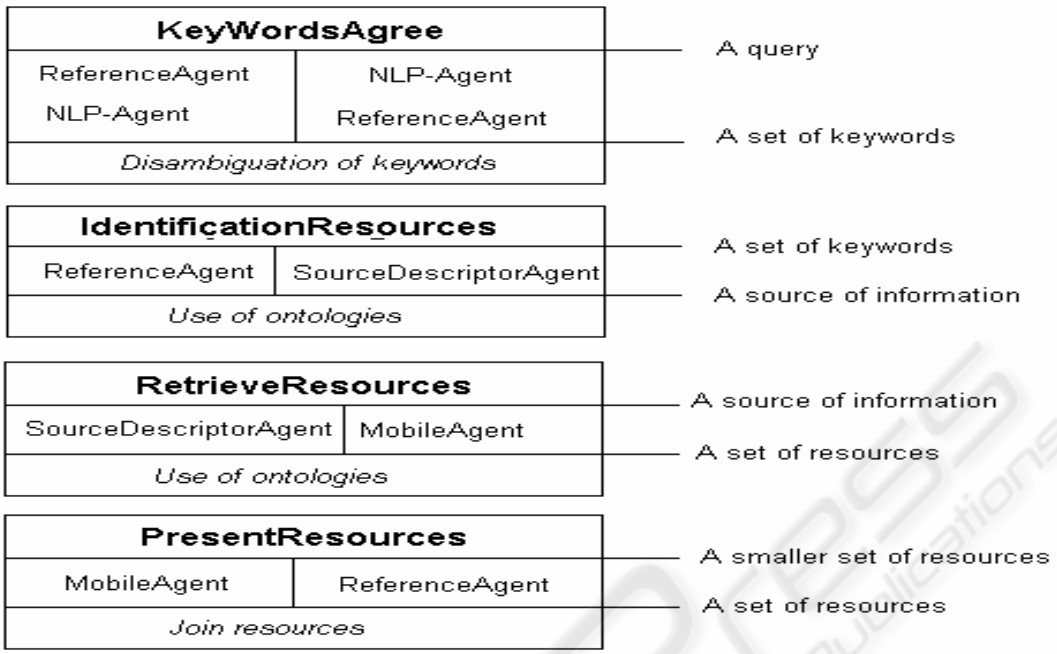

Figure 1: The interactions model

4 RELATED WORK

The problem of having access to heterogeneous sources of information has been analyzed from diverse points of view. This section includes some of the most relevant ones are briefly described in this section.

FIPA (Foundation for Intelligent Physical Agents) fosters agent-based applications. It provides specifications to support interoperability between agents and a special specification to manage ontologies provided by an Ontology Agent (abbreviated as OA). Some of the services of an OA are: searching and accessing public ontologies, maintaining a set of ontologies and a translation mechanism between ontologies. It is able to answer queries about relationships between terms. At FIPA, it is not mandated that every OA will be able to offer these services; the implementation of these services is left to developers [FIPA 2001].

[Saavedra 2003] proposed a method to federate a set of documental databases based on ontologies. It accomplishes with the requirements of scalability, adaptability. Ontologies are used to conciliate schema of the databases and also to build a friendly user interface at run time. The concepts are represented in a structure called *concept trees*. Users select these trees when he or she introduces a query through the interface. Due to XML files are used to guide software execution; any change in the databases does not require recompilation of any module of the system, providing physic and logic independence of databases.

The semantic web project is focused on providing intelligent access to heterogeneous and distributed information. In this project, agents operate as mediators between user needs and available information resources. Ontologies play a key role in this project [Fensel 2000], [Fensel 2001].

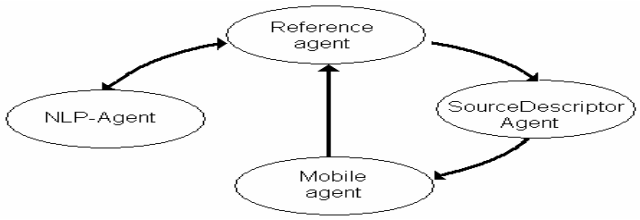

Figure 2: Representation of aquaintance model

5 CONCLUSIONS

In this paper, we have briefly described Gaia methodology and we have presented its application to design a set of agents to retrieve information of a set of federated digital libraries. This design has been proposed as an alternative to have access to these heterogeneous and distributed sources of information. All the stages and the different of models were included.

Gaia methodology was chosen due to its adaptability characteristics in such a way that can be applied to a wide range of agent-based systems. The design of the agents presented in this paper with this methodology provides us of a detailed description of the system in such a way that the different kinds of agents can be implemented easily; likewise, it can be perceived as a reliable document to support system maintenance. According to the requirements statement, this design can be a base for the stages of validation, verification and test.

REFERENCES

Louis, R., 1999. Software agents activities. In *ICEIS'99, 1st International Conference on Enterprise Information Systems*. ICEIS Press.

Chandrasekaran, B.; Josephson, R. What are ontologies, and why do we need them? In IEEE Inteligent systems, 1999.

Busse S., Kutsche R., Lesser U., Weber H. 1999. Federated information systems: concepts, terminology and architectures. Technical Report 99-9. TU Berlin. April.

Farquhar, A., Fikes, R., Pratt, W., and Rice, J. (1995). Collaborative ontology construction for information integration. Technical report, Stanford University.

Fensel D., 2000. The semantic web and its languages. *IEEE Computer Society 15,* 6 (November /December), 67-73

Fensel D., 2001. *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce.* Springer.FIPA 2001. Foundation for Intelligent Physical Agents.FIPA Ontology Service Specification. March. Available on: http://www.fipa.org/specs/fipa00086/XC00086D.html

Luck M., McBurney P, and Preist C. 2003. Agent technology: Enabling next generation computing: A roadmap for agent-based computing. AgentLink report, available from www.agentlink.org/roadmap.

Medina, M. A., Sánchez, J. A. 2002. Agents at the reference desk: Serving information needs and constructing knowledge for wide communities of users. Memorias del XII Congreso Internacional de Electrónica, Comunicaciones y Computadoras (Conielecomp 2002, Acapulco, México, Feb.), 74-78.

Medina Ma. Auxilio, Chávez-Aragón A. Inclusion of Reference Agents in a Virtual Reference Environment, ENC 2003, Apizaco Tlaxcala, September 2003.

Pomerantz, J., & Silverstein, J. (2001, 13 November). The Importance of Ontology: Standards for a Referral Tool. Paper presented at the 3rd Annual VRD Conference: Setting Standards and Making it Real, Orlando, FL. Available on: http://www.vrd.org/conferences/VRD2001/proceedings/silverstein.shtml.

Saavedra A. 2003. Arquitectura para federación de bases de datos documentales basada en ontologías. Disertación doctoral. Universidade Da Coruña. Departamento de Computación. Enero.

Sánchez, J. A., García, A. J., Proal, C., Fernández, L. 2001. Enabling the collaborative construction and reuse of knowledge through a virtual reference environment. Proceedings of the Seventh International Workshop on Groupware, (Darmstadt, Germany, Sept. 6-8). IEEE Computer Society Press, Los Alamitos, California. 90-97.

Van de Sompel H., Lagoze C. Notes from the interoperability front: a progress report on the open archives initiative. Proceedings of the 6th European Conference on Research and Advance Technology for Digital Libraries (ECDL 2002, Rome, Italy, September), 144-157.

Wooldridge, M., Jennings N. Kinny D. 2000. The Gaia Methodology for Agent-Oriented Analysis and Design. Autonomous Agents and Multi-Agent Systems, 3, 285-312. Kluwer Academic Publishers. The Netherlands.