Intelligent Agents: a Tool for Managing and Retrieving Information in Distributed Systems

F. Bergadano¹ A. Puliafito² S. Riccobene² G. Ruffo¹ L. Vita³

¹ Dipartimento di Informatica
Università di Torino
Torino, ITALY 10149

² Istituto di Informatica
Università di Catania
Catania, ITALY 95025

³ Facoltà di Ingegneria
Università di Messina
Messina, ITALY 98166

Abstract

Intelligent agents represent a powerful tool to implement distributed functionalities into a network environment. This paper describes the current status of a research project whose target consists in using this new technology for automatic information retrieval, network management and security purposes within a distributed system. Specifically, we describe the software agents dedicated to information retrieval, focusing on architectural aspects and implementation issues.

1 Introduction

An Intelligent Agent is a software module equipped with Artificial Intelligence mechanisms which is capable, autonomously or in cooperation with other agents, of solving a certain problem or carrying out a certain task. The more an agent is autonomous, user-friendly, capable of dealing with unexpected situations, of learning and improving its performance on the basis of its mistakes, or observations or instructions from the outside, the more intelligent it is considered to be.

Literature provides several examples of applications of the agent programming paradigm: it has been used in the field of intelligent navigation within the Web (Internet), at the University of Washington [4, 5], in order to produce access mechanisms to the information using natural language (MIT) and in order to optimize management systems of satellite images [7].

This paper is based on research activity currently in progress at the Universities of Catania, Messina and Turin. Within a distributed network system, the research proposes to specify and implement software agents dedicated to:

- optimizing and rationalizing the use of shared resources (management);
- maintaining the security of the system at an adequate level;
- providing support for intelligent navigation within the system.

In this paper we describe the functional architecture and implementation of the module dedicated to the information retrieval task.

2 Information Retrieval and Intelligent Agents

The techniques currently used to facilitate retrieval operations, in particular finding any information in a reasonable period of time, are based on the use of index files ([6]). It has been observed in [5] that indexing agents can deliver quick responses, but they have a number of technological limitations.

The ideal situation is one in which a generic user gives a high-level description of the information she requires and the system, on the basis of heuristic techniques and experience acquired, automatically retrieves this information: the whole network can thus be considered as a virtual archive. It is the authors’ opinion that the paradigm of agent programming can provide an efficient support for this kind of approach and for its automation.

We envisage our world divided into subnetworks or domains, each of which comprises web nodes, an agent we call a Broker, and possibly other nodes which are not WWW servers but where the human user can be located.

The following agents are used, in order to obtain the goal described above:

- User Agent (UA);
- Information Retrieval Agent (COLOMBO);
- Database Agent (DBA);
- Broker Agent.
2.1 Agent Description

In each node in our domains there is always an agent called a User Agent (UA) which interfaces with the user, "translating" her requests as appropriate; it activates the search process which returns a list of URL addresses at which the information being sought can presumably be found.

A central node in the retrieval process is COLOMBO which, when duly "instructed" by the User Agent, can locate the URL addresses which have to be returned to the user, possibly "migrating" from one Web node to another, according to the advice given by the Broker, and interacting with the local DataBase Agents (DBAs) in each Web site. A Broker Agent can be seen as a name server, that can be asked for the Web site addresses where the information required by the user is most likely to be found; the Broker is also capable of "advising" COLOMBO to consult brokers in other domains in order to meet the user’s request.

3 Learning Mechanisms

We implemented C-procedures able to index each file of a Web node belonging to a domain, with a series of attributes and weights; they can be used in order to characterize local knowledge bases (which are written in C-Prolog). Actually these procedures use a well established technique in information retrieval: the term-frequency inverse-document-frequency calculation [8]. The knowledge bases of COLOMBO and the other agents can be dynamically modified whenever a new query is satisfied or whenever a user notifies the UA that she is not satisfied with the response obtained. Classic Machine Learning methods of reward/punishment are able to do this, especially in a context where an attribute-weight representation is used.

Another way to do this is to use ILP (Inductive Logic Programming [2, 3]) techniques. By definition, ILP makes it possible to learn logic programs with the aid of examples. We may observe the user's reaction to the system's suggestions in order to produce "examples". From these examples we can derive more general rules written in a logic programming language, as Prolog. These rules will substitute or complete the knowledge bases, which are written in Prolog.

4 Implementation Notes and Conclusion

The agents' system we have just described has been partially developed and is currently being implemented inside a joint cooperation between the University of Catania, Messina and Turin. We used 3 Sun SparcStations with Solaris 2.4 as the operating system, but it is possible to install and compile the whole system in any machine supporting Unix. The knowledge bases of the agents and the user's query are translated in C-Prolog 1.5 plus; in particular it must be noted that the migrating code of COLOMBO is also written in this language. In order to implement migration and inter-agents communication procedures, it has been used C and KAPI 2.6e tool; the whole communication system is to be implemented in Java. Note that the standard language for the exchange of messages between agents that the KAPI tool implements is KQML (Knowledge Query Manipulation Language).

An information retrieval system based on intelligent agent technology has been described which is part of a wider project currently under development. The system architecture is formerly based upon the activity of COLOMBO, which is able to "migrate" in remote nodes in order to retrieve the information needed. Machine learning techniques based on Inductive Programming Languages are being implemented to increase the agent capabilities to learn from the experience of previously satisfied queries.

References


