

WIC France Research Center Activity Report

Pierre Morizet-Mahoudeaux
University of Compiègne
UMR-CNRS 6599 Heudiasyc
B.P.20529, 60206 Compiègne CEDEX, France
pmorizet@hds.utc.fr

1. Introduction

The research activity of the Heudiasyc laboratory in the domains related to WIC concerns four projects : (i) intelligent agents, (ii) e-learning, topic maps, and semantic Web, (iii) Web topology, and (iv) digital music preservation.

2. Intelligent agents [3]

We are interested in the architecture of complex systems that integrate automatic modules and human agents. The selected domain concerns distributed artificial intelligence, and more precisely, that of cognitive agents. We address software issues and human factors (cognitive sciences, and communications). Test applications include (i) Cooperative Work in Design and Concurrent Engineering, and (ii) Management of Scientific and Technical Knowledge.

Previous research have shown that blackboard architectures were too limited to be useful in areas like concurrent engineering or for the management of corporate knowledge. Therefore, it is more efficient to develop systems of cognitive agents for the development of large distributed systems. In such cases, problems like size factors, distribution, integration of the human actors, or continuous functioning during several years, have not received satisfactory answers. Our research thus presents two aspects: (i) experimental work, and (ii) theoretical work. Our experimental work is intended to acquire first hand experience about systems of cognitive agents, in the framework of actual large complex applications. Our OMAS experimental multi-agent platform allows us to test the various mechanisms proposed by other research teams (e.g., communication protocols, models of interaction and cooperation, agent knowledge, transfer of knowledge among agents). We developed two new platforms: SMAS for simulating complex Multi-Agent Systems and OMAS for a real time execution. The designer of the MAS first debugs her system in the SMAS environment, and then executes it in the real time distributed OMAS environment.

Our theoretical work aims at designing a methodology, allowing a casual user to build systems of cognitive agents easily, while relying on a sound theoretical model. Among the recent theoretical result, we developed a three-level model of multi-agent systems and proposed the concept of coterie. We also introduced the concept of MASH (Multi-Agent Systems with Humans) as an environment linking various human persons and including service agents.

We also worked on the concept of personal assistant agent. Its role is to translate human requests into messages that can be understood by the other agents of the system. We have defined a conceptual model of a personal assistant, and demonstrated how a personal assistant could be used to assist the user through syntax-free dialogs. The concept was further developed, to improve the dialog interface, to introduce automatic learning of the user's profile, and to propose to replace the personal assistant with a local MAS. We are now interested in replacing the keyboard interaction with a vocal interaction.

3. The Memora project [1]

Building and managing a training or a course unit, in a university or in a firm, are often complex and expensive operations, which need the production of multiple documents and the cooperation of numerous persons. Knowledge and documents are not usually capitalised during these operations. This leads to a waste of effectiveness for teachers who are often obliged to recreate documents, although they may already exist. This also leads to a waste of richness for courses, because apart from course notes or textbooks, other documents can be useful for students. Moreover students have often difficulties to identify and to access resources that could be useful.

Our goal within the MEMORAe project is to examine to what extent knowledge engineering could tackle this problem. Seeing a course as an organization the knowledge of which has to be managed, we propose the concept of "learning organizational memory". This memory allows,

on one hand, to capitalize the learning knowledge, and on the other hand, to better index resources, taking into account the learning context. This concept is particularly adapted to e-learning. In this kind of situation, the utility of a shared memory is reinforced since the interactions do not usually take place in the same place and at the same time. This research work is thus situated at the crossroad of three domains: knowledge engineering, pedagogical design and semantic web. The setting of what to learn and of the links between the learned notions relates to pedagogical design. The choices of organization, the management of resources in an ontology-based learning organizational memory concerns knowledge engineering. Finally, the choice of the ISO Topic Maps standard for representing the memory and allowing its consultation is connected with the semantic web domain. The MEMORAE project (that stands in french language for ORganizational MEMory for E-learning) is supported by the STEF program of the Picardie district.

4. Web topology [4]

How to extract and visually explore the topology of an open, large scale, hypertext system such as the web ? We address this issue by developing an experimental tool for extracting, exploring and analyzing aggregates of web documents. This tool, called TARENTE, includes a crawling technology, and algorithms for both content analysis and authority graphs calculations, linked with visualization solutions. We provide series of experimental results on different topics that allow us to describe the Web's structure in terms of topic aggregates.

Our goal is to validate some assumptions on the aggregate topology of the Web. Some general, theoretic principles have already been explored the most important of which is Kleinberg's assumption that the structure of broad topics on the web are aggregates organized around a core of hubs and authorities. Our objective is to test these principles by describing the aggregates at the local level: their volume (or the scale of their resource), their inner structures, and their borders (aggregates neighborhoods). Graph theory, data mining, crawling techniques and other methods are at stake to discuss these theories by correlating hypertext connectivity and topical content. The experimental aspect of this work consists in developing a flexible software (TARENTE) which is able to connect statistical analysis and visualization modules to produce visual as well as statistical evidences. This work should give hints to formalize a general theory of aggregates (especially mathematically), including different kinds of topological models. In particular, those models could help in qualifying web communities. In further work, we will study the optimization of larger crawls including time-dynamic graph to observe emergent patterns.

5 The MUSTICA project [2]

MUSTICA, an international team of archivists, musicologists, and communication and information scientists formed in the spring of 2003 to research issues of digital music preservation while drawing on resources provided by the fields of information technology and social sciences. MUSTICA's research activities will operate as a case study within the framework of InterPARES 2, a large interdisciplinary, international research group based out of the University of British Columbia that is concerned with the preservation of authentic interactive digital records of artistic, scientific, and government activities. The goal of InterPARES 2 is to provide two series of preservation guidelines, one for the creators of records and one for archivists, that are adaptable to a variety of legal, social, and cultural norms (see <http://www.interpares.org>). MUSTICA will assist in the drafting of these guidelines by collaborating with three French organizations, the *Groupe de Recherches Musicales* (the GRM), a research center within the *Institut National de l'Audiovisuel of France*, the *Institut de Recherche et Coordination Acoustique* (IRCAM), and the UMR-CNRS 6599 Heudiasyc laboratory of the University of Compiègne.

MUSTICA's objectives are based on two principles. A technological principle: the musical content of a music document is an inscription on a physical carrier. In the context of contemporary music composition using computers, this carrier is generally in digital format. Its inscriptions will only be accessible through the use of technological mediation to preserve and reconstitute the different formats and materials involved in its creation. A musicological principle: the musical content of a music document is an object requiring comprehension and interpretation in the context of the conditions of its production, reception, and interpretation.

References

- [1] M. Abel, C. Barry, A. Benayache, B. Chaput, D. Lenne, and C. Moulin. Ontology-based organizational memory for e-learning. *Educational Technology & Society*, in print, 2004.
- [2] B. Bachimont, J.-F. Blanchette, A. S. Andrew Gerzso, O. Lescurieux, P. Morizet-Mahoudeaux, N. Donin, and J. Teasley. Preserving interactive digital music: A report on the mustica research initiative. In *Third International Conference on WEB delivering of Music (WEDELMUSIC'03)*, Leeds, United Kingdom, September 15 - 17 2003.
- [3] F. Enembreck and J.-P. Barthès. A new approach for learning agents. *Journal of Autonomous Agents and Multi-Agent Systems*, in print, 2004.
- [4] F. Ghitalla, C. Maussang, E. Diemert, and F. Pfaender. Tarante: an experimental tool for extracting and exploring web aggregates. In *IEEE-ICTTA04*, Damas, Syria, April 2004.