Web Intelligence in Mexico

Jesús Favela
CICESE, Mexico
favela@cicese.mx

Manuel Montes
INAOE, Mexico
mmontesg@inaoep.mx

Edgar Chávez
UNMSH, Mexico
elchavez@fismat.umich.mx

Abstract

The Mexico Research Centre of the Web Intelligence Consortium was established in 2003 motivated by Mexico being selected as the host of the 2nd Atlantic Web Intelligence Conference. It currently has 18 members including faculty and doctoral students from 7 different institutions. The WIC-Mexico includes groups working in the areas of Intelligent Web Information Retrieval, Web Mining and Farming, Knowledge Management, and Agents in Ubiquitous Computing.

1. Introduction

Web Intelligence is a growing field of research in Mexico. Although research in related areas has been done for a number of years, it was not until 2003 that the WIC-Mexico was formed motivated by the selection of Mexico as host of the 2nd Atlantic Web Intelligence Conference (AWIC'04) held in Cancun in May of 2004. This event has helped to increase the awareness of the different efforts being made in this area and should create a fertile ground to initiate more ambitious projects within this community and collaborate with other members of the WIC around the world.

In this paper we describe some of the work being done in the field of Web Intelligence in Mexico and opportunities we envision for further collaboration. Additional information can be found the Center's Web site: http://ccc.inaoep.mx/wicmx/index.html

2. Research Projects in Web Intelligence

Web Intelligence research in Mexico centers around the following areas: Intelligent Web Information Retrieval, Web Mining and Farming, Knowledge Management, and Agents in Ubiquitous Computing.

2.3 Intelligent Web Information Retrieval

Perhaps the most active area of WI research in Mexico relates to information retrieval. Several projects are being pursued in this area:

Context aware news portal with non-intrusive user profiling. We collect news from newspapers, news portals, etc. We use user profiling based on clickstreams under the holomorphic distances [1]. The news portal changes the locus of news articles depending on the user profile.

Agent-based contextual retrieval. We are working in the contextual retrieval of medical information within ubiquitous computing environments. Links from previous patient's records and information from medical libraries relevant to the case at hand are offered to the user through implicit queries gathered from contextual information [4].

Question answering on the Web. We are working in two different approaches. On the one hand, an statistical QA system supported on Web redundancy rather than on spphisticated linguistic analyses of either questions and documents [6]. On the other hand, a QA model based on the representation of the docuements by a set of instances of concepts from a top level ontology expressed in DAML+OIL [7].

2.4 Web Mining and Farming

Collaborative crawling for scientific information gathering. We provide a set of Web tools (context management tools) and agents to maintain Web pages with individual curriculum vitae of researchers. A number of DTD's are defined to save separately data and content. Each scientific product is maintained as a separate record with syndication tags. A central server harvests all the information of individual pages and feeds a citeseer-like index. We are working on extracting the curriculum vitae ontology from pure text. We want to work on similarity measures between scientific papers.

Using the Web for word sense disambiguation. We propose to use the Web as a virtual corpus. Currently we visualizate two main uses: (i) learn a language model

from the Web, and (ii) extent the available tagged corpora collecting aditional training examples from the Web.

2.5 Knowledge Management

Adaptive resource management for Web-based collaboration. The management of shared resources Web-based co-authoring environments raises important technical questions due to the constraints imposed by Web technology. An elaborated group awareness function is being developed to allow each author to notify his contributions to other authors and control de way by which other contributions are integrated into his environment [2].

2.6 Agents in the Web

Ontology sharing for Web services. Agent technology is being used to facilitate the development of dynamic open service environments for Web based systems [5]. Ontologies constitute the centerpiece of the knowledge retrieval, sharing and reuse mechanisms used on the Web or among agents.

Agents for ubiquitous computing. Autonomous agents can enable spontaneous collaboration by representing users, as well as devices or services available through the Web, which has become an ubiquitous medium for information sharing. An agent acts on behalf of the user while he is disconnected, and represents services added to the environment, thus allowing the physical integration and interoperability of these entities. We have developed the SALSA framework [3], which allows developers to implement simple agents for ubicomp systems. These agents use an expressive communication language based on XML, which provides protocols for locating and interacting with Web services even when the user is disconnected.

Agents for document authoring. We are working in a multi-agent system that partially automates the authoring of Web documents. This system extracts descriptive information from a set of documents and constructs a set of meta-information descriptions based on the Dublin Core specifications [8].

3. AWIC 2004

WIC-Mexico was in charge of organizing the 2nd Atlantic Web Intelligence Conference (AWIC2004). AWIC provided a forum for specialists from different disciplines of computer science to exchange their experiences and ideas in this growing field of research.

The selection of papers followed a strict, double-blind refereeing process by a renowned international committee. We received 57 contributions from which 22 papers were selected to be presented and included in proceedings published by Springer in its Lecture Notes in Artificial Intelligence series (LNAI 3034).

In addition, we were pleased to have as invited speakers Dr. Prabhakar Raghavan, Chief Scientist from Verity, Inc. and consulting professor of Computer Science at Stanford University, and Prof. C. Lee Giles, Davide Reese Professor of the School of Information Sciences and Technology at The Pennsylvania State University.

4. Conclusions

Web Intelligence is a growing area of research in Mexico. Since the creation of WIC-Mexico and by hosting the AWIC Conference we have increased the level of awareness of the work being done in this area. This has ignited shared efforts not only between the groups working in Mexico, but gradually with other countries as well, mainly the U.S., Spain, France, and Chile. Through our participation in the Web Intelligence Consortium we expect to find new opportunities for collaboration.

References

[1] Chavez, E. "Knowledge-based Distances for Sequence Comparison". SEBD 2002: 16-24. Keynote. Isola de Elba, Italy, 2002.

[2] Mendoza, S., Decouchant, D., Martinez, A., Moran, A.L. "Adaptive Resource Management in the PIÑAS Web Cooperative Environment". In AWIC 2004, LNAI 3034, pp. 33-43, 2004

[3] Muñoz, M. Rodriguez, M., Favela, J., Gonzalez, V., and A. I. Martinez, "Context-aware mobile communication in hospitals," IEEE Computer, vol. 36, no. 8, pp. 60-67, 2003.

[4] Rodriguez, M. and Preciado, A., "An Agent Based System for the Contextual Retrieval of Medical Information". In AWIC 2004, LNAI 3034, pp. 64-73, 2004.

[5] Sheremetov, L., Contreras, M., Valencia, C. "Intelligent Multi-Agent Support for the Contingency Management System". International Journal of Expert Systems with Applications. 26(1). Pp. 57-71, 2004.

[6] Alejandro Del-Castillo, Manuel Montes, Luis Villaseñor. "QA on the Web: A Preliminary Study for Spanish Language". To appear in the proccedings of the ENC04, Mexico, 2004.

[7] M. Pérez-Coutiño, T. Solorio, M. Montes-y-Gómez, A. López-López and L. Villaseñor-Pineda. "Toward a Document Model for Question Answering Systems". In AWIC 2004, LNAI 3034, Springer, 2004.

[8] Manuel Pérez-Coutiño, Aurelio López-López and Manuel Montes-y-Gómez. "A Multi-Agent System for Web Document Authoring". In AWIC 2003, LNAI 2663, Springer, 2003.