Jornada de Seguimiento de Proyectos, 2010
Programa Nacional de Tecnologías Informticas

SOLERES — A Spatio-Temporal Environmental Management Information System based on Neural-Networks, Agents and Software Components TIN2007-61497

Luis Iribarne *
Departamento de Lenguajes y Computación
Universidad de Almeria

Abstract

This project includes and develops multidisciplinary applied research (e.g., software engineering, knowledge engineering, artificial intelligent, and ecology) for the modelling of a new *Environmental Management Information System* (EMIS) through an ontology-driven *agent* architecture perspective. This project aims to study the automatic generation of ecological maps from satellite images through *neural-networks*, useful for critical actions of environmental management in natural disasters (i.e. floods or earthquakes). Throughout this project, we are studying solutions on *component-based development* (e.g., *trading* services, or COTS composition) and *model-driven development* perspectives in EMIS.

Keywords: Components, Agents, Neural Networks, EMIS, Ontologies, CBD, MDD.

1 Project Goals

Some organizations of environmental scope (like EGMASA or URCI[1]) use ecological maps for their activities of environmental impact assessment. The elaboration of these ecological maps has high costs due to their preparation starting from fieldwork and periodical updating (years). Nevertheless, satellite images are faster to be acquired and give enough information to be interpreted. In recent years, the research team has been working on the design of an EMIS called SOLERES in which it is established a framework for the "correlation" and management of satellite maps and ecological cartography. EMIS is a special case of GIS (*Geographical Information System*) [21]. Our aim was to model an EMIS divided in two parts: on the one hand the platform that supports the correlation and information representation (KRS), and

*Email: luis.iribarne@ual.es

[1]EGMASA is a government company belonging to the Regional Andalusian Government. URCI is a private company working on environmental impact projects. Both companies take part as EPO in the project.

on the other hand, the platform that supports aspects of interface and user interaction (HCI). For correlation, we aimed to plan a three-phase methodology: (a) sectorization of cartography, (b) satellite image classification and (c) neural-networks training and design.

The specific objectives explained in the Project are briefly the following: (a) Studying the correspondence between the ecological maps information and satellite images information; (b) Developing a new language for specifying environmental maps (for instance, ecological maps obtained from satellite images); (c) Studying and demonstrating viability of neural-networks in the ecological field; (d) Defining a hierarchical trading model for searching processes (location and retrieval) in an EMIS; (e) Studying the intelligent user interfaces design to build cooperative systems; (f) Carrying out research in the field of *component-based development* (CBD) and *model-driven development* (MDD); (g) Carrying out research of neural-networks with software engineering techniques in CBD and MDD; (h) Studying service dynamic adjustment and composition algorithms in commercial components architectures (COTS) based on trading; (i) Approaching, developing and reinforcing techniques and practices in the computing disciplines of *software engineering* and *knowledge engineering*.

Figure 1 shows the chronogram with the scheduled tasks of the Project including the estimated time to be executed and level of completeness for each task. The nine objectives explained before are covered in six general tasks. Therefore, tasks A, B, C and D cover the objectives (a), (c) and (g); task E covers objectives (b), (d) and (f); task F, objectives (e) and (f); and task G, objectives (f) and (h). Objectives (f) and especially (i) are transversal to practically all tasks. To reach these objectives, we managed to gather a multidisciplinary human team of 16 people (8 ecology researchers and 8 engineers, the latter being specialists in Computing, Mathematics, Physics and Electronics): 2 foreign collaborators, one of whom is Full Professor at Delft University of Technology (The Netherlands) and another, Associate Professor at the National Technological University (Argentina); 4 Associate Professors at the UCM (*Universidad Complutense de Madrid*); and 7 Associate Professors and 3 doctoral researchers. The team has enough experience on neural-networks, satellite images, trading, CBD, MDD, agents and GIS. The tools to fulfil the tasks have been: (a) ERDAS and toolkit[2] for satellite; (b) Matlab for neural-networks; (c) and for components and agents, XMLSchemas, OWL, Protege, SPARQL, Jade, Visual Paradigm for Eclipse, ATL, COTStrader. We also have a server of three NVIDIA GPUs at 256 core precision, especially acquired for the project.

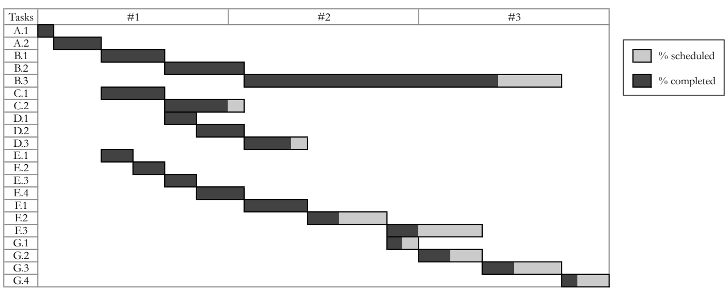

Figure 1: Task allocation: estimated time and level of completeness

[2]The Environmental Ministry of the Regional Andalusian Government, through EGMASA, provided us with a set of satellite images from the study area (Almeria and Granada).

2 Successful level of the project

This section is organised as follows. Firstly we will present some general considerations regarding difficulties found in the project. Then in different subsections we will explain in detail the project status arranged in four parts according to the four thematic groups in which the work team was organised, that is: (a) ecology, (b) satellite, (c) neural-networks, and (d) modelling.

New objectives have been arising throughout the Project which were not taken into account in the first planning; others have been reorganised in time.These changes are the following: (a) tasks linked to the analysis of new diffuse classifiers of satellite images (task C) established that this kind of classifier does not provide better solutions than traditional classifiers. We decided to start an alternative task to be able to develop a satellite image digital classifier based on *Cellular Automata*[3] and Software Agents. This new objective did not delay the remaining tasks because we used traditional classifiers to prepare data in the methodological process of correlation. (b) In the original proposal we planned the environmental information modelling through conventional XML templates. Nevertheless, we finally decided to use ontologies for further extensions of the system. This fact did not substantially delay the project tasks either; (c) for scientific purposes unscheduled in the planning, the acquisition of ecological variables prior to final sectorization was delayed in the methodology that the sub-team of ecology researchers used. Therefore, this fact has delayed other tasks especially associated with the neural-networks sub-team (task D, correlation process). Even so, these tasks have been fulfilled and results are in publication stages; (d) lastly, —and given that ontologies favour this— the modelling sub-team is interested in standardizing the environmental information system to work in Web environments. For this reason, the team has focussed on consolidating the investigations in order to bring together agent proposals, trading and ontologies in the Web.

2.1 Ecology

As previously advanced, the correlation methodology requires treatment on cartographic maps. This treatment consists of the development of a classification of ecological variables, which is called sectorization. Sectorization is an already developed methodological process that the ecological sub-team applied in the project to obtain ecological variables (necessary for the correlation process). This process has been carried out with ArcGIS.

Sectorization of the study area (Almeria-Granada) has given rise to five different maps: (a) Climatology. In order to elaborate these maps, we considered variables of evapotranspiration, temperature and rainfall; a total of 30 variables were considered. We dealt with 21,905 grids and obtained 4 climatic types in the study; (b) Litology. It is a subclassification of the previous map. We dealt with other variables for each climatic area (of the four obtained). There were 18 variables in total; (c) Geomorphology. We considered the altitude (11 variables), slope (6 variables) and orientation (8 variables). 25 variables in total; (d) Sectorization of (edaphic) soils. We took into account 32 variables for the analysis. From them, we obtained 6 types of dominant soils; (e) Sectorization of soils (vegetation). We considered 45 variables and 6 types of dominant vegetation. It was this layer of sectorization that we took into account for correlation with satellite images through neural-networks. All results of the sectorization process of the project (maps, tables and the whole material) are availabled at http://www.ual.es/acg/soleres/.

[3]Both foreign researchers have experience in Cellular Automata.

2.2 Satellite

The methodology also requires the treatment of satellite maps. We began from three images to be able to cover the whole study area (Almeria and Granada): a 7981x5195 pixel Landsat image and a spatial resolution of 30x30m. The satellite image had to be adapted to decrease the spatial resolution to the same size as the ecological map (1km2). We also had to calculate the mean, media and #30 percentile in a matrix. Thereby, this treatment gave rise to 21 variables later used for the process of correlation. The treatment has been carried out with ERDAS and programming Toolkit. On the other hand, the satellite sub-team investigated a new image classifier based on cellular automata, as alternative to the studies of diffuse classifiers (proposed as task in the project). In the scope of teledetection, cellular automata have been used, above all, for modelling cases of clouds or rivers behaviour, fire propagation simulation, deforestation processes modelling, etc. The project research team created a methodology that applies the cellular automata into classification processes. It has been implemented with Toolkit of Erdas Imagine 9.1 tool. To evaluate the results obtained, we used a confusion matrix with the Kappa index. Such results showed that this methodology improved the classification with supervised classic algorithms: it improved the percentage almost equivalent to the noise percentage in the image. The satellite team is preparing some publications with the results obtained. There exists a previous publication in [16] with preliminary studies.

2.3 Neural-Networks

For the correlation process, we carried out an implementation of neural-networks based on *Radial Basis Function* (RBF) networks. RBFs were used as tools to approximate field-data according to the information gathered from Landsat satellite. The information is mapped to ecological variables associated with vegetation in order to substitute them in the process of map generation. The test was made in an area covering Southeast of Spain and we could obtain ecological variables by using the satellite information in 43 from 45 cases (95% success rate).

The correlation process was the following. A set of neural-networks were trained to relate each ecological variable with the satellite processed information. For each ecological variable there were 21,905 soil samples that represented 1x1Km sectors in which statistical information of 7 satellite bands has been generated. We repeated 32 times the experiment of building a RBF network in the process of modelling every ecological variable. Once the networks had been built, we compared the results obtained with the estimated and we calculated two measures: (a) the precision obtained and (b) the variance between the results obtained and expected that enabled us to know in which ecological variables the approximation confidence was high enough. We implemented the experiment in Matlab on a 3-Core NVIDIA GPUs, 256 cores each one. We used the CUDA technology of GPUs manufacturer. The execution of the program took 48 days of intensive calculation. Now the experiment involves an improvement of calculation time by using the CUDA programming paradigm of NVIDIA company. As this document is being written, NN team is preparing some publications with the results obtained.

2.4 Modelling

Around the correlation methodology, we modelled a framework to keep the knowledge generated. The information involved in this process is stored in a three-level format [8]. Level

1 (data level) represents a set of repositories where the original sources of basic information are stored, that is, satellite images and cartography. These sources may be local resources or external resources to the information system. Level 2 (metadata level) has a group of templates repositories written in OWL that represent metadata of satellite images, cartographical maps and classifications. To obtain OWL templates, we followed an MDD approach where they were designed in UML and then translated into OWL [4]. Later on, in [18] we described an ontological model in UML class diagrams, and presented a formalization of the ontology in *First Order Logic* (FOL) encoded in *Description Logic* (DL) as a previous stage for UML2OWL model transformation.

For design reasons, we followed a trading-based approach to store and query OWL templates [1]. This approach is a traditional solution of a "yellow pages" directory for the interoperability of objects in open and distributed systems; here we have adapted this directory for OWL documents, instead of objects. This "directory" is a repository with meta-metainformation (level 3) of OWL templates (metadata) associated to a trader for subsequent tasks of search, location and retrieval of original information (L1 and/or L2). In the project we defined an Ontological Web Trading agent in a Multi-Agent System (MAS) [2, 19]. OWT presents the ontology needed by some agents to communicate with the trading agent. It uses a Web "Query-Searching/Recovering-Response" model, and the SPARQL query language and the OWL ontology description language to operate information retrieval. Ontologies are used in two different contexts in the project: (a) to represent the application domain information itself (data ontology); and (b) to represent the services that some agents request from others during their interaction (behavioural ontology).

Regarding the user interface module (HCI), we have been working on mechanisms to model user interaction and WIMP (Windows, Icons, Menus and Pointing) user interfaces through a model-driven methodology [7, 12] and we established a general HCI model [17]. On the one hand, in [7] we studied how transforming UML models into class diagrams to represent the objects involved in the user interface by means of class icons. We believe that more complex user interfaces such as static user interfaces (i.e.,HTML, Visual, IDEs in general), active user interfaces (i.e., Applet, JavaScripts, general Scripting in), or dynamic user interfaces (i.e., JSP, ASP, PHP, CGIs in general) could be modeled following our technique.

On the other hand, the *user interface* (UI) —as a significant part of most applications— should be modeled using UML and MDD perspectives, and automatic CASE tools may help to generate user interfaces from UML designs. In [12] we described how to *use* and *specialize* UML diagrams in order to describe the user interfaces of a software system based on WIMP. *Use case diagrams* are used to extract the *main user interfaces.* Use cases are described by means of *user-interaction diagrams*, a special kind of *activity diagrams* in which states represent *data output actions* and transitions represent *data input events.* Input and output interactions in the *user-interaction diagrams* help the designer to extract the user interface components used in each user interface. Furthermore, in order to accomplish the description of user interfaces following a model-driven development (MDD) perspective, in [6] we proposed three specialized UML diagrams called *user-interaction*, *user-interface* and *GUI-class* diagrams. These diagrams can be seen as the UML-based UI models of the system, which concern with code-generation to implement user interfaces. In [5] we also presented an Eclipse GMF tool for modelling user-interaction diagrams —an specialization of the UML state-machines for UI design— as a means to describe the behaviour of user interfaces.

In [17] we designed the interaction issues by means of two kind of agents: (a) interface agents and (b) environmental information agents. The first ones deal with data representation policies (presentation layer, user interface) and the second ones deal with the way in which the system is interrogated. The query implies enquiry preparation, information location and retrieval. Here, the queries are internally defined in SPARQL language [1]. Again in this layer we intended to adapt trading processes. We are now working for building user interfaces according to the dynamic composition of user interface components (COTS-IU) [10], taking into account interaction aspects related to a cooperative system. This perspective is currently in progress and it is expected to obtain results throughout 2010.

Following one of the main lines proposed in the project on MDD research, in [11] we presented a framework for using logic programming (in particular, Prolog) in the specification of model transformations in the UML context. Our approach describes how the UML metamodel can be represented in Prolog, and how model transformations can be defined by means of Prolog rules. It uses rules to specify queries in source models and rules to express how to build the target model. Our approach showed how Prolog can be used to verify constraints on source and target models by means of a model verification language. We validated our proposal by means of a prototype developed under SWIProlog.

3 Result Indicators

According to the achievement degree of estimated outcomes, nearly all of them (neural-networks and modelling) have been successfully achieved: approximately 85% of the project. Although scientific results were not expected in the satellite module in the beginning, the work sub-team developed a new, original classifier for noisy satellite images that is over 5% better than traditional classifiers. Its behaviour is based on Cellular Automata implementation, usually used for dynamic systems simulations (i.e., fire propagation or floods).

Additionally, it has been proved that mapping between cartographic and satellite variables is possible. We believe that this outcome may involve a great advance in the interpretation of new satellite image information, which had been possible only with cartography so far. This project substantially improves time and reduces the cost of ecological sectorization studies, restricting field surveys and automating the creation of these variables. On the other hand, this technique may unlimitedly be applied to any other geographical site with satellite coverage and upon which we wish to obtain referenced vegetation information. Despite the study has been carried out only for 45 variables (related to vegetation), we believe that it is possible to find correspondence in other variables. We think these two results (classifier and correlation) may have social impact as they may introduce a change in the usual methodological procedures of specialists of some environmental management organizations.

From a technological perspective, a comprehensive ontological approach for an EMIS as established for SOLERES, with three-level environmental information modelling through ontologies, using ontology traders and ontology-driven agent architectures, gave rise to a standard information representation framework that may be extended to any other environmental information apart from satellite and cartography. In this regard, the EGMASA company is interested in implementing this solution for environmental information agents.

According to the scientific contribution in the scope of software engineering, there have been great advances for user interface modelling and user interaction through UML and MDD-based approaches. We believe that this contribution may have impact in user interfaces developers, as it represents a formalization mechanism and provides development process.

Given the four modules of the Project, we expected to have scientific results —therefore publications— in the modules of neural-networks and modelling. The other two (ecology and satellite) are a means of achieving the results. In this regard, most publications are located in the module of modelling (where there is approximately 75% of the scientific and technological workload and variability of the project). Neural-networks module is currently pre-published: throughout the year 2010 it is expected at least one publication in international congress and an impact journal. So far the project has been supported by 14 scientific and technological publications, nine of which are contributions into proceedings of conferences (7 international) [1, 2, 4, 5, 8, 10, 11, 16, 17], three are journal articles (all with JCR impact) [7, 12, 3], one is a book chapter (IGI Global Publisher) [6], and one scientific and technical book written by all members of the work-team [17]. There are also 3 journal articles that are currently in the third revision stage [18, 19] and second revision stage [20]. Finally, there exists a collateral research line in structures modelling, where some UML modelling experiences of SOLERES project have been applied with 2 journal publications [3, 14][4].

There are four open lines of research that will lead to doctoral thesis works, two of which are about to be finished and other two started during the year 2009. Regarding the first two lines, one is a proposal for an ontology-driven multiagent architecture that brings together the KRS subsystem of SOLERES as study case in EMIS. The other line is a methodology for spectral classification of satellite images by means of cellular automata and software agents. Both doctoral theses are estimated to be finished by the period 2010-2011. There have been two recent research works already started, one for the automatic generation of user interfaces through intelligent-model transformation and another one on a trading service development for the creation of user interfaces based on interface-component composition. Both works are aimed at searching for solutions in evolutive and cooperative user interfaces.

On the other hand, there exists a close cooperation with our two foreign researchers.Thereby, some team members are expected to stay in foreign organizations throughout the year 2010. Additionally, as an initiative by one of the foreign researchers, in February 2009 we made up a preliminary team to apply for a MED European Project (within European territorial cooperation programme) between Italy, Greece and Spain. The proposal finally failed for team configuration problems by Greek partner. We hope to restart a similar proposal in the future.

Finally, as previously advanced, the team was arranged in four sub-teams to develop the Project (thematic groups): ecology, satellite, neural-networks and modelling. Nevertheless, coordination meetings between sub-teams have been held throughout the project. We organized two workshops (http://www.ual.es/acg/soleres/events/) as events associated with delivery of annual deadlines scheduled in the project. Regarding the staff and execution costs, they are being fulfilled in accordance with the original report submitted to the Ministry.

Acknowledgments. The team of TIN2007-61497 project would like to express his gratitude to OTRI and "*Vicerrectorado de Investigación, Desarrollo e Innovación*" (University of Almeria), as well as EPOs that support the project. Our special thanks to EGMASA and "*Consejería de Medio Ambiente*" (*Junta de Andalucía*, Spain) for providing satellite images from the study area. Our general thanks to MICINN (*Ministerio de Ciencia e Innovación*).

[4]TIN2007-61497 Project and the Ministry are cited in these JCR publications.

References

[1] Criado J, Padilla N, Iribarne L, Asensio JA, Muñoz FJ. Ontological Trading in a Multi-Agent System. *Proc. Int. Conf. on Practical Applications of Agents and Multi-Agent Systems* (PAAMS'2010). 26-28 Apr., Salamanca (Spain), Vol. 2, pp. 268–273, 2010.

[2] Iribarne L, Padilla N, Asensio JA, Muoz FJ, Criado J. Involving Web-Trading Agents & MAS. An implementation for searching and recovering environmental information. In press, *Proc. Int. Conf. on Agents and Artificial Intelligence* (ICAART'2010). 22-24 Jan., 2010, Valencia, Spain.

[3] Iribarne L, Ayala R and Torres JA. A DPS-based system modelling method for 3D structures simulation in manufacturing processes. *Simulation Modelling, Practice and Theory J.*, 17(5):935–954, 2009. doi:10.1016/j.simpat.2009.03.001.

[4] Asensio JA, Iribarne L, Padilla N, Muoz FJ , Ayala R, Cruz M, and Menenti M. A MDE-based Satellite Ontology for Environmental Management Systems. In press 2010, *Information Technology, Information Systems and Knowledge Management.* Springer book. ISBN: 978-0-387-88750-0.

[5] Almendros J, Iribarne L, Asensio JA, Padilla N, and Vicente-Chicote C. An Eclipse GMF Tool for Modelling User Interaction. WSKS 2009. LNCS/LNAI 5736, pp., 405–416. Springer-Verlag Berlin, 2009 doi:10.1007/978-3-642-04754-1_42.

[6] Almendros J and Iribarne L. User interaction and interface design with UML. (Reprinted selected manuscript) Chapter XXXVIII in the book *Human Computer Interaction: Concepts, Methodologies, Tools, and Applications.* P. Zaphiris and C. S. Ang (eds.), IGI Global. Pages 404–431, 2009.

[7] Almendros J and Iribarne L. UML modeling of user and database interaction. *The Computer Journal*, 52(3):348–367, 2009. doi:10.1093/comjnl/bxn028.

[8] Padilla N, Iribarne L, Asensio JA, Muoz FJ, and Ayala R. Modelling an Environmental Knowledge-Representation System. WSKS 2008. 24-28 Sept., Athens (Greece), LNCS-5288, pp. 70–78, 2008. doi:10.1007/978-3-540-87781-3.

[9] Iribarne L, Asensio JA, Padilla N, and Ayala R. SOLERES-HCI. Modelling a human-computer interaction framework for open EMS. WSKS 2008. 24-28 Sept., Athens (Greece), Springer CCIS 19, pp. 320–327, 2008. doi:10.1007/978-3-540-87783-7_41

[10] Asensio JA, Iribarne L, Padilla N, and Ayala R. Implementing trading agents for adaptable and evolutive UI-COTS components architectures. ICE-B'2008. 26-29 July, Porto (Portugal), pp. 259–262, 2008.

[11] Almendros J and Iribarne L. A Framework for Model Transformation in Logic Programming. PROLE'2008. 7-10 Oct, Gijón, Asturias (Spain), pp. 29–39, 2008.

[12] Almendros J, and Iribarne L. An extension of UML for the modeling of WIMP user interfaces. *Journal of Visual Languages and Computing*, 19(6):695–720, 2008. doi:10.1016/j.jvlc.2007.12.004.

[13] Almendros J and Iribarne L. Describing use case relationships with sequence diagrams. *The Computer Journal*, 50(1):116–128, 2007. doi:10.1093/comjnl/bxl053.

[14] Iribarne L, Torres JA, and Peña. Using computer modeling techniques to design multi-tunnel greenhouses. *Computers in Industry J.*, 58(5):403–415, 2007. doi:10.1016/j.compind.2006.09.001.

[15] Almendros J and Iribarne L. User interaction and interface design with UML. Chapter in the book *Visual Languages for Interactive Computing: Definitions and Formalizations.* Fernando Ferri (eds.), IGI Global. Pages 202–222, 2007.

[16] Espinola M, Ayala R, Leguizamion S, and Menenti M. Classification of Satellite Images Using the Cellular Automata Approach. WSKS 2008. 24-28 Sept., Athens (Greece), Springer CCIS 19, pp. 320–327, 2008. doi:10.1007/978-3-540-87783-7_66.

[17] Iribarne L, Torres JA y Aguilera P (eds.) Tecnicas Avanzadas de la Informatica. Fundamentos y Aplicaciones en el Medio Ambiente. *Aconcagua Libros*, Sevilla, 2008, ISBN: 978-84-96178-15-1.

[18] Iribarne L, Padilla N, Asensio JA, Muñoz FJ, Ayala R, Cruz M, Almendros J, and Menenti M. An Open-Environmental Ontology Modeling. *IEEE Trans. Syst. Man Cybern. Part A-Syst. Hum.* In 3rd revision.

[19] Iribarne L, Padilla N, Asensio JA, Muñoz FJ, Ayala R, and Criado J. An Ontological Web-Trading agent approach for information retrieval. *World Wide Web Journal.* In 3rd revision.

[20] Almendros J and Iribarne L. A Prolog-based Approach for Model Transformation. *IEEE Trans. Knowl. Data Eng.* (TKDE). In 2nd revision.

[21] ISO14001. Environmental management systems — Specification with guidance for use. Nov., 2004.