AGENT-BASED INTERDISCIPLINARY FRAMEWORK FOR DECISION MAKING IN COMPLEX SYSTEMS

Marina V. Sokolova
Kursk State Technical University, Russia
University of Castilla-La Mancha, Department of Computing Systems, Spain
Instituto de Investigación en Informática de Albacete, Spain

Antonio Fernández-Caballero, Francisco J. Gómez
University of Castilla-La Mancha, Department of Computing Systems, Spain
Instituto de Investigación en Informática de Albacete, Spain

Keywords: Decision support system, Agents, Complex system, Framework.

Abstract: We offer a framework for the creation of decision support and expert systems for complex natural domains. This is due, on the one hand, to the numerous advantages of intelligent methods of data manipulation and, on the other hand, to the abilities of the computational agents to make decentralized decisions, which are crucial for complex systems modeling and simulation.In our approach, the qualitative improvement in decision making is obtained by using computational agents and interdisciplinary approach. The frameworks combines, on the one hand, the numerous advantages of intelligent methods for data manipulation and, on the other hand, the abilities of the computational agents to make decentralized decisions, which is crucial for complex system modeling and simulation. The approach contributes to decentralization and local decision making within the standard workflow. We demonstrate our framework in a case study and discuss obtained results.

1 INTRODUCTION

When the idea of creating a symbiotic human-computer system to increment accessible knowledge for decision making in complex problems appeared, it was firstly applied for managerial and business domains. Later, the initial domains of decision support systems (DSS) application were widened, and the concept of DSS was spread out to manifold spheres and fields of human activities, extending not only to technical but also to complex ill-determined domains (environmental, medical, social issues, etc). With time diverse DSS have appeared, which enhanced a number of models for use such as preprocessing, optimization, hybrid and simulation models.

The use of DSS is of great importance now for complex natural phenomena studies, because they allow specialists to quickly gather information and analyze it in order to understand the real nature of the processes, their internal and external dependencies, and the possible outcomes while making actions and correcting decisions. The technical areas where these systems could help vary from the storing and retrieving of necessary records and key factors, examination of real-time data gathered from sensors, analysis of tendencies of complex natural processes, retrospective time series, making short and long-term forecasting, and in many other cases (Riaño et al., 2001),(Athanasiadis and Mitkas, 2004),(Sokolova and Fernández-Caballero, 2008), (Sokolova and Fernández-Caballero, 2009).

In this article we will present our framework for creation of agent-based decision support and expert systems, focusing firstly on its principal formalisms and phases of decision making process for a complex natural domain (part 2) and revise related works (part 3), then presenting our approach for creation a framework for complex system design in complex domains (part 4) and demonstrating an application of the framework for the case study and discuss obtained results (part 5) and, lastly, we will comment directions of future work (part6).

V. Sokolova M., Fernández-Caballero A. and J. Gómez F. (2010).
AGENT-BASED INTERDISCIPLINARY FRAMEWORK FOR DECISION MAKING IN COMPLEX SYSTEMS.
In *Proceedings of the 2nd International Conference on Agents and Artificial Intelligence - Agents*, pages 96-103
DOI: 10.5220/0002732000960103

2 DECISION MAKING FOR COMPLEX SYSTEMS

Human activity increases constantly, and both the scale and speed of the human influence upon the natural, social, economical, and other processes grows significantly, so, it is impossible now not to take it into account as one of the motive forces in ”human - nature - technology” arena.

The science of today has reached significant results in modeling and control over man-made technical systems. Notwithstanding, effective managing natural complex phenomena often lies beyond of our possibilities.

Generally speaking, a complex systems (CS) is a composite object, which consists of many heterogeneous (and on many occasions complex as well) subsystems, and has emergent features, which arise from interactions within the different levels. Such systems behave in non-trivial ways, originated in composite functional internal flows and structures of the CS. As a rule, researchers encounter difficulties when trying to model, simulate and control complex systems.

Due to this, it would be correct to say that to puzzle out the natural complex systems paradigm is one of the crucial issues of modern science. Because of high complexity of CS, traditional approaches fail in developing theories and formalisms for their analysis. Such study can only be realized by a cross-sectoral approach, which uses knowledge and theoretical backgrounds from various disciplines as well as collaborative efforts of research groups and interested institutions.

The term ”system” has a number of definitions and one of them sounds like ”a set of interdependent or temporally interacting parts”, where parts are generally systems themselves, and are composed of other parts in their turn (Levin, 2006). If any part is being extracted from the system, it looses its particular characteristics (emergency), and converts into an array of components or units.

Thus, an effective approach to CS study is to follow the principles of the system analysis (Rechtin, 1999), when we have to switch over to the abstract view of the system and perform the following flow of tasks:

- Description of a system. Identification of its main properties and parameters.
- Study of interconnections amongst parts of the system.
- Study of external system interactions with the environment and with other systems, etc.
- System decomposition and partitioning.
- Study of each subsystem or system part.
- Integration of results obtained on the previous stages.

However, it is obvious that it is impossible to create an overall tool, which would play a role of a ”magic wand” to study any complex adaptive problem. The process of DSS and expert system (ES) design is laborious, and has many requirements, the crucial one are: collaboration of the specialists from various domains (experts, developers, analysts, mathematics, programmers, etc.). The discussed peculiarities and difficulties of CS study constrain multiple requirements on experts and developers, those have to do not with the problem in question, but with theories and formalisms from different disciplines.

Thus, the process of decision making and expert systems design should be facilitated, and a portfolio of methods and tools have to be offered to specialists working with complex adaptive systems.

3 CONTEMPORARY APPROACHES TO COMPLEX SYSTEM STUDY

Recognition that a complex natural phenomenon has to be studied using the principles of the system approach, induced some theoretical insights and their practical realization. But, it should be said, that many researchers still tend to use mono-discipline approaches and do not zoom in the complex emergent behavior of CS and their causal-effect relations with the environment.

Having revised a number of research works and publications, we can mark out some general tendencies for complex systems study.

Some research group offer ”island solutions”, which, as a rule, are oriented to evaluation or assessment of a few parameters or indicators, in other words, they are dedicated to resolve specific goals. For example, a system which provides indicator assessment for a particular case study or for a limited area. Domain ontologies for such systems are limited though may suffer from possible heterogeneity. As a rule, such systems are effective when working within the application domain and very sensitive to any unforeseen changes (Karaca et al., 2009),(Urbani and Delhom, 2005).

Multi-functional systems provide multiple analysis of input information, can be based upon hybrid techniques, possess tools and methods of data pre- and post-processing, modelling, simulation. They are less sensitive to changes in application domain, as they have tools to deal with uncertainty and heterogeneity (Nastar and Wallman, 2009), (Terry Bossomaier and Thompson, 2005).

Methodologies for software applications development may support the mayor part of the system life cycle phases, starting with the initial system planning, which include system analysis and domain (problem) analysis phases, and then assist and provide EIS design, coding, testing, implementation, deployment and maintenance. In this case, consolidate cooperation of specialists from various domains and with various backgrounds is necessary (Gorodetski et al., 2004), (Rotmans, 2006).

Basically, in spite of the diversity of existing approaches of DSS creation, obtained and reviewed outcomes, it is not possible to elaborate a uniform overall tool, capable of dealing with various domains and create adequate solutions. However, the quality of solutions and multi-function tools upsurges, because they perform better results when these tools are oriented to limited and specific domains.

Even then, exists a necessity to elaborate a general methodology for DSS design oriented to distributed heterogeneous domains, and powerful enough to facilitate work not only for small, but also for distributed and numerous research groups.

4 OUR APPROACH TO THE FRAMEWORK CREATION

4.1 Interdisciplinary Approach

In general terms, a framework facilitates development of an informational system, as it offers a consequence of goals and works to do. In our case the informational system in the concept definition given above (see part 2), represents a CS or a natural phenomenon.

Following the stages, determined in a framework, developers and programmers will have more possibilities to specialize in the domain of interest, meet software requirements for the problem in term, thereby reducing overall development time.

An effective framework should, on the one hand, be general enough to be applicable for various domains and, on the other hand, be adaptable for specific situations and problem areas. Moreover, the framework should be based of the interdisciplinary approach, which results from the melding of two or more disciplines. Being applied to a complex system, the approach includes the principal works: (1) the complex system is decomposed into components if necessary, the process of decomposition is repeated; (2) then, each of the subsystems is studied by means of the ”proper” techniques, belonging to the respected discipline, or/and of hybrid methods; (3) managerial process of decision making is realized, with feedback and possible solutions generation.

As we have accentuated in the previous part, the most important principles of CS is that they can not be studies from a mono-discipline viewpoint, and it is necessary to provide a complex hybrid application of methods and techniques from various disciplines. Using intelligent agents seems to be an optimal solution in this case (Weiss, 1999).

Actually, MAS helps to create cross-disciplinary approaches for data processing, and, hence, for CS study. An agent may include nontraditional instruments to bear from different domains. Roles played by an agent depend on the system (or subsystem) functions and aims. There are no restrictions or limitations put on knowledge and rule base, used by each agent (Sokolova and Fernández-Caballero, 2009).

4.2 The Main Phases of The Framework

The purpose of our framework is to provide and facilitate complex systems analysis, simulation, and, hence, their understanding and managing. From this standpoint, and taking into account results and insights, given in the previous parts, we implement the principles of system approach.

The overall approach we use is straightforward: we decompose the system into subsystems, and apply intelligent agents to study them, then we pool together obtained fragments of knowledge, and model general patterns of the system behavioral tendencies (Sokolova and Fernández-Caballero, 2007). The framework consists of the three principal phases:

Preliminary Domain and System Analysis. This is the initial and preparation phase when an analyst in collaboration with experts study the domain of interest, extract entities and discover their properties and relations. Then, they state main and additional goal of the research, possible scenarios and functions of the system. During this exploration analysis, they search answer to the questions: *what* the system has to do and *how* it has to do it. As a result of this collaboration it appears meta-ontology and knowledge base.

This phase is supported by the Protege Knowledge editor, that implements the meta-ontology, and by the Prometheus Design Kit, which we use to design multi-agent system and generate skeleton code for further implementation of the DSS.

System Design and Codification. The active ”element” of this phase is a developer. As supporting software at this phase, we used the Jack^{TM}Intelligent Agents and Jack^{TM}Development Environment. Once the codification is finished and the system is tested, th second phase is completed.

Simulation and Decision Making. This is the last phase of the framework and it has a special mission. During this phase the final users - decision makers - can interact with the system, constructing solutions and policies, estimating consequences of possible actions on the basis of simulation models.

5 THE CASE STUDY: ENVIRONMENTAL IMPACT ASSESSMENT ON HUMAN HEALTH

5.1 Intelligent Agents

The general system organization is coherent to the principal phases,described above.

Data search, retrieval, fusion and pre-processing are realized by two intelligent agents: the *Data Aggregation agent* (DAA) and the *Data Preprocessing agent* (DPA), which do a number of tasks, following the workflow:

Information Search - Data Source Classification - Data Fusion - Data Preprocessing - Believes Creation

Data mining procedures are based on the *Function Approximation agent* (FAA) and its team of agents. The principal tasks to be solved here are: " to state the environmental pollutants that impact on every age and gender group and determine if they are associated with examined diseases groups; " to create the models which explain dependencies between diseases, pollutants and groups of pollutants.

Here we are aimed to discover the knowledge in form of models, dependencies and associations from the pre-processed information which comes from the previous logical layer. The workflow of this level includes the following tasks:

State Input and Output Information Flows - Create models - Assess Impact - Evaluate models - Select Models - Display The Results

Decision generation, simulation and human-computer interaction are realized by the *Computer Simulation agent* (CSA). The system works in a cooperative mode, and it allows the decision maker to modify, refine or complete decision suggestions, provided by the system and validate them. This process of decision improvement is repeated until the consolidated solution is generated. The workflow is represented below:

State Factors for Simulation - State the Values of Factors - Simulate - Evaluate Results - Check Possible Risk - Display The Results - Receive Decision Maker Response - Simulate - Evaluate Results - Check Possible Risk - Display The Results

Agents communicate to each other and are triggered by events and messages that they send. Agents also share common data. Preliminary system design was realized in the Prometheus Development Kit.

5.2 Methods used by Agents

The DAA has a number of subordinate agents under its control; these are the *Domain Ontology agent* and the fusion agents: the *Water Data Fusion agent*, the *Petroleum Data Fusion agent*, the *Mining Data Fusion agent*, the *Traffic Pollution Fusion agent*, the *Waste Data Fusion agent* and the *Morbidity Data Fusion agent*. At the beginning of the execution the DAA firstly send a message to the Domain Ontology agent, which reads information from metaontology, and returns in to the DAA. Then, the DAA starts to searches for information sources and reviews them trying to find if there was a key ontological concept there. If the file contains the concept, the Data Aggregation agent sends an internal event to start data retrieval to the specialized fusion agent. The plan responsible for execution with the identified concept starts reading the information file and searching for terms of interest.

The DPA provides data preprocessing and has a number of subordinate agents which specialize in different data clearing techniques: the *Normalization agent*,the *Correlation agent*, the *Data Smoothing agent*,the *Gaps and Artifacts Check agent.* They perform all data preprocessing procedures, including outliers and anomalies detection, dealing with missing values, smoothing, normalization, etc.

The FAA has a hierarchical team of subordinate agents, which serve to support the roles: "Impact Assessment", "Decomposition" and "Function Approximation". FAA has under its control a number of data mining agents: the *Regression agent*, the *ANN agent,* and the *GMDH agent*, which work in a concurrent mode, reading input information and creating models. Then, if any agent from this group finishes modeling, it calls for the *Evaluation agent*, which evaluates obtained models, and return the list of the accepted ones, while the others are banned and deleted. The FAA pools the output of the agents work, creates the list with the accepted models and then, once the *Regression agent*, the *ANN agent,* and the *GMDH agent* finished their execution, calls for the *Committee Machine agent*, which creates the final models in form of committees for each of the dependent variables, and validates them.

The methods which execute agents of the FAA team are the following:

- Statistical modelling - include linear, non-linear simple and multiple models and their evaluation.
- Artificial Networks modelling - include feed-forward neural networks, trained by back-propagation algorithm, genetic algorithms, radial-based function networks.
- Group Method of Data Handling (GMDH), which is one of the powerful methods of self-organization. We used the combinatorial algorithm of GMDH (Madala, 1994).
- Committee Machines, which provide creation of weighted hybrid models with the condition that the best models selected for the committee have more impact on the final result (Haykin, 1998).

The CSA interacts with user and performs a set of tasks within ”Computer Simulation”, ”Decision Making” and ”Data Distribution” roles. It has the agent team, which includes *Forecasting agent*, *Alarm agent* and *View agent*.

The process of simulation and generation of possible solutions is interactive. The human-computer interaction sessions are organized by the *View agent*, which offer window -forms, graphical and textual documentation as supporting tools. The *Forecasting agent* calculated predicting values for the CSA, and the *Alarm agent* checks if these values satisfy permitted standards. Figure 1 gives a look at agent teams.

5.3 Simulation Results

The MAS has an open agent-based architecture, which would allow us an easy incorporation of additional modules and tools, enlarging a number of functions of the system. The system belongs to the organizational type, where every agent obtains a class of tools and knows *how* and *when* to use them. Actually, such types of systems have a planning agent, which plans the orders of the agents' executions. In our case, the main module of the JACK^{TM}program carries out these functions.

The *View agent* displays the outputs of the system functionality and realize interaction with the system user. As the system is autonomous and all the calculations are executed by it, the user has only access to the result outputs and the simulation window.

Table 1: Part of the Table with the outputs of impact assessment.

N	Disease	Pollutant
1	Neoplasm	Nitrites in water; Miner products; DBO5; Asphalts; Dangerous chemical wastes; Fuel-oil; Petroleum liquid gases; Water: solids in suspension; Non-dangerous chemical wastes;
2	Diseases of the blood and blood- forming organs, the immune mechanism	DBO5; Miner products; Fuel-oil; Nitrites in water; Solids in suspension; Dangerous metallic wastes Miner products;
3	Pregnancy, childbirth and the puerperium	Kerosene; Petroleum liquid gases; Solids in suspension; Petroleum; DQO; Fuel-oil; Asphalts; Gasohol; DBO5; Water: Nitrites. Petroleum; Petroleum autos;

To evaluate the impact of environmental parameters upon human health in Castilla-La Mancha, in general, and in the city of Albacete in particular, we have collected retrospective data since year 1989, using open information resources offered by the Spanish Institute of Statistics and by the Institute of Statistics of Castilla-La Mancha.

As indicators of human health and the influencing factors of the a environment, which can cause negative effect upon the noted above indicators of human health, the levels of contamination of potable water, air, soil, and indicators of traffic activity, hazardous wastes, energy and used annually, etc.

The DSS has recovered data from plain files, which contained the information about the factors of interest and pollutants, and fused in agreement with the ontology of the problem area. It has supposed some necessary changes of data properties (scalability, etc.) and their pre-processing. After these procedures, the number of pollutants valid for further pro-

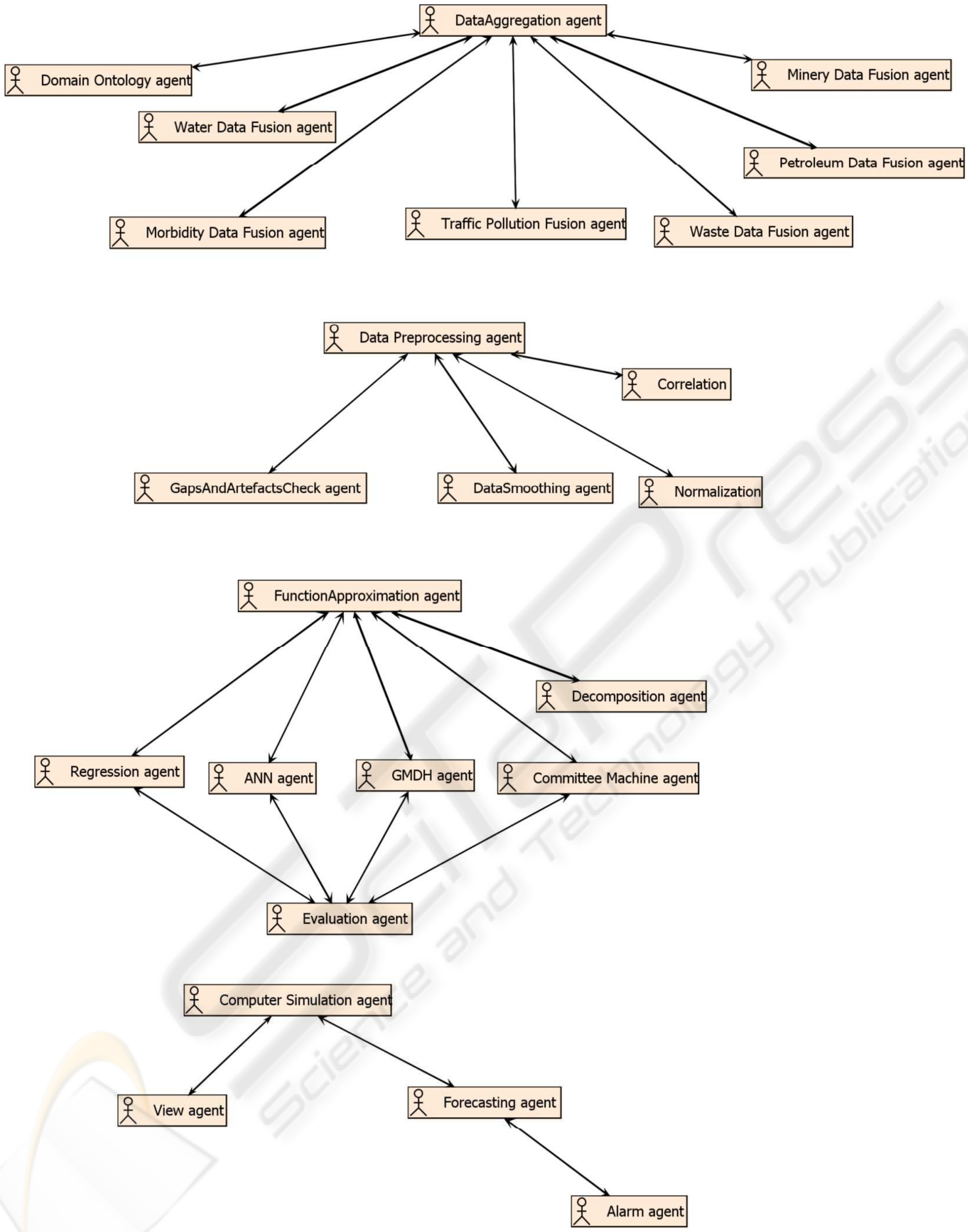

Figure 1: The agent teams created for the case study.

cessing has decreased from 65 to 52. This significant change was caused by many blanks related to several time series, as some factors have started to be registered recently.

After considering this as an important drawback, it was not possible to include them into the analysis. The human health indicators, being more homogeneous, have been fused and cleared successfully. The impact assessment has shown the dependencies between water characteristics and neoplasm, complications of pregnancy, childbirth and congenital malformations, deformations and chromosomal abnormalities.

The MAS has a wide range of methods and tools

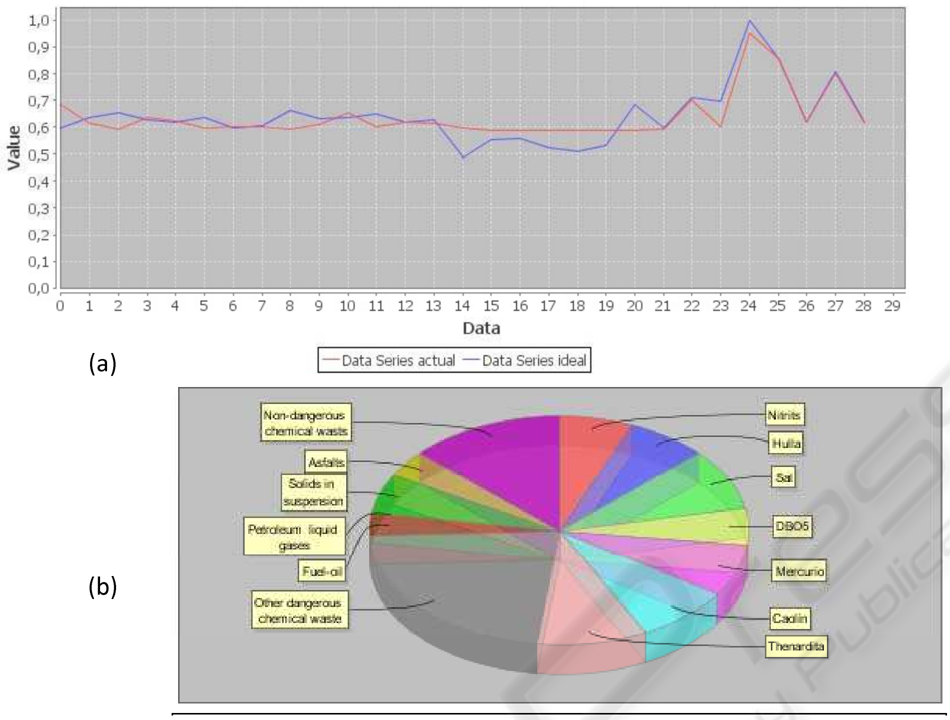

Figure 2: (a) The final model for one of the diseases of the case study. (b) The outcomes of the impact assessment for one of the diseases of the case study.

for modeling, including regression, neural networks, GMDH, and hybrid models, based on committee machines. The function approximation agent selected the best models, which were: simple regression - 4381 models; multiple regression - 24 models; neural networks - 1329 models; GMDH - 2435 models. The selected models were included into the committee machines. We have foretasted diseases and pollutants values for the period of four years, with a six month step, and visualized their tendencies, which are going to overcome the critical levels. Control under the "significant" factors, which cause impact upon health indicators, could lead to decrease of some types of diseases.

Committee machines provide universal approximation, as the responses of several predictors (experts) are combined by means of a mechanism that does not involve the input signal, and the ensemble average value is received. As predictors we used regression and neural network based models.

An example of the final model, received by committee machine for the "'Neoplasms"' for the region of Castilla-La Mancha, Spain, is given on Figure 2a, and the results of the impact assessment are shown on Figure 2b.

6 CONCLUSIONS

Complex system analysis and modelling is still a complicated problem, and an efficient framework which supports decision making and expert systems creation facilitates research efforts for various research groups. The framework we have demonstrated in the article, is one of the approximations to solve this complicated task. And, to conclude with, we should note some essential advantages we have reached, and some directions for future research.

First, the framework is interdisciplinary and is flexible for changes: it can be adapted to comply with specific features of the domain of interest. The prototype of the DSS, created in accordance with the framework, supports decision makers in choosing the behaviour line (set of actions) in such a general case, which is potentially difficult to analyze and foresee. As for any complex system,the human choice is decisive.

Second, in spite of our time consuming modeling work, we are looking forward to both revising and improving the system and deepening our research.

Third, we consider making more experiments varying as with data structure, trying to apply the system to the other.

The framework supports all the necessary steps for standard decision making procedure by utilizing intelligent agents. The teams of intelligent agents, that are logically and functionally connected, have been presented. Real-time interaction with the user provides a range of possibilities in choosing one course of action from among several alternatives, which are generated by the system through guided data mining and computer simulation. The system is aimed to regular usage for adequate and effective management by responsible municipal and state government authorities.

We used both traditional data mining techniques, hybrid and specific methods, with respect to data nature (incomplete data, short data sets, etc.). Combination of different tools let us gain in quality and precision of the reached models, and, hence, in recommendations, which are based on these models. Received dependencies of interconnections and associations between the factors and dependent variables help to correct recommendations and avoid errors.

ACKNOWLEDGEMENTS

This work was partially supported by Spanish Ministerio de Ciencia e Innovacion TIN2007-67586-C02-02, and Junta de Comunidades de Castilla-La Mancha PII2I09-0069-0994 and PEII09-0054-9581 grants.

REFERENCES

Athanasiadis, I. N. and Mitkas, P. A. (2004). An agent-based intelligent environmental monitoring system. *CoRR*, cs.MA/0407024.

Gorodetski, V. I., Karsaev, O., Samoilov, V., Konushy, V., Mankov, E., and Malyshev, A. (2004). Multi-agent system development kit. In *Intelligent Information Processing*.

Haykin, S. (1998). *Neural Networks: A Comprehensive Foundation*. Macmillan, New York.

Karaca, F., Anil, I., Alagha, O., and Camci, F. (2009). Traffic related pm predictor for besiktas, turkey. In Athanasiadis, I. N., Mitkas, P. A., Rizzoli, A. E., and Gómez, J. M., editors, *ITEE*, pages 317–330. Springer.

Levin, M. S. (2006). *Composite Systems Decisions (Decision Engineering)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Madala, H.R., I. A., editor (1994). *Inductive Learning Algorithms for Complex Systems Modelling.* CRC Press Inc., Boca Raton, Ann Arbor, London, Tokyo.

Nastar, M. and Wallman, P. (2009). An interdisciplinary approach to resolving conflict in the water domain. In *Information Technologies in Environmental Engineering Proceedings of the 4th International ICSC Symposium Thessaloniki, Greece*.

Rechtin, E. (1999). Systems architecting of organizations: why eagles can't swim. CRC Press.

Riaño, D., Sànchez-Marrè, M., and R.-Roda, I. (2001). Autonomous agents architecture to supervise and control a wastewater treatment plant. In *IEA/AIE*.

Rotmans, J. (2006). Tools for integrated sustainability assessment: A two-track approach. *Integrated Assessment*, 6(4).

Sokolova, M. V. and Fernández-Caballero, A. (2007). A multi-agent architecture for environmental impact assessment: Information fusion, data mining and decision making. In *ICEIS 2007 - Proceedings of the Ninth International Conference on Enterprise Information Systems, Volume AIDSS, Funchal, Portugal*.

Sokolova, M. V. and Fernández-Caballero, A. (2008). Facilitating mas complete life cycle through the protégé-prometheus approach. In *Agent and Multi-Agent Systems: Technologies and Applications, KES-AMSTA, Incheon,Korea.*

Sokolova, M. V. and Fernández-Caballero, A. (2009). Data mining driven decision making. In *ICAART 2009 - Proceedings of the International Conference on Agents and Artificial Intelligence, Porto, Portugal*.

Sokolova, M. V. and Fernández-Caballero, A. (2009). Modeling and implementing an agent-based environmental health impact decision support system. *Expert Syst. Appl.*, 36(2):2603–2614.

Terry Bossomaier, Denise Jarratt, M. M. A. T. S. and Thompson, J. (2005). Data integration in agent based modelling. *Complexity International*, 11.

Urbani, D. and Delhom, M. (2005). Water management policy selection using a decision support system based on a multi-agent system. In *AI*IA*.

Weiss, G., editor (1999). *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT Press, Cambridge, MA, USA.