

Evolutionary Computation and Case-Based Reasoning Interoperation in IEDSS through GESCONDA

Miquel Sànchez-Marrè¹, Karina Gibert², Radha K. Vinayagam³, Beatriz Sevilla-Villanueva¹

¹Knowledge Engineering & Machine Learning Group (KEMLG), Computer Science Dept.,
Universitat Politècnica de Catalunya-BarcelonaTech, Jordi Girona 1-3, 08034 Barcelona, Catalonia
(miquel@lsi.upc.edu)

²Knowledge Engineering and Machine Learning Group (KEMLG), Statistics and Operations Research
Dept., Universitat Politècnica de Catalunya-BarcelonaTech, Jordi Girona 1-3, 08034 Barcelona,
Catalonia (karina.gibert@upc.edu)

³SASTRA University, Tamil Nadu, India
(vrk.1993@gmail.com)

Abstract: Interoperability of Intelligent Environmental Decision Support Systems (IEDSS) is one open challenge in IEDSS field. This paper shows the interoperability of Evolutionary Computation, concretely Genetic Algorithms (GA), and Case-Based Reasoning (CBR) in IEDSS through the GESCONDA tool. GESCONDA is a tool for the deployment of Intelligent Decision Support Systems. This interoperability has been tested with several domains with different purposes like classification tasks, predictive tasks, etc. In the paper, the application in one environmental domain is described and analysed. The experimentation results indicate that this interoperation of both methods can improve the results of the application of one single method, CBR or GA. Thus, the potential of this kind of interoperation seems to be very good and it is an illustrating example of the benefits of Interoperable Intelligent Environmental Decision Support Systems.

Keywords: Genetic Algorithms, Case-Based Reasoning, Model Interoperability, Intelligent Environmental Decision Support Systems.

1. INTRODUCTION

IEDSS are built using some *empirical models* (from AI field or from Statistical field) and *mechanistic models* or *integrating several models* to be more powerful, jointly with other components like GIS, environmental/health ontologies, and some economic components. In addition, some knowledge/data can be mined through the intelligent analysis of large databases coming from historical operation of the environmental process. Knowledge/data mining (model production), as well as reasoning over the produced models are crucial for IEDSS. And reasoning over those models requires high degrees of interoperation among them for achieving reliable and more accurate IEDSSs.

Single AI models provide a solid basis for construction of reliable and real applications, but combining the contributions of different models together can produce better and more accurate IEDSSs exploiting the different characteristics of the different models. Interoperability is the key to make possible the interconnection of different models to solve complex problems in environmental systems. Therefore, the interoperability of AI/Numerical models is one of the main open challenges in this field. Interoperability is the ability of two or more systems or components to exchange information and to use the information that has been exchanged (IEEE, 1990). Additionally, when the components share a common understanding of the information model behind the data being interchanged, semantic interoperability is achieved (Manguinhas, 2010). GESCONDA (Sànchez-Marrè et al., 2010; Gibert et al., 2006) is a tool for the deployment of Intelligent Decision Support Systems, including all main steps like data cleaning, data mining, model validation, reasoning abilities to generate solutions, and predictive models to support final users.

As a particular case, interoperability between Case-Based Reasoning and Genetic Algorithms is tackled in this work, as Case-Based Reasoning and Genetic algorithms can interact in several ways. In the Case-Based Reasoning (CBR) step of adaptation or reuse, a Genetic Algorithm (GA) can be used as an adaptation scheme for optimizing the proposed solution of CBR. On the contrary, a GA algorithm could be used to select the most appropriate configuration of a CBR method, to solve a concrete problem. In this case, each individual of the population represents a different configuration of the CBR system (k value, dissimilarity measure, weighting scheme, etc.), and the fitness function will be the accuracy of the CBR execution for each configuration. By now, the first interaction mentioned before has been implemented in the GESCONDA tool.

2. RELATED WORK AND BACKGROUND

Semantic integration and/or semantic interoperation has been the focus of some research works coming from environmental modelling field. Some pioneer work was done in semantic integration of Environmental models for application to global information systems and decision making, specially related to GIS components and models (Mackay, 1999; Wesseling *et al.*, 1996). In addition, some work was done in model and data integration and re-use in EDSS (Rizzoli *et al.*, 1998) and an overview of model integration was analysed in (Argent, 2004). An interesting work was presented in (Sottara *et al.*, 2012) using the Drools Rule-based integration platform using a unified data model and execution environment. An approach for a framework to develop IEDSS is presented in (Sánchez-Marrè, 2014). In the information systems field several recent works were done in semantic integration of business components (Elasri and Sekkaki, 2013; Kzaz *et al.*, 2010). Other interesting works focused on the semantic interoperability through service-oriented architectures (Vetere and Lenzerini, 2005). Although there are some architecture proposals in the literature to combine some of these models, there is not a common framework to be taken into account as first guideline to deploy Interoperable IEDSS providing an easy way to integrate and (re)use several AI models or statistical/numerical models in a whole IEDSS. Until now, most of the interoperability of the models is achieved by a manual *ad hoc* model interaction. In our research, an interoperable IEDSS has been constructed. The interoperability has been preliminary achieved by an *ad hoc* interconnection through the software codification as a first approach, but there is an open research work aimed to provide a useful and *systematic approach* to interoperate different models at different steps in the IEDSS solving process (Sánchez-Marrè, 2014) sharing a common interchange format, providing a semantic interoperation.

3. GESCONDA TOOL

GESCONDA tool was designed as a four-layer architecture for the development of IEDSS, connecting the user with the environmental system or process. These 4 layers are: data filtering & pre-processing, recommendation and meta-knowledge layer, data mining layer, and knowledge management and reasoning layer. The architecture is depicted in figure 1.

GESCONDA provides a set of heterogeneous techniques that will be useful to acquire relevant knowledge from environmental systems, through available databases. This knowledge will be used afterwards in the implementation of reliable IEDSS. The portability of the software is provided by a common Java platform. The *Data filtering & pre-processing layer* provides a set of tools for data cleaning. Statistical one-way and two-way are provided, missing value and outlier value management, graphical visualization and variable transformation operators are integrated in this layer. Cleaned data will be used afterwards at the other layers to produce data mining models, which will be executed at the reasoning level. The *recommendation and meta-knowledge layer* includes two modules: the recommender and meta-knowledge module, and the feature relevance module. The former one (Gibert *et al.*, 2010) let the user to be assisted to select the most suitable methods to be applied in front of a real situation, by taking into account the goals of the user, main features of the domain and their structure. In addition, the meta-information associated with the data can be managed, both from the features and the observations. The feature relevance module provides GESCONDA with a set of implemented feature weighting algorithms, which determine the weight or relevance of each one of the features describing the data. There are several unsupervised methods and some supervised methods. The *data mining layer* is the layer joining the modules containing several data mining models, which can be induced from the data. There are four modules developed: the clustering

techniques module, the decision tree techniques module, the rule induction module, and the statistical modelling module. Currently an Association Rule module is being implemented. At the *knowledge management and reasoning layer* is where the new reasoning capabilities provide GESCONDA with analytical, synthetic and predictive skills (model execution). Until now GESCONDA was an Intelligent Data Mining and Discovery tool, which was able to produce several knowledge models, but no problem solving skills were given to the user. The past integration of a rule-based reasoning and a case-based reasoning (Sánchez-Marrè et al., 2010) module has led GESCONDA to become an actual Intelligent Decision Support tool, which assists the user to all the steps of the problem solving cycle: diagnosis, planning/solution generation, prediction, and finally decision support.

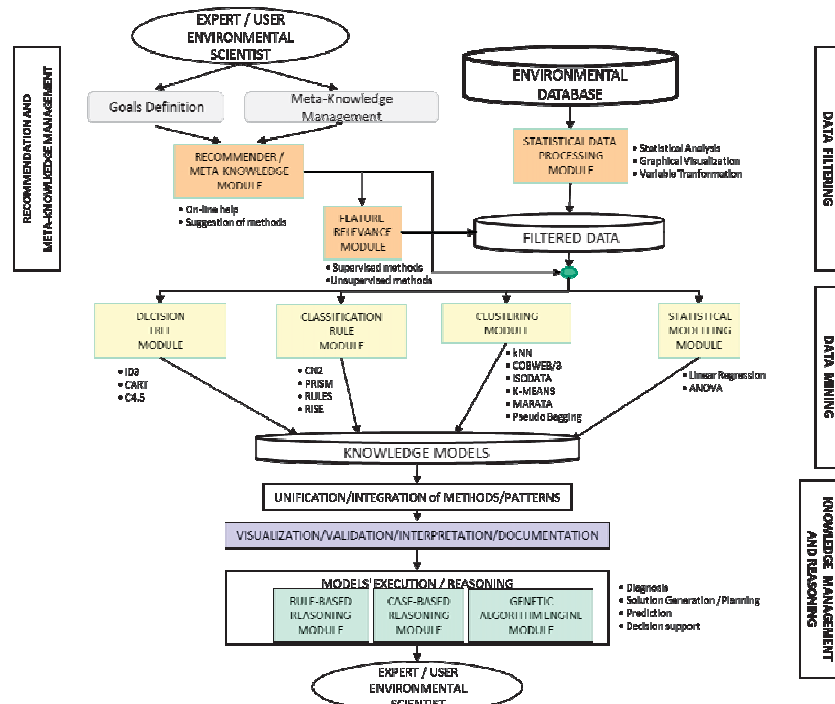


Figure 1. The GESCONDA architecture

Currently, the rule-based reasoning module is being interconnected with the decision tree module and the classification rule module to be able to use models produced by both modules achieving interoperability between model producers and model executors. The Clustering module can interoperate with the classifier modules providing the class variable information to them. This interoperation is among model producers. Now, a new evolutionary computing reasoning module (e.g. a genetic algorithm reasoning engine) has been implemented. This is an interoperation among model executors, which was not yet accomplished in the GESCONDA tool.

4. INTEROPERATING CASE-BASED REASONING AND GENETIC ALGORITHMS

The two reasoning modules have been interconnected to enhance the abilities of the tool. To achieve interoperation, an ad-hoc interconnection have been made, because a more general interoperability approach is just being designed (Sánchez-Marrè, 2014), to be implemented in the near future.

4.1 Case-Based Reasoning

Case-Based Reasoning (Kolodner, 1993) is a flexible reasoning and problem solving technique based on using previous similar solved problems (experiences or cases) to solve new problems, without starting to construct the solution from scratch. One of the key points is the adaptation step where the solution/s of the most similar problem/s are reused to derive the solution of the current new case.

The Case-Based Reasoning implemented in GESCONDA is conceived to be a Flexible Data Intensive Based Reasoning. Flexible in terms of extending its functionality and configuration of its functions (Sevilla & Sánchez-Marrè, 2010). This system could be launched interacting with the user at any step (retrieve, reuse, revise and retain) or defining the new case and compute the whole cycle. The main functionalities offered are the following: regarding *Case Structure*. Since the data is shared in the GESCONDA tool, it is necessary to define which attributes belongs to the description part or to the solution part. It could be introduced by a file or by the interface.

Retrieval step: Search for the most similar cases to the new case. It depends on the *indexation* of the case base: in future releases, hierarchical structures and self-organizing maps will be integrated; on a threshold: Maximum number of cases to retrieve or/and maximum distance; and on an specific distance to be used, from those available in GESCONDA

Reuse step: Adaptation of the retrieved cases solutions to the new solution. The options are: Copy, where the old solution is directly transferred; Mean/Mode, where the mean or mode of the solution is used; Weighted Mean, where the cases are weighted by its distance to the new case or a given attribute (i.e. utility or evaluation attribute); Formula, where the user can introduce a formula for the numerical solution attributes

Revise step: Evaluation of the proposed solution. By the end-user; Approximation of the evaluation assessed by the evaluation of the cases that have been used to create the new one. In test mode, it is possible to compare the real solution with the proposed one.

Retain step: The system learns the proposed case, adding it to the Case Base. This process could be managed by the end-user. Also, it is possible to define a list of conditions that must be satisfied to store the new case in the case base. Each condition is defined in a XML file, and has its own Java class where it is assessed. The introduction of new conditions is trivial. Already implemented: based on distance, evaluation and a recursive cluster elimination (RCE).

At the *execution*, there is the possibility to run more than one case.

- Individual: one case is solved, from the case base or defined by the user.
- Battery: a list of cases is solved, from the case base or a file.

There is also the Mode of execution, which could be normal mode or test mode: in the test mode it creates new solution attributes in order to compare with the existing ones and to compute the percentage of success or error measures for numerical solution attributes.

4.2 Genetic Algorithm Reasoning Engine

Evolutionary Computation (EC) is a bio-inspired approach mimicking *natural selection process* in biological populations. Genetic Algorithms were proposed by John Holland (Holland, 1975) based on the concepts of Charles Darwin's natural evolution theory.

A Genetic Algorithm (Goldberg, 1989) is a biologic random search technique. It begins with a set of k randomly generated individuals, called the *population*. Each individual is coded as a string over a finite alphabet (commonly a binary code). Next generation of the population is produced after some *genetic operators* (selection, crossover, mutation, etc.) have been applied to some probabilistically-selected individuals. Each individual is rated according to an evaluation function, named the *fitness function* which is correlated to their associated probabilities. Selected individuals for *reproduction* are combined using a randomly *crossover* point of cut among the position in the string. The *offspring* are created by crossing over the parent strings at the crossover point. Finally, each new individual is subject to random *mutation* of some positions with some small probabilities. Since best individuals are selected and reproduced every time, convergence to the best individual (the optimal) is expected.

Researchers in artificial intelligence field tried applying genetic algorithms to solve several problems like optimization problems, search problems, classification tasks, rule induction, feature selection, etc. This is a probabilistic method which searches for the best solution from the population of possible solutions. Genetic algorithms have three basic *operators*: selection, cross over and mutation. Genetic algorithms work on encoded individuals, thus representation of solutions in the form of encoded individuals is the first step. Each of these operators has several different choices.

Representation: Representation of individuals depends on the applications. There are several representation techniques like binary encoding, value encoding, tree encoding, etc.

Fitness function: To evaluate the quality of the individuals a function called fitness function is needed. Fitness function depends on the application too.

The three basic operators currently implemented in GESCONDA are:

- *Selection*: This operator selects individuals from current population to process them for producing the new generation of population called offspring. Its objective is to improve the

quality of individuals in further generations. There are several techniques available in selection: *Fitness proportional*: Probability of selection of individuals is based on their fitness; *Tournament selection*: By conducting tournament between individuals.

- *Crossover*: This operator acts on the parent individuals to generate children such that they have traits of both the parents. The choices available in this operator are: Single point cross over, Multi point cross over, and Uniform cross over
- *Mutation*: This operator randomly changes the characters of the offspring. The mutation depends on the encoding of chromosomes.

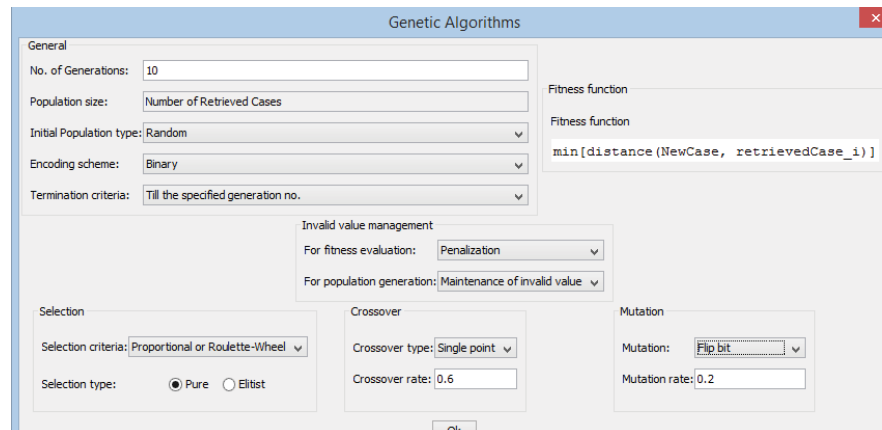


Figure 2. The GESCONDA interface of the Genetic Algorithm engine parameter selection

Apart from these choices, there are several extensions of genetic algorithms like: Elitism, Island model, Fitness sharing, etc.

All these characteristics of different genetic operators provide number of parameters to solve a problem. Deciding which parameters are the best depends on the problem to be solved. Thus, the implementation of all these choices produces a true generic purpose genetic algorithm engine. In figure 2 there is the interface for selecting the parameters of the general GA engine.

4.3 The interoperability among both models

Case-Based Reasoning computes the new solution for a new problem (case) based in the adaptation of the older solution/s of similar problem/s. This step, as well as the retrieval step is very important. It depends on the domain knowledge available to propose adaptation strategies, but some general procedures can be used. These procedures are the copy solution, mean strategy, weighted mean strategy or the use of some mathematical formula introduced by the user/expert. All these general procedures were implemented in GESCONDA. Each one of the solution variables of a case could be adapted according to a different adaptation scheme. Thus, now one solution variable could be adapted using a new method: a genetic algorithm. Indeed, the genetic algorithm will provide a value for all the solution variables. Figure 3 shows this interaction between both methods in GESCONDA.

The rationale is the following: given a new problem to be solved (i.e., the solution variables are unknown), the case base is retrieved to search for the most k similar cases to the new case (C_{new}) in the system. Those cases will form the initial population for the genetic algorithm, which will optimize the following fitness function:

$$FitnessFunction(C_i) = Sim(C_i, C_{new})$$

Were $C_i, i=1, \dots, k$, are the most similar cases to the current new case (C_{new}) and are the individuals belonging to the corresponding population of the genetic algorithm. Thus, the genetic algorithm searches for the optimal individual maximizing the *FitnessFunction*. This way, the genetic algorithm will converge to the best individual (most similar) to the new case we want to solve, which has the solution variables known. Thus, those variables of the best individual will be the solution values for the required unknown solution variables, and this way the genetic algorithm has been used to adapt the solution variables of the new case.

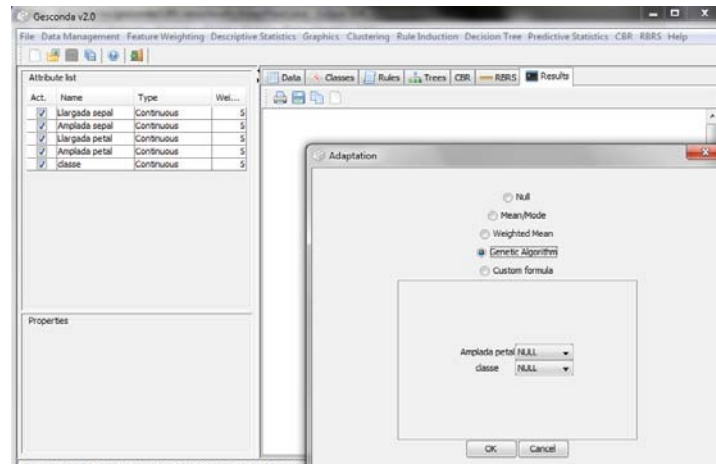


Figure 3. Updated Interface of the Case-Based Reasoning Module

5. RESULTS IN A CASE STUDY

This section describes a case study of the application of the GESCONDA tool to assess the usefulness and the results obtained using the combined interoperation of both a case-based reasoning model and a genetic algorithm model in the management of a wastewater treatment plant.

An Activated Sludge Wastewater Treatment Plant (WWTP) is a complex process involving several chemical, biological, physical, mechanical, electrical features. The quality of the effluent water must be always maintained in a good condition to minimise any environmental impact. In the WWTP operation, problems frequently appear such as solids separation problems, biological foam or bulking episodes in the bioreactors, or overloading derived from storms and heavy rains.

Dissolved Oxygen in the bioreactor (DO in mg/l), the Recirculation Flow (RF in m³/day) and the Waste Flow (WF in m³/day) are 3 of the most common control variables used in WWTPs to try to lead the process to normal/good operation. Real data available from a WWTP was used for the experimentation. The data contained 253 instances with 15 variables, including the three control variables described before. Twelve variables described the process: Inflow (I in m³/day), Chemical Oxygen Demand (COD in mg/l) and Suspended Solids (SS in mg/l) measured at the inflow; after the primary treatment; and at the outflow, Suspended Solids Liquor Mixture (SSLM in mg/l), Sludge Volumetric Index (SVI), Retention Time (RT in days), Food/Mass ratio (F:M), a Diagnostic label (Diag). The validation of the approaches was carried through a leave-one-out validation scheme, which means that all the instances less one are used as the Case Base or the whole population, and the remaining one as the test case. The final values described in table 1 are the average values across all the 252 executions. Three main strategies for the management of the WWTP were tested:

- Using only the case-based model. The three parameters were predicted using null option ($k=1$), mean adaptation or weighted mean adaptation scheme over the k ($k=3$, $k=5$, $k=10$, $k=20$) most similar cases. The customized formula option could not be used, because this knowledge encoded in formulas was not available. The similarity measure used was the Euclidian one, because all descriptive variables were numerical.
- Using only the genetic algorithm model. The three solution variables were computed as the minimization process of the Fitness function mentioned in the preceding section. In this approach, the initial population was the whole dataset of cases. Several trials with different number of generations (10, 20) and selection rate values (0.2, 0.6, 0.8) were used. Mutation rate was set to 0.2. The similarity measure used was also the Euclidian.
- Using the interoperation of both models. The case-based reasoning model was used for retrieving, for each new test case, the most k ($k=1$, $k=3$, $k=5$, $k=10$, $k=20$) similar cases. These k most similar cases were used as the initial population of a genetic algorithm to minimize the Fitness function mentioned above. The similarity function used was the Euclidian measure too. Several parameters of the genetic algorithm were tested.

The objective was to adjust the three control variables (DO, RF, WF) to the suitable values to lead the WWTP to a normal operation state, or to optimize these values whenever the normal operation was met. The predicted values were compared with the available real data, and the Normalized Root Mean Square Error (NRMSE) was used. In addition, the percentage of instances with a Relative Error less than 10% were counted to give a complementary accuracy measure.

Strategy	WAS		RAS		DO	
	NRMSE	P10	NRMSE	P10	NRMSE	P10
CBR null (k=1)	0,1420	14,23	0,1252	47,43	0,1857	14,62
CBR mean k=2	0,1348	14,23	0,1040	43,87	0,1882	16,60
CBR mean k=3	0,1272	15,41	0,0984	41,89	0,1763	12,65
CBR mean k=5	0,1221	14,23	0,0950	40,71	0,1621	13,04
CBR mean k=10	0,1203	16,60	0,0920	39,13	0,1548	19,36
CBR mean k=20	0,1204	15,41	0,0885	37,15	0,1487	26,08
CBR Wmean k=3	0,1272	15,41	0,0984	41,89	0,1762	12,65
CBR Wmean k=4	0,1221	14,23	0,0950	40,71	0,1698	13,04
CBR Wmean k=10	0,1203	16,60	0,0921	39,13	0,1548	19,76
CBR Wmean k=20	0,1204	15,41	0,0885	37,15	0,1487	25,69
GA 10/0.2/0.2	0,1524	18,18	0,1638	31,62	0,2154	12,25
GA 10/0.6/0.2	0,1392	15,02	0,1645	34,78	0,2032	15,41
GA 10/0.8/0.2	0,1531	14,23	0,1402	38,34	0,2076	15,81
GA 20/0.2/0.2	0,1719	7,90	0,1088	37,15	0,1863	16,60
GA 20/0.6/0.2	0,1720	16,20	0,1400	38,73	0,1972	14,62
GA 20/0.8/0.2	0,1722	13,83	0,1270	30,83	0,2211	13,04
CBR-GA K=1/10/0,2/0,2	0,1420	14,23	0,1252	47,43	0,1857	14,62
CBR-GA K=3/10/0,2/0,2	0,1540	14,62	0,0889	44,66	0,2109	15,02
CBR-GA K=5/10/0,2/0,2	0,1537	15,41	0,1254	44,66	0,2210	18,58
CBR-GA K=10/10/0,2/0,2	0,1513	15,41	0,0899	43,48	0,2627	14,23
CBR-GA K=20/10/0,2/0,2	0,1553	12,25	0,1086	42,29	0,2183	16,21
CBR-GA K=1/10/0,6/0,2	0,1420	14,23	0,1252	47,03	0,1857	14,62
CBR-GA K=3/10/0,6/0,2	0,1421	11,46	0,0899	44,66	0,2020	13,44
CBR-GA K=5/10/0,6/0,2	0,1472	11,07	0,0889	46,64	0,2250	14,62
CBR-GA K=10/10/0,6/0,2	0,125	14,23	0,1085	47,82	0,2082	17,79
CBR-GA K=20/10/0,6/0,2	0,1398	11,86	0,0900	44,27	0,2138	15,02

Table 1. NRMSE values and percentages of instances with less than 10% of error in the experimentation. The best values are marked in bold. The parameters of GA are size of population(k)/number of generations/selection rate/mutation rate.

The CBR-GA approach predicts better the RAS variable. The DO variable is predicted better with the CBR approach, and the WAS is slightly better predicted by some GA approaches. Nevertheless, it can be observed that the GA approaches get higher NRMSE values in all three variables, probably due to its higher variability. From a computational perspective, the CBR-GA strategy increases the computation time of the CBR approach by a factor proportional to the product of the number of generations by the number of descriptive variables by the parameter *k*.

7. SUMMARY AND DISCUSSION

The paper shows that including interoperability skills in an intelligent decision support system permits the real interaction between the different components of the system and provides the possibility to exploit the whole potential of the system by increasing the performance of the decision support.

GESCONDA is an IEDSS tool including many different modelling components and some of them might communicate through different interoperability schemes. For those components that might interact in a sequential way and the output of one of them might be considered the input of another one, file-based gateways based on formal specification of the information transferred among models might be sufficient, like in *model producers*. This is the case of communicating a clustering model with a classifier model. Another kind of interoperation is between *model producers* and *model executors*, like a classification rule model with a rule-based engine model. Again, this interoperation could be done with file-based interchange protocols.

However, for some other components that might communicate in a non-sequential way when *model producers* are in execution, other interoperability skills might be required. Therefore, in this intra-task model producers' interoperation, other communication schemes should be needed. In this particular

paper, the contribution of a Genetic Algorithm engine to improve a Case-Based Reasoning paradigm is focused on, and specific modification of the GESCONDA software has been implemented to permit this interoperation. In the paper, the application of using this integrated collaborative scheme where Genetic Algorithms and Case Based Reasoning interoperate at runtime to provide decision support is tested over a real case study. The methodology proposed to use a genetic algorithm as an adaptation strategy for a case-based reasoning engine. This way the genetic algorithm explores and optimizes, among the population of most similar cases to the new cases to be solved, to find the solution variables. The case study has illustrated a real scenario to manage a WWTP. The preliminary evaluation of the approach has showed that the results were the best ones in one variable solution (RAS) with the interoperation and integration of both case-based reasoning and the genetic algorithm approach. Moreover, in the other two solution variables, the interoperation of CBR-GA improved the GA approach, but not the CBR approach.

REFERENCES

- Argent, R.M. (2004). An Overview of Model Integration for Environmental Applications-components, frameworks and semantics. *Environmental Modelling & Software* 19:219-234, 2004.
- Elasri, H. and Sekkaki, A. (2013). Semantic Integration process of Business Components to Support Information System Designers. *Int. Journal of Web & Semantic Technologies* 4(1):51-65.
- Gibert, K., Sánchez-Marrè, M. and Codina, V. (2010). Choosing the Right Data Mining Technique: Classification of Methods and Intelligent Recommendation. 5th International Congress on Environmental Modelling and Software. iEMSs' 2010 Proceedings, Vol 3., pp. 1940-1947.
- Gibert, K., Sánchez-Marrè, M. and Rodríguez-Roda, I. (2006) GESCONDA: An Intelligent Data Analysis System for Knowledge Discovery and Management in Environmental Data Bases. *Environmental Modelling & Software* 21(1):116-121.
- Goldberg, D.E. (1989). Genetic Algorithms in Search, Optimization & Machine Learning. Addison-Wesley.
- Holland, J. H. (1975). Adaptation in natural and artificial systems. The University of Michigan, 1975.
- Institute of Electrical and Electronics Engineers (1990). IEEE Standard Computer Dictionary: a Compilation of IEEE Standard Computer Glossaries. New York.
- Kolodner, J. (1993). Case-Based Reasoning. Morgan Kaufman, 1993.
- Kzaz, L., Elasri, H., and Sekkaki, A. (2010). A Model for Semantic Integration of Business Components. *Int. Journal of Computer Science & Information Technology* 2(1):1-12.
- Manguinhas, H. (2010). Achieving Semantic Interoperability using Model descriptions. *Bulletin of IEEE Technical Committee on Digital Libraries*, Vol. 6, No. 2, Fall 2010.
- Mackay, D. S. (1999). Semantic Integration of Environmental Models for Application to Global Information Systems and Decision-Making. *ACM SIGMOD Record* 28(1):13-19, March 1999.
- Ouksel, A.M. and Sheth, A. (1999). Semantic Interoperability in Global Information Systems: a Brief Introduction to the Research Area and the Special Section. *ACM SIGMOD Record* 28(1):5-12.
- Rizzoli, A. E., Davis, J.R. and Abel, D.J. (1998). Model and Data Integration and re-use in Environmental Decision Support Systems. *Decision Support Systems* 24:127-144, 1998.
- Sánchez-Marrè, M. (2014) Interoperable Intelligent Environmental Decision Support Systems: a Framework Proposal. Submitted to 9th International Congress on Environmental Modelling and Software (iEMSs'2014). San Diego, CA, USA. June 2014.
- Sánchez-Marrè, M., Gibert, K. and Sevilla, B. (2010). Evolving GESCONDA to an Intelligent Decision Support Tool. 5th International Congress on Environmental Modelling and Software (iEMSs'2010). iEMSs' 2010 Proceedings, Vol. 3, pp. 2015-2024.
- Sevilla, B. and Sánchez-Marrè, M. (2010). Providing Intelligent Decision Support Systems with Flexible Data-Intensive Case-Based Reasoning. 5th International Congress on Environmental Modelling and Software (iEMSs'2010). iEMSs' 2010 Proceedings, Vol. 3, pp. 2063-2072.
- Sottara, D., Bragaglia, S., Mello, P., Pulcini, D., Luccarini, L and Giunchi, D. (2012). Ontologies, Rules, Workflow and Predictive Models: Knowledge Assets for an EDSS. 6th International Congress on Environmental Modelling and Software. iEMSs' 2012 Proceedings, pp. 204-211.
- Vetere, G. and Lenzerini, M. (2005). Models for Semantic Interoperability in Service-oriented architectures. *IBM Systems Journal* 44(4):887-903, 2005.
- Wesseling, C.G., D. Karssenberg, P.A. Burrough, and W.P.A. Van Deursen. (1996). Integrating dynamic environmental models in GIS: the development of a dynamic modelling language. *Transactions in GIS*, 1(1), 40-48.