OntoWEDSS: an ontology-underpinned decision-support system for wastewater management

Luigi Ceccaroni[a*], **Ulises Cortés** [b] and **Miquel Sanchez-Marre**[b]

[a]*Fujitsu Laboratories of America*
595 Lawrence expressway, Sunnyvale, CA 94085, USA
lc@ a.fujitsu.com

[b]*Departament de Llenguatges i Sistemes Informatics, Universitat Politecnica de Catalunya*
Modul C5, Campus Nord, C/ Jordi Girona 1-3, 08034 Barcelona, Catalunya, Spain
{ia, miquel}@lsi.upc.es

Abstract: This paper characterizes part of an interdisciplinary research on artificial intelligence techniques applied to environmental decision-support systems. The architecture's design of OntoWEDSS, a decision support system for wastewater management, is presented. This system augments classic rule-based reasoning and case-based reasoning with a domain ontology. The integration of the newly created WaWO ontology provides a more flexible management capability to OntoWEDSS. The construction of the decision support system is based on a specific case study but the system is also of general interest, given that its ontology-underpinned architecture can be applied to any wastewater treatment plant and, at an appropriate level of abstraction, to other environmental domains. The OntoWEDSS system helps improve the diagnosis of faulty states of a treatment plant, it provides support for wastewater-related complex problem-solving, and it facilitates knowledge modeling and reuse by means of the WaWO ontology. In particular, the following issues are dealt with: (1) the improvement of the modeling of information about wastewater treatment processes and the clarification of a part of the existing terminological confusion in the domain, (2) the incorporation of ontology-modeled microbiologic knowledge related to the treatment process into the reasoning process and (3) the creation of a decision support system with three layers (perception, diagnosis and decision support) which combines knowledge through a novel integration between KBSs and ontologies, thus providing better results.

Keywords: ontologies; environmental modeling; artificial intelligence.

1 INTRODUCTION

This paper presents the design of an *ontology*-based environmental decision-support system (named OntoWEDSS[2]) applied to the domain of wastewater treatment. This is a *new and interdisciplinary approach* to the management of knowledge in the problem-solving processes related to environmental issues. In fact, even if the application studied is specific, the architecture presented could serve as a basis for any environmental system.

1.1 Environmental Decision-Support Systems

Environmental Decision-Support Systems (EDSSs) are useful when dealing with complex environmental processes, which are not easily modeled because the knowledge is still incomplete and uncertain. *Rule-based expert systems* (RBES), *case-based reasoning systems* (CBRS) and ontologies have proven, individually, to be able to cope with some known difficulties and to have successfully faced several problems related to the wastewater domain. The synthesis of these different modeling and reasoning systems can result in great improvements in decision support. The introduction of an ontological component in an EDSS, in particular, allows to develop issues which will contribute to the improvement of the current state of the art in wastewater

---

[*] Luigi Ceccaroni carried out the most part of the research related to this paper at Universitat Politècnica de Catalunya.

[2] Ontology-based Wastewater Environmental Decision-Support System

management and in particular to: (1) a more stable wastewater treatment operation through ontology-based supervision and (2) the *portability* of the management system of a *wastewater treatment plant* (hereafter, WWTP).

Wastewater treatment plants are the physical element of the domain modeled by the ontology and managed by OntoWEDSS. They serve to decontaminate wastewaters prior to their discharge into a natural body of water. For that, they use techniques of physical, chemical and biological treatment. The wastewater-treatment process is very *complex* and it is difficult to develop a reliable supervisory technology based only on a classic chemical-engineering control approach.

1.2 Ontologies

Ontologies are being developed in AI to facilitate knowledge sharing and reuse. With respect to the research involved in this study, ontologies can provide: (1) a shared and common understanding of the knowledge domain that can be communicated among agents and application systems, and (2) an explicit conceptualization that describes the semantics of the data (Fensel et al. [2000]). Ontologies are considered to be critical in allowing software programs to communicate among themselves in meaningful ways, and attract attention not only from academic disciplines such as computer science, information science and artificial intelligence, but also from industries as diverse as the high-tech, financial, medical, educational and environmental sectors. A recent comprehensive document covering the main aspects of ontologies in AI research is the technical roadmap of the ontology field in Europe and worldwide produced by the OntoWeb project[3]. In perspective, on one hand ontologies represent a first step on the way for *real portability* of a system towards other similar domains and they could be effectively employed to address the problem of general model-construction (*generalization*); while on the other hand it is possible to instantiate/adapt an ontology to the specific configuration of a WWTP and to automatically construct and validate specific models (*specification*).

2 ONTOWEDSS

OntoWEDSS is a research tool built to explore the possibilities and the potential of introducing ontologies into decision support systems, using an environmental domain as case study. In this section, OntoWEDSS is described through its profile, model and grounding.

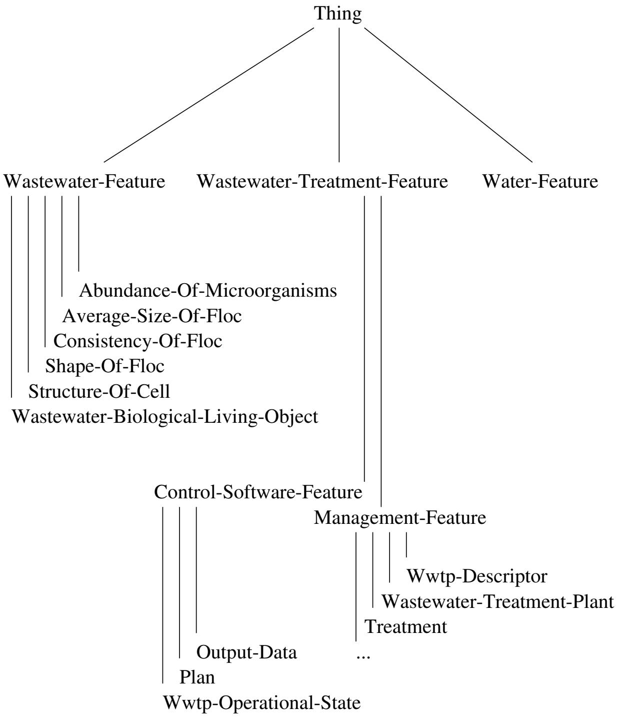


Figure 1: Top-level categories in the October 2001 version of the WaWO ontology.

2.1 Profile

The *profile* of OntoWEDSS describes what it does at a high level and what it requires: hardware, software, knowledge, input types.

Description. The architecture of OntoWEDSS integrates various kinds of data and several AI techniques (including an ontology). Given an adequate amount and type of data, it is flexible enough to deal with the complexity of the wastewater treatment process. In OntoWEDSS, the domain is represented in detail and the evolution of micro-organism communities (a key element in the biological treatment process) is taken into account. With OntoWEDSS it is possible to capture, understand and describe the knowledge about the whole physical, chemical and microbiological environment of a WWTP. OntoWEDSS makes use of an ontology in domain modeling and for the clarification of existing terminological confusion in the wastewater domain. It automatically, reliably discovers and manages problematic states in the wastewater treatment domain. OntoWEDSS reuses and composes different reasoning systems (rule-based, case-based and ontology-

---

[3] Deliverable 1.1.1 of IST Project IST-2000-29243 OntoWeb, http://www.ontoweb.org/deliverable.htm.

based), which interoperate among themselves. It supervises the process through a management distributed in *3 layers*: perception, diagnosis and decision support. OntoWEDSS incorporates wastewater microbiological knowledge into the reasoning process and represents cause-effect relations. Finally, it resolves existing *reasoning-impasses*, such as lack of diagnosis.

The OntoWEDSS system uses its internal knowledge-bases and inference mechanisms to process information about a WWTP. It diagnoses the ongoing state of the treatment plant and predicts the evolution of that state. Eventually, the output of the system is represented by statements about actions to be taken or statements to support a human manager's decisions, in order to maintain the correct operation of the plant. OntoWEDSS uses the *WaWO ontology*[4] (Ceccaroni et al. [2000]), which has been designed and built following current mainstream ideas about ontology construction. WaWO is a hierarchically-structured set of terms and relations describing the domain of wastewater treatment. It is the manifestation of a shared understanding of the wastewater domain that is agreed among a number of experts in environmental and chemical engineering. The introduction of this agreed-upon ontology in the domain of wastewater treatment facilitates: (1) an accurate, effective *communication and sharing* of meanings, which leads to benefits such as knowledge reuse, (2) an advancement in environmental technologies for the management of biological and biochemical processes, and (3) an enhancement of the knowledge about the specific microbial ecology of environmental processes developing in treatment plants. Even though WaWO was designed on the basis of the specification of a few particular plants, the knowledge which it embodies is valid for any treatment plant of the same class.

Functionalities. The *input* (from the user or a file) for modeling and execution in OntoWEDSS is represented by (1) the list of descriptors to use and (2) the descriptor values of a new-problem. The user can use a predefined set of descriptors and can define new ones (in this second case, however, only case-based reasoning is readily available for diagnosis). Optionally, the weight of the descriptors can be provided. What follows is an *example* of input descriptors[5] together with their range of possible values: *foam at aeration tank* (none, very little, some, abundant, very abundant), *floc situation at clarifier* (deflocculation, floating sludge, denitrification), *settler qualitative assessment* (bad, good, very good).

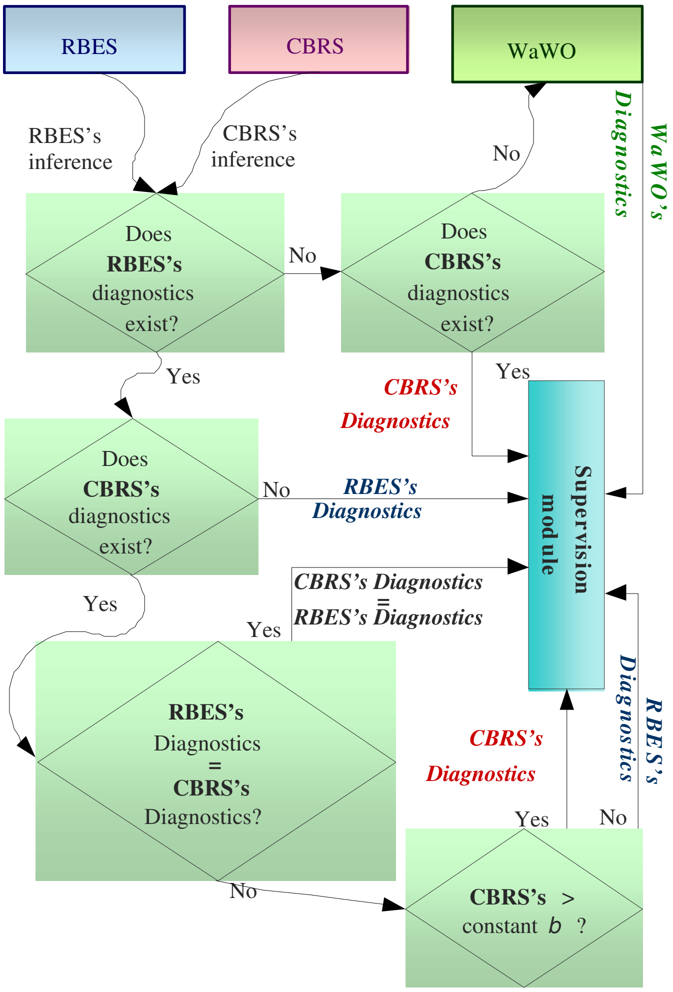


Figure 2: Data flow of general diagnosis-integration.

The *output* (to the user) of OntoWEDSS execution is represented by (1) a diagnosis of the current state of the WWTP (with a reliability factor), (2) a trace of the reasoning carried out, and (3) a list of actions to be taken according to the situation diagnosed.

Functional parameters. The suggested activation cycle for OntoWEDSS is 1 hour (5 min in the case an emergency is detected). The accuracy in the classification of the current state of a WWTP (based on focused evaluation) ranges between 73% and 100%. The cost of this kind of system is related to the programming language used for implementation (Allegro Common LISP 5.0.1 in our case).

---

[4] *WasteWater Ontology* (WaWO) availability: http://www-ksl-svc.stanford.edu and http://babylon.com.

[5] OntoWEDSS uses dozens of physical, chemical and biological descriptors, such as presence of various microorganisms (e.g., Acineria uncinata, Acineta spp., Aspidisca cicada), biochemical oxygen-demand (at effluent, inflow and primary-effluent), and BOD/N ratio, just to name a few (see also Table 1).

2.2 Model

The *model* of OntoWEDSS is how it works: the internal architecture, what happens when it is executed and how to monitor the execution.

The architecture of the system (Ceccaroni [2001], pp. 111) has a modular design, to improve understandability, reliability and, above all, the ability to make modifications. Many distinct, specialized subsystems (such as perception, rule-based reasoning, case-based reasoning, ontology, modeling, planning and execution) are defined and then grouped into three layers, whose details are as follows: *perception* (data gathering and knowledge acquisition), *diagnosis* (reasoning and learning), *decision support* (prediction, evaluation of alternative scenarios, advising, actuation and supervision).

Perception. During the perception phase, all available information is gathered: this includes quantitative and qualitative descriptors (see Table 1).

Table 1: Some of the descriptors used in WWTPs.

Type	Descriptor	Sampling rate
On-line	dissolved oxygen, pH, water and sludge flow rates	seconds
Off-line	COD, BOD, TSS, TKN, NO_2^-, NO_3^-, PO_4^{3-}, Cl^-, conductivity, turbidity, MLSS, T, NH_4^+, V30, in situ observations.	daily
	floc characterization, protozoa, metazoans, filamentous bacteria	weekly
Calculated	COD and TSS removal SRT, HRT, F/M, SVI	daily

Diagnosis. The diagnosis phase is characterized by a chicken-and-egg paradox related to process-modeling. The situations (set of descriptors' values) cannot be defined without first knowing what diagnostics they correspond to. And most diagnostics can be hard to define as such until the corresponding situations have been identified. To overcome this problem, experts often have to use trial-and-error methods. Once the process is modeled, three modules, covering *rule-based reasoning*, *case-based reasoning* and the *ontology*, are used for diagnosis. With respect to the rule system, it is designed to be implemented in *two separate layers*. A more general one, which can be reused across WWTPs, and a more specific one to be used only

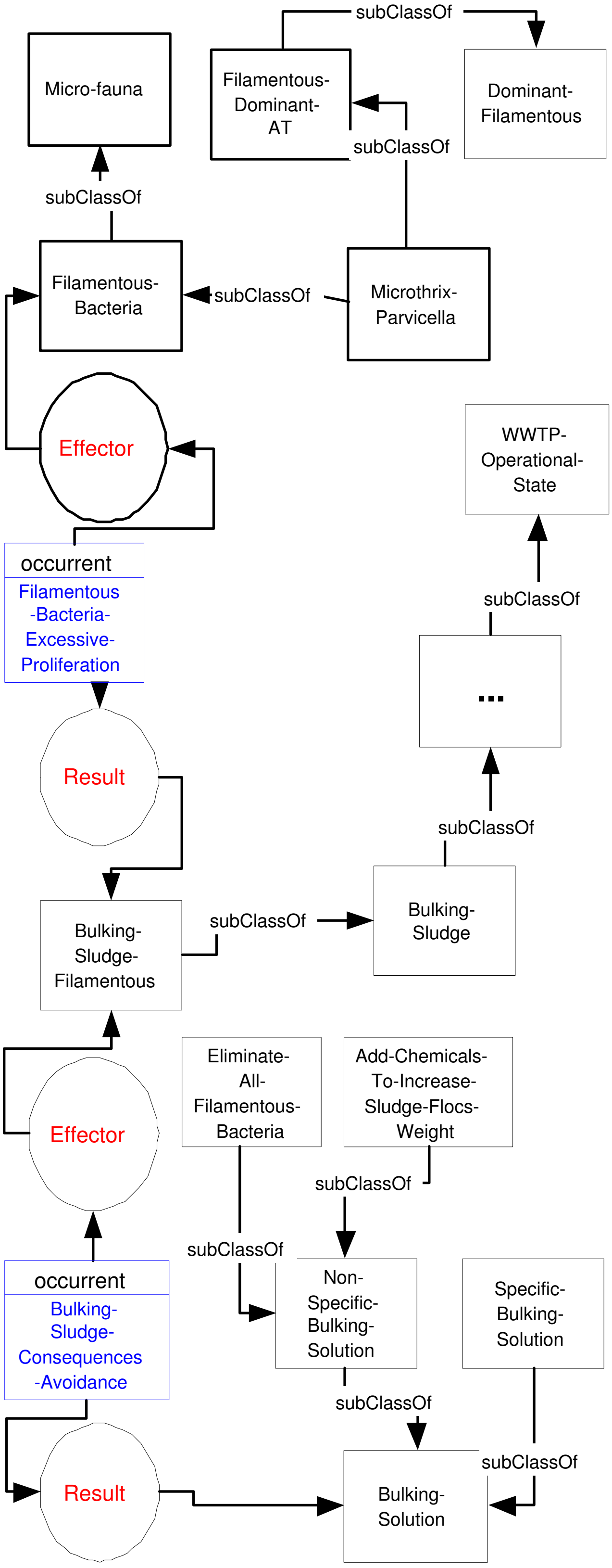
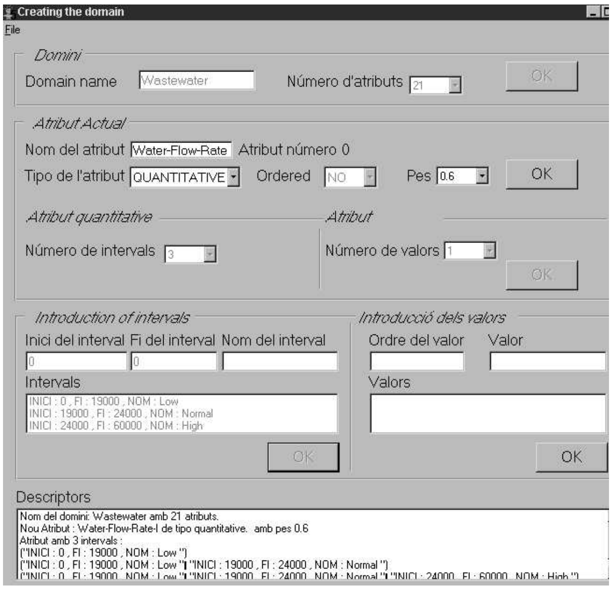


Figure 3: Reasoning with the ontology.

in a particular WWTP. The essential addition of an *ontology* in the diagnosis layer helps to model the wastewater treatment process (see Figure 1), paying a special attention to the management of the qualitative knowledge, that is, the environmental information on micro-organisms presence. Within the layer, the RBES and the CBRS work independently and they both produce as output a diagnostics about the state of the plant. This output is passed to a diagnosis integrator, which shows the two outputs (together with an associated confidence value) to the user and then starts the reasoning schema illustrated in Figure 2. If the diagnostics of the two KBSs is the same, this result is communicated to the decision support layer. If the diagnostics exist and are different, the system prioritizes as follows. If the case similarity is higher than a predefined threshold (b), the case-based reasoner's diagnostics prevails. Otherwise, the rule-based expert system's diagnostics prevails. In case of impasse (no diagnosis), OntoWEDSS turns first to the ontology and then, if it fails, to the plant manager, demanding a diagnosis based on their microbiological deep knowledge. This external solution is always learned.

The WaWO ontology used is specialized on the wastewater domain. In WaWO, for example, *Storm* is an *Operational-Problem*, *Bacterium* is a *Wastewater-Biological-Living-Object*, and the only *Metazoans* represented are *Nematode* and *Rotifer*. The use of the ontology for reasoning is completely experimental and there are no other documented studies in a similar context. The hierarchical organization of categories of the WaWO ontology is expressed in the Ontolingua knowledge-representation language, and KIF axioms are used for answering queries, language analysis and general reasoning. The axioms of the ontology have been partially modeled, but not implemented in KIF, so that most of the reasoning is simulated.

In Figure 3, an example of the reasoning with the ontology is depicted. It can be partially read, using the terminology of Sowa [2000], as a sequence of occurrences. Simple rectangles are **role categories** (or phenomenon categories, or classes) and are always part of concept hierarchies; circles are **relations**. In the example, *Filamentous-Bacteria* is what causes (i.e., is the effector of) the *Filamentous-Bacteria-Excessive-Proliferation* **occurrent**. Being the effector part of taxonomic (*Micro-fauna* branch) and operational (*Microthrix-Parvicella* branch) semantic structures, the occurrent can then be linked both to a class of micro-organisms and to the *Filamentous-Dominant-AT* state of the WWTP. The result of the occurrent is the *Bulking-Sludge-Filamentous* situation, which is itself the effector of the *Bulking-Sludge-Consequences-Avoidance* occurrent. The result of this last occurrent is the search for a *Bulking-Solution* within its concept hierarchy. The axioms (which are essentially rules) guide the search towards a specific or a non specific solution, according to the original effector. The final result can be, for instance, that the presence of a certain amount of filamentous bacteria triggers the addition of certain chemicals to the wastewater. Even if it is true that the same result can be achieved through a complex RBES, the very structure of the ontology guarantees much greater readability, easier consistency-checking, and reuse of knowledge in a way that is becoming a more and more recognized standard. Furthermore, the ontological knowledge representation is very Web oriented and could be useful in a future scenario of knowledge sharing over the Internet.



Domain name: Wastewater
Número d'atributs: 21
Nom del atribut: Water-Flow-Rate
Atribut número 0
Tipo de l'atribut: QUANTITATIVE
Ordered: NO
Pes: 0.6
Número de intervals: 3
Número de valors: 1
Inici del interval: 0
Fi del interval: 0
Nom del interval:

Intervals

```
INICI : 0 , FI : 19000 , NOM : Low
INICI : 19000 , FI : 24000 , NOM : Normal
INICI : 24000 , FI : 60000 , NOM : High
```

Descriptors

```
Nom del domini: Wastewater amb 21 atributs.
Nou Atribut : Water-Flow-Rate-I de tipo quantitative.  amb pes 0.6
Atribut amb 3 intervals :
("INICI : 0 , FI : 19000 , NOM : Low ")
("INICI : 0 , FI : 19000 , NOM : Low "|"INICI : 19000 , FI : 24000 , NOM : Normal ")
("INICI : 0 , FI : 19000 , NOM : Low "|"INICI : 19000 , FI : 24000 , NOM : Normal "|"INICI : 24000 , FI : 60000 , NOM : High ")
```

Figure 4: Window for the creation of a domain in the CBRS of OntoWEDSS.

Decision support. This layer exploits available data and information to provide, through a friendly interface, active decision-support about the following key actuations in the WWTP: (1) execution monitoring of RBES and CBRS, and (2) concrete action suggestions, such as "*Change Sludge-Recirculation-External to 120*" or "*Destruction of filaments via chlorine addition*" or "*Addition of inorganic coagulant*".

2.3 Grounding

The *grounding* of OntoWEDSS is how the user accesses the system, that is, the interface for data exchange. It allows the user to communicate with the RBES and CBRS in a friendly way, via the selection of options from menus and buttons. The grounding has three main functionalities: (1) the introduction of the data of the problem in question (see Figure 4 for the case of CBRS), (2) consulting the RBES about the state of facts or rules, and (3) asking the user to confirm actions or about data values.

2.4 Implementation and performance

OntoWEDSS has been fully implemented in LISP, but not deployed in a full-scale facility. However, its application and evaluation are based on data coming from full-scale facilities. The ontology, in its current state, is fully reusable if transferred to any real facility. The same is true for the CBRS, while in the case of the RBES only the rules of the upper layer (see *Diagnosis* in section 2.2) can be reused without modification. The *evaluation* of the system was focussed on the most representative problematic situation that is possible to come upon in wastewater treatment: the presence of bulking sludge due to filamentous micro-organisms. The objective of the evaluation was to quantify the performance of the various paradigms and of the whole system when they react to a specific problem. In the three experiments carried out, the successful diagnosis coming from the system without the ontology ranges from 60% to 73% and impasse situations correspond to a set of 10 instances (out of 57). When the system operates with the WaWO ontology, these results improve. The advance in diagnosis is due to the following two circumstances: first, WaWO activates when an impasse situation has been reached and for this reason it includes more flexible axioms with respect to the RBES; second, WaWO has usually at its disposal additional information about micro-organisms that was not used in the earlier evaluation because the RBES and the CBRS are not able to deal with it. The final successful diagnosis coming from the system with the ontology ranges from 73% to 100%. We acknowledge that, in one of the experiments, the ontology could not improve the performance, which was 73%.

3 CONCLUSIONS

In the paper, we introduced the use of ontologies for the solution of complex problems related to environmental science and engineering. In particular, we studied the integration of an ontology with case-based reasoning and rule-based reasoning into an environmental decision-support system. This integration improved the modeling of the information about wastewater treatment processes and resolved existing *impasses* in the reasoning cycle. Specifically, we presented an ontological representation of two kinds of cause-effect relations: *micro-organisms $\leftrightarrow$ problematic situations* and *state of the plant $\leftrightarrow$ suggested actions*. We also used the ontology to improve the communication among different elements and agents of an environmental decision-support system, thus reducing ambiguities. The implementation of the system and an evaluation of the advantages related to the proposed approach are described in detail in Ceccaroni [2001].

ACKNOWLEDGMENTS

This research is supported by ATEAM (IST-1999-10176) and Agentcities (IST-2000-28384) European projects, and the CICyT TIC-2000-1011 Spanish project. We wish to thank Jonathan Dale for the feedback after having read a preliminary version of this paper.

REFERENCES

Ceccaroni, L. *OntoWEDSS - An Ontology-based Environmental Decision-Support System for the management of Wastewater treatment plants*. PhD thesis, Universitat Politècnica de Catalunya, http://www.lsi.upc.es/~luigic/thesis.pdf, 2001.

Ceccaroni, L., U. Cortés, and M. Sànchez-Marrè. WaWO - An ontology embedded into an environmental decision-support system for wastewater treatment plant management. In *Proceedings of ECAI2000 - W09: Applications of ontologies and problem-solving methods*, pages 2.1–2.9, Berlin, Germany, 2000.

Dieng, R. et al., editors. *Knowledge Acquisition, Modeling, and Management, Proceedings of the European Knowledge Acquisition Conference (EKAW2000)*. Lecture Notes in Artificial Intelligence (LNAI). Springer-Verlag, 2000.

Fensel, D., I. Horrocks, F. Van Harmelen, S. Decker, M. Erdmann, and M. Klein. *OIL in a nutshell*, pages 1–16. 2000. In Dieng et al. [2000].

Sowa, J. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole, 2000.