# WaWO - An ontology embedded into an environmental decision-support system for wastewater treatment plant management

**Luigi Ceccaroni, Ulises Cortés, Miquel Sànchez-Marrè**[1]

**Abstract.** We present an *ontology* (named WaWO - Waste Water Ontology), integrated into an existent environmental decision-support system (named DAI-DEPUR+), designed to model the information about a wastewater treatment process through the definition of the basic terms and relations comprising the vocabulary of the wastewater treatment area. The WaWO ontology is built following the ideas of Uschold and Grüninger [14] [13] and is a hierarchically structured set of terms and a set of axioms for 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 agents: mainly, experts in environmental and chemical engineering. This ontology incorporates some new features for the wastewater domain, such as: semi-automatic generation of knowledge bases, modeling of the cause-effect relations existing within a wastewater treatment plant, integration with predictive reasoning based on transition networks.

## 1 INTRODUCTION

The general issues we would like to address in the paper are:

* optimizing wastewater treatment operation by a *more reliable management* and an automatic control, and
* making easier the *portability*[2] of the management system.

### 1.1 Environments

Contamination levels of waters constantly increase due basically to industrial development and to the increase of population density in certain zones. Wastewaters, either industrial or urban, have to be decontaminated until an adequate level, so that they could be poured to the surrounding hydric medium without causing problems of environmental deterioration. For that, there are *wastewater treatment plants* (WWTPs) with techniques of physical-chemical and biological treatment.

Wastewater treatment plants with biological processes are the physical environment modeled by our ontology.

The DAI-DEPUR+ system, in which the WaWO ontology is being embedded, receives on-line inputs[3] from sensors all over the WWTP as well as off-line inputs from the WWTP laboratories and human operators. The system uses its internal knowledge-bases[4] and inference[5] mechanisms to process and understand this information, to diagnose the ongoing WWTP-state, and to predict the evolution of that WWTP state.

In case of diagnosis impasse, DAI-DEPUR+, before turning to the plant manager, will try to solve the problem exploiting the connection, in the WaWO ontology, between data and states of the WWTP.

### 1.2 Inspirations and ambitions

Rule-based *knowledge-based systems* (KBSs) (one of the broadly applied paradigms of AI) proved able to cope with some known difficulties and to face several WWTP-domain problems, even if they are not the definitive solution to the treatment problem as a whole. On the other hand, large, multifunctional, available ontologies would significantly improve current KBSs and tutoring systems because they contain the broad knowledge of a domain required to perform multiple tasks and to explain domain knowledge from multiple viewpoints. A lot of ontologies are nowadays being built in many research centers around the world, but a few of them are specialized in biology or ecology, let alone wastewater management and WWTP microbiology.

A great improvement in addressing this kind of problems can come from the *integration* of different modeling and reasoning systems, such as specific *ontologies*, *rule-based reasoning*, *case-based reasoning* and *reactive planning*. The architecture we adopt (DAI-DEPUR+, which derives from the DAI-DEPUR system [11]) integrates two knowledge-based systems and an ontology, and is flexible enough to deal with the complexity of the wastewater treatment process, given an adequate amount and kind of data. In the DAI-DEPUR+ architecture, with the embedded ontology for the wastewater treatment, the representation of a deeper knowledge of the domain is permitted and the evolution of WWTP-microorganism communities can be taken into account. In this way, the management of biological problems arising in a treatment plant is more effective.

For the first time, by the creation of this ontology and its integration with the two KBSs, it will be possible to capture, understand and

[1] Universitat Politècnica de Catalunya, Departament de Llenguatges i Sistemes Informàtics, Campus Nord, Modul C6, Jordi Girona 3, 08034 Barcelona, Spain, email: {luigic , ia , miquel @lsi.upc.es}

[2] For a software system, 'portability' means the possibility to run once transported to another software environment.

[3] Input/output devices are any of various devices used to enter information and instructions into the DAI-DEPUR+ system for storage or processing and to deliver the processed data to a human operator or, in some cases, a machine controlled by the system. Such devices comprise sensors and effectors.

[4] Both static (rule-based), dynamic (case-based) and ontological knowledge-bases.

[5] Inference is the process of drawing conclusions about a particular parameter of the domain.

describe the knowledge about the whole physical, chemical and microbiological environment of a wastewater treatment plant. The basic terms and axioms of the ontology will entail a model of wastewater domain and a classification of the microorganisms according to known biological taxonomy, and will include a complete description of the microorganisms themselves (physical aspect, abundance and behavior in the treatment plant). The relations of the ontology will include, besides the classic hierarchical affiliations, all the interesting bindings among microorganisms and between them and the state of the plant (potential diagnosis).

In perspective,

- by one side the ontology will represent a first step on the path for a real portability of the system towards other similar domains; it could be effectively employed to address the problem of general model-construction in domains close to the one of wastewater (generalization);
- while on the other side it will be possible to instantiate / adapt the ontology to the specific configuration of a treatment plant and it will be useful to automatically construct and validate WWTP specific models (specification).

### 1.3 General overview

In this paper we present the WaWO ontology for wastewater management, we describe its structure in some detail, discuss the approach it takes to some important issues in domain modeling and show how it handles a variety of example problems. This ontology belongs to a more general decision support system for the supervision of wastewater treatment plants, which, in turn, is part of the knowledge and technology needed for the rational management of water resources.

We start by describing related work on ontologies in general (section 2). In section 3 we briefly present our environmental domain of study. Section 4 introduces the WaWO ontology; its architecture and features are considered, as are the associated problems, and the functioning of its components is discussed. Section 5 eventually explains the contributions of this work to AI research.

## 2 ONTOLOGIES

### 2.1 AI definitions

AI literature is full of different definitions of the term *ontology*. Each community seems to adopt its own interpretation according to the use and purposes that the ontologies are intended to serve within that community.

- One of the early definitions: 'An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary.' [10]
- A widely used definition (Gruber): 'An ontology is an explicit specification of a conceptualization.' [6]
- An elaboration of Gruber's definition: 'Ontologies are defined as a formal specification of a shared conceptualization.' [2]

### 2.2 Ontology development and ontology editors

Even if it is now widely recognized that constructing ontologies, or domain models, is an important step in the development of KBSs, what is lacking is a clear understanding of *how* to build ontologies. However, there exists a small but growing number of methodologies that specifically address the issue of the development and maintenance of ontologies. Ontolingua is the ontology development environment most related to the ontology of the DAI-DEPUR+ system. For a comprehensive survey of the work which has been done so far on methodologies see [7] [5].

The Ontolingua language, which is used by our WaWO ontology, is expressive, declarative, portable, domain independent and semantically well defined.

There are a number of more or less generic editors to create and manage ontologies, but the Stanford Ontolingua Ontology Editor (Stanford KSL Network Services:
http://www-ksl-svc.stanford.edu:5915/&service=frame-editor) is the most standard editor to create ontologies.

### 2.3 Ontologies and the environment

Environmental ontologies are just instantiations of the general ontology concept which assist in understanding a domain related to the natural environment and in modeling the processes involved. No ontology application exists yet in the field of WWTPs and no ontology modeling the evolution of microbiological systems has been defined. We think that the representational power of ontologies can be exploited to deepen the knowledge about the microorganisms of WWTP activated-sludge and the wastewater domain in general, and can be integrated together with other reasoning methods to better the whole supervision of WWTPs.

## 3 ENVIRONMENTAL DOMAIN OF STUDY

### 3.1 Wastewater treatment plants

In a wastewater treatment plant (WWTP), the main goal is to reduce the level of pollution of the inflow water, that is to remove, within certain limits (depending on local legislation), abnormal amounts of pollutants in the water prior to its discharge to the natural environment. This can be done in a number of different ways, corresponding to different kinds of WWTP. The most widespread classes of WWTP are:

- plants with only physical-chemical treatment;
- plants with additional biological reactor (for better organic matter removal), which can be of two main sub-type, depending on the sort of growth of microorganisms [1]:
  - suspended growth: with the microorganisms mixed with the wastewater and dispersed in the form of free cells or of bioflocks (activated sludge reactors).
  - attached growth: with the microorganisms anchored, in the form of biofilm, to inert surfaces (biological-film reactors).

This paper focuses on WWTPs with activated sludge, which is now the most common case in the European Union. Such wastewater treatment process schematically consists of:

- Preliminary treatment (pre-treatment): the wastewater flows into the WWTP and undergoes a screening process (e.g. grit removal of voluminous solids), chemical additions, (pre)aeration, odor control and flow measurement.
- Primary treatment: in a settler, readily settleable solids sediment and are taken apart in the form of sludge (which exits the system); in this way part of the organic matter associated with these solids is also removed.

- Secondary treatment: the wastewater flows into an aerated tank where biological treatment takes place; the microorganisms of the activated sludge come into contact with the organic matter dissolved in the wastewater and use it to develop; new activated sludge forms and it is poured into a secondary settler (clarifier), where suspended solids sediment and are separated from the water, which now is clean and can be released to the external environment.
- Sludge cycle: part of the concentrated sludge (RAS, Recirculated Activated Sludge) settled in the clarifier is recycled back to the aeration tank to preserve an optimum concentration of microorganisms there, while the rest of the sludge (WAS, Waste Activated Sludge) exits the system as waste.

### 3.2 WWTP management

With activated-sludge treatment plants it is normally possible to meet the general requirements regarding the final quality the water should have. In many cases, however, WWTP standard management wastes resources and energy, and is not cost effective in reaching acceptable effluent quality levels [15]. Moreover, in case of exceptional events (e.g., a storm or a heat wave), it can be very difficult to maintain the quality of the effluent water.

That is why the application of AI techniques and advanced water treatment technology coupled with on-line quality control techniques in wastewater management is needed. The integration of different AI models could provide an innovative way to:

- improve the efficiency of the treatment process;
- eliminate chemical pollutants and biological threats;
- reduce energy and chemicals consumption;
- reduce the formation of disinfection by-products.

Local legislation usually regulates all kind of dumping (including the quality of the effluent from wastewater treatment facilities) and defines the maximum allowed concentrations of every single component released from a WWTP into the receiving water body. Regarding this point, the introduction of improved AI decision-support techniques [4] can help to manage complex problems, such as stabilizing the composition of the effluent, preventing microbiological pollution, minimizing general environmental-impacts from wastewater treatment outflows, and evaluating the safety of water re-use.

### 3.3 Aerobic biological treatment processes

This paper deals with WWTPs which have an aerobic biological reactor (the unit that contains the activated sludge). The choice of aerobic biological processes is appropriate for the treatment of urban wastewaters, due to their origin and composition. This choice is also valid for some industrial wastewaters, whose treatment however requires a special preliminary study.

The activated-sludge process consists essentially in a mixed microbial culture which grows aerobically on the (organic and inorganic) components of wastewater, producing new microorganisms which are separated, in the settlers, from treated water. The biological purification process could indeed develop also in anaerobic environments, but in any case it needs such conditions as to promote the growth of microbial material. All microorganisms show common nutritional demands: they need a carbon source from which to obtain energy to grow and reproduce; besides, they need nitrogen, phosphorus, potassium, oxygen (if the process is aerobic) and water; and finally they can show specific demands for different chemical elements and organic compounds (Ca, Mg, Fe, S, vitamins and hormones). The final products of microorganism activity, besides new biomass, are: catabolism by-products, water and carbon dioxide (in aerobic processes), or methane and carbon dioxide (in anaerobic processes).

To control the process of the activated-sludge unit, knowledge about the biological and chemical characteristics of the water and about the microorganisms of the sludge is needed. It is also important to have some knowledge about urban washing and rain water, because they contribute a lot of pollutants, and high concentrations of toxic polluting-substances in the WWTP are able to prejudice the efficiency of the treatment and the equilibrium of the sludge environment.

All biological processes in the activated-sludge reactor are characterized by the nature of the substrates, by the kind of microorganisms and by the final products that form. In the biological treatment of a WWTP, the demolition of the organic matter (substrate) contained in the wastewater by the microorganisms takes place according to processes analogous to the ones of water-stream auto-purification, but in such a way as to make the transformations occur with much higher rate and performance. The cellular metabolism of WWTP-microorganisms is made up of two phases:

- anabolism: constructive phase of the metabolism in which microorganisms transform the substrate into the material that constitutes the protoplasm and the reserve substances;
- catabolism: transformation, generally destructive, of the substrate, with release of the necessary energy for the synthesis and cellular preservation processes.

The biochemical activity exerted by microbial cells on the substrate takes place through enzymes produced by the very cells. These enzymes are in part discharged into the extra-cellular environment and in part held in the cells or on the cells. The enzymes take part in the reactions which govern energy freeing and protoplasmic synthesis; the transformation of the organic matter constituting the substrate occurs through the formation and subsequent decomposition of enzyme-substrate complexes.

### 3.4 Microorganisms in WWTP management

The metabolic activity of WWTP-microorganisms is very high, but they are also very sensitive to the parameters that affect cell reactivity and kinetics (temperature, radiation, pressure, pH and substrate concentration). In the paper we will not go into the details of the kinetics of the biological reactions (modeled by Monod in [9]), but we will deal with the relation among the presence of certain microorganisms, the running of the biological purification process and the state of the WWTP.

To evaluate the presence of microorganisms, a periodic microscopic observation of the activated sludge is essential. Some microbial populations, such as the one of the filamentous bacteria or of the *actinomycetes* which produce biological foams, are important for their negative influence (on wastewater treatment), while other ones, such as bio-P bacteria (or phosphorus-accumulating bacteria), are important for their positive role. Now, even if various kinds of microorganisms can be found in a WWTP, the biodegradation activity is mostly carried out by bacteria (although protozoa and small metazoans can contribute to effluent pureness), to which the prokaryotic nature gives extreme versatility in the use of a wide range of substrates and very reduced duplication times. In the aerobic processes, aerobic and facultative-anaerobic bacteria are active, while, in strictly anaerobic processes, facultative and obligatory anaerobic bacteria (like the ones involved in methane formation) are active. Moreover, in plants with special nutrient removal units, there are also aerobic bacteria with the possibility of anaerobic respiration (denitrifying bacteria, which carry out, in the anoxic reactors, the nitrate reduction).

The microbial composition of the activated-sludge (a micro-community in the form of biological flocks) and its activity depend on many factors: kinetic constants of the different species, availability and nature of the carbon substrate and physical conditions (pH, temperature, dissolved oxygen, mixing system, plant configuration). Besides, it is necessary to consider also the interactions among different components of the micro-community, such as predation, commensalism and food competition.

While simple visual microscopic observations give sufficient information on activated-sludge microbial populations, in anaerobic biological treatment the main tool of characterization of anaerobic microorganisms are estimations by means of microbiological analyses (which are more difficult and expensive) of the most important microbial populations.

Considering the variability of several parameters which are involved and the complexity of the activated-sludge system, it is a very arduous task to reach an equilibrium and to keep steady the composition of biological aggregates while guaranteeing specific results in terms of purification. Indeed, despite the progress in the study of the microbial ecology of activated sludge we still do not have a complete knowledge of the interventions to be adopted when problems like bulking[6] or foam abnormal-production arise.

With the aid of suitable AI systems, including an ontology, the automatic association of the presence and evolution of microorganism communities with the control and management of WWTPs seems to be feasible. For example, it is possible to identify, on a morphological basis, numerous filamentous-bacteria which cause bulking or foaming problems and with such information, even having a limited case history, to outline an operational modification of the plant in order to restrain the growth of the responsible microbial species and then to recover from the malfunction possibly generated [12].

Now that the domain we have to deal with has been defined, a detailed analysis follows of the ontology which have been developed. As before, in next section we keep the focus on the activated-sludge variant of the treatment process.

# 4 WaWO ONTOLOGY

## 4.1 Why an ontology for wastewater treatment?

**Terminological confusion** In the domain of wastewater treatment, many different experts work together. Their disciplines of expertise are: artificial intelligence, chemical engineering, chemistry, computer science, environmental engineering, microbiology.

- Each expert uses a particular vocabulary (a precise common terminology does not exist).
- There are no rules helping in the use of each term.
- Synonyms exist.
- Some term can be used in different disciplines with similar, but not identical, meanings (semantic differences appear using the same term in different disciplines).

All these reasons suggest the need for the creation of a unified, complete and consistent terminology, which can be used in different formal contexts and applications related to the wastewater treatment domain. Ontologies are a practical way to achieve this goal.

But the ontology that we propose, named WaWO, models not only the terminology but also **it models the entire domain of wastewater processes and wastewater treatment management**, with particular attention to the activity of the microbiological component.

## 4.2 Goals

The goal of the WaWO ontology is to create a model that:

1. provides a shared terminology for the wastewater domain that each agent can jointly understand and use,
2. defines the meaning of each term (part of the semantics) in a precise and as unambiguous manner as possible,
3. encodes a deeper microbiological knowledge with respect to existing knowledge-based systems,
4. links terms and concepts with taxonomic / hierarchical relations,
5. implements the semantics in a set of axioms that will enable WaWO to automatically deduce the answer to many questions about the wastewater domain, such as cause-effect questions,
6. has axioms which permit diagnosis-impasse solving,
7. exploiting axioms' deductions, semi-automatically generates a knowledge base of new rules to be used by the rule-based expert system of DAI-DEPUR+,
8. uses the Ontolingua formalism for depicting a term or the concept constructed thereof in a graphical context,
9. integrates with some temporal reasoning, based on transition networks, to obtain a qualitative simulation of the evolution of WWTP states.

**Axioms** We approach goals 5, 6 and 7 by defining a set of axioms (or rules) that describe wastewater processes. Axiom deductions should be determined by a set of questions used to decide the competence of WaWO representation. Since there does not exist a standard for determining the competence of a model, we will define a set of questions about wastewater processes and the axioms used to answer them.

**Basic entities** The basic entities in the WaWO model are represented as objects with specific properties and relations. Objects are structured into a taxonomy and the definitions of objects, attributes and relations are specified according to the Ontolingua version of the frame ontology.

## 4.3 Peculiar features

WaWO integrates the knowledge about WWTP management with the current knowledge about aerobic processes and aerobic microorganisms, and it will help to create a general model for wastewater treatment process management.

This ontology is built on the KSL Ontology Server (in the Ontolingua language). It is then translated into Lisp and embedded into DAI-DEPUR+ system.

In the following we discuss the main features of the WaWO ontology and some issues that arise in the representation of a specific domain, such as wastewater treatment, through the construction of an ontology. Then, in next sections, we describe the steps followed to create WaWO.

The main features of WaWO ontology are:

[6] Situation of poorly settling sludge.

**Table 1.** The Ontolingua library

| | | |
|---|---|---|
| 3d-Tensor-Quantities | 20 September 1995 | (3 definitions) |
| Abstract-Algebra | 1 September 1997 | (25 definitions) |
| Agents | 11 December 1999 | (8 definitions) |
| Basic-Matrix-Algebra | 20 September 1995 | (26 definitions) |
| Bibliographic-Data | 28 January 1999 | (79 definitions) |
| Chemical-Crystals | 28 January 1999 | (121 definitions) |
| Chemical-Elements | 20 April 1999 | (170 definitions) |
| Cml | 12 January 2000 | (96 definitions) |
| Component-Assemblies | 16 January 2000 | (15 definitions) |
| Components-With-Constraints | 30 September 1997 | (7 definitions) |
| Device-Ontology | 30 September 1997 | (28 definitions) |
| Dme-Cml | 25 October 1999 | (6 definitions) |
| Documents | 28 October 1999 | (31 definitions) |
| Enterprise-Ontology | 28 January 1999 | (177 definitions) |
| Frame-Ontology | 6 February 1999 | (66 definitions) |
| Hp-Product-Ontology | 30 September 1997 | (124 definitions) |
| Hpkb-Upper-Level-Kernel | 14 August 1999 | (2180 definitions) |
| Interface-Definition-Language | 2 September 1997 | (27 definitions) |
| Interface-Ontology | 4 August 1999 | (51 definitions) |
| Job-Assignment-Task | 28 January 1999 | (20 definitions) |
| Kif-Extensions | 30 August 1997 | (12 definitions) |
| Kif-Lists | 28 August 1997 | (22 definitions) |
| Kif-Meta | 28 January 1999 | (39 definitions) |
| Kif-Numbers | 28 August 1997 | (89 definitions) |
| Kif-Relations | 30 September 1997 | (18 definitions) |
| Kif-Sets | 22 November 1997 | (31 definitions) |
| Mace-Domain | 1 September 1997 | (26 definitions) |
| Mechanical-Components | 20 September 1995 | (8 definitions) |
| Okbc-Ontology | 7 January 2000 | (46 definitions) |
| Parametric-Constraints | 28 January 1999 | (24 definitions) |
| Physical-Quantities | 28 January 1999 | (28 definitions) |
| Product-Ontology | 28 January 1999 | (46 definitions) |
| Simple-Time | 16 March 1999 | (254 definitions) |
| Slot-Constraint-Sugar | 28 January 1999 | (14 definitions) |
| Standard-Dimensions | 9 October 1997 | (45 definitions) |
| Standard-Units | 23 November 1997 | (61 definitions) |
| Tensor-Quantities | 30 September 1997 | (25 definitions) |
| Unary-Scalar-Functions | 20 September 1995 | (14 definitions) |
| Vt-Design | 30 September 1997 | (12 definitions) |
| Vt-Domain | 30 September 1997 | (769 definitions) |
| Wawo-Wastewater | 15 January 2000 | (126 definitions) |

#### *4.3.1 Objects*

In January 2000, WaWO class hierarchy has 100 classes defined. The final number of classes we plan to identify will be between 100 and 200. For a comparison, the average number of classes of the ontologies loaded in the Ontolingua Server is less than 100 and the Upper Cyc ontology has 2180 definition (see Table 1).

The WaWO ontology includes general terms about the wastewater domain, but part of the ontology is microbiology oriented and encodes a deep knowledge about microorganism taxonomy.

#### *4.3.2 Impasse solution*

The axioms of the ontology connect the descriptors to the states of the WWTP. More general states are linked to the descriptors by more general axioms, while less general states or substates are linked to the descriptors by more specific axioms. In the case of microorganism descriptors, this means linking the state of the plant to higher or lower levels of microorganism taxonomy. (Example: in case of presence of two microorganism *species* and lack of axioms (or discordant axioms), the connection to the state of the WWTP is done at the taxonomic-*class* level.) In this way, we endow the ontology with inference processes. And we improve the diagnosis and the decision support.

**Diagnosis** WaWO is integrated with the KBSs mentioned in section 1.2 above. With WaWO it is possible to capture, understand and describe the knowledge about the whole physical, chemical and microbiological environment of a WWTP. The hierarchical structure and the axioms of the ontology can help to diagnose the situation in case of *impasse* of the other KBSs.

The ontology is normally static. It activates its inference mechanisms (axioms) only under specific petitions from the diagnosis integrator (see next paragraph). The result of the inference of the ontology is:

- an answer about the diagnosis impasse (e.g.: 'We have a foaming situation' or 'I do not have information to solve the impasse'),
- an explanation of the answer (e.g.: 'I received information related to the answer from the activation of the following axioms ...' or 'The answer was obtained searching the following classes ...').

The activation of the ontology always means that there was an impasse in KBS diagnosis. If the answer of the ontology to the petition is 'I do not have information to solve the impasse', then a primary alarm is activated.

**Diagnosis integration** The rule-based expert system (RBES) and the case-based reasoning system (CBRS) work independently (and in parallel, if necessary) and they both produce as output a diagnosis on the state of the plant. This output is passed to the diagnosis integrator, a subsystem between the diagnosis and the decision support layers.

**General integration schema** If the diagnosis of the two KB systems is the same, it is passed to the decision support layer. If the diagnoses exist and are different, the system prioritizes as follow:

- If the case library contains a predefined minimum historical series and the case similarity is higher than a predefined value, the case-based reasoner's diagnosis prevails.
- Otherwise, the rule-based expert system's diagnosis prevails.

In case of impasse (no diagnosis), DAI-DEPUR+ turns first to the ontology and then, if it fails, to the plant manager, demanding an off-line diagnosis based on their microbiological deep knowledge. This external solution is learned.

**Detailed integration schema**

1. $\notin$ CBRS diagnosis and $\notin$ RBES diagnosis: impasse, the diagnosis integrator turns to the ontology.
2. $\notin$ CBRS diagnosis and $\in$ RBES diagnosis:
   - RBES diagnosis-certainty $\geq a$: RBES diagnosis passed to decision support layer.
   - RBES diagnosis-certainty $< a$: impasse, the diagnosis integrator turns to the ontology.
3. $\in$ CBRS diagnosis and $\notin$ RBES diagnosis:
   - CBRS case-similarity $\geq b$: CBRS diagnosis passed to decision support layer.
   - CBRS case-similarity $< b$: impasse, the diagnosis integrator turns to the ontology.
4. $\in$ CBRS diagnosis and $\in$ RBES diagnosis:
   - CBRS case-similarity $\geq b$: CBRS diagnosis passed to decision support layer.
   - CBRS case-similarity $< b$:
     - RBES diagnosis-certainty $\geq a$: RBES diagnosis passed to decision support layer.
     - RBES diagnosis-certainty $< a$: impasse, the diagnosis integrator turns to the ontology.

**Example (case 1)** We have a certain perception state *A*. The case-based reasoning system (CBRS) and the rule-based expert system (RBES) activate.

The CBRS finds a case similar to state *A*. The similarity value is 0.1 and it is less than the minimum acceptable value *b* (e.g. b=0.2). Therefore there is no diagnosis output from the CBRS.

The RBES finds no rules leading to a diagnosis starting from state *A*. Therefore there is no diagnosis output from the RBES.

The diagnosis integrator acknowledges a case of missing diagnosis from the KBSs and send a petition to the ontology with the description of state *A*.

The output of the ontology is: Answer = 'We have a foaming situation', Explanation = 'I received information related to the answer from the activation of the following relation (Nocardia *is a* Actinomycetes) and the following axioms (Actinomycetes *is cause of* Foaming sludge)'.

Comment: State *A* is characterized by a strong presence of Nocardia bacterium, but this bacterium has no direct relation (in the static knowledge-base of the RBES) to any state of the WWTP. Even in the ontology the Nocardia class has no link to any state of the WWTP, but its parent class (Actinomycetes) has a cause-effect link to the general state of the WWTP 'Foaming sludge'. One reason for this could be that Nocardia is not causing foaming, but another (not detected) bacterium of the same taxonomic class is.

The diagnosis integrator receives the diagnosis from the ontology and pass it to the decision support layer.

#### *4.3.3 Representation of cause-effect relations*

Through the definition of axioms we are able to represent two kinds of relations. We can associate microorganisms to the problematic situations that they cause, and we can associate the actual state of the plant to the actions that need to be performed in order to reach the normal state from that actual state.

#### *4.3.4 Knowledge-base production*

The semi-automatic translation from the ontology in the Ontology Server to the Lisp code give implicitly rise to the relative ontological knowledge base, after some code depuration.

Furthermore, axiom deductions can be exploited for a semi-automatic generation of a static knowledge-base to be used as a rule-based expert system prototype.

### 4.4 Methodology

The WaWO ontology construction follows these steps:

1. We first study the potential utility of the ontology (see section 4.1).
2. We then define the purposes for the ontology (see section 4.2).
3. We next carry out the knowledge acquisition process (see [3]). We are able to reuse some existing specific encoded knowledge and for the rest we go through interviews with experts and texts analysis.
4. In the conceptualization step, we identify a preliminary list of the objects in our domain of discourse: the wastewater.
5. Then we choose as development environment for the ontology the quasi standard Ontolingua on-line ontology-editor.
6. Next we decide which other existent ontologies to reuse in the creation of WaWO and start the formalization.
7. Objects are represented by classes and individuals in the Ontolingua language. We identify the properties of these objects and the relations that exist over them (represented by relations and functions in the Ontolingua language).
8. We next define a set of axioms in Lisp or KIF to represent the constraints over the objects and properties in the ontology. This set of axioms constitutes a microtheory [8] and provides a declarative specification for the various tasks we want to model.
9. Eventually, we need to prove results about the properties of our microtheory in order to provide a characterization and justification for our approach; this enables us to understand the scope and limitations of the approach.

In the following sections, we describe some of these steps in detail.

### 4.5 Conceptualization

In building WaWO, we first identify the objects in our domain of discourse: the wastewater. One part of the objects is related to the WWTP (e.g.: actuator, descriptor, body of water, treatment, wastewater, WWTP); another part is related to the AI control software (e.g.: output datum, plan, state of the WWTP); and a third part constitutes the microbiological taxonomy (e.g. superkingdom, kingdom, class, order, species), according to known and accepted classifications. A few more detailed examples follow:

- Descriptor
  - Descriptor-Off-Line
    - Descriptor-Qualitative
      - Appearance-Floc
      - Appearance-Surface-Clarifier
      - Biomass-Color

    * Descriptor-Quantitative
      - BOD
      - Chlorine
      - COD
  - Descriptor-On-Line
    * Water-Flow
- State-Wwtp
  - State-Wwtp-Abnormal
    * Deflocculation
    * Denitrification
    * State-Abnormal-Secondary-Treatment
      - State-Abnormal-Biological-Origin
      - State-Abnormal-Non-Biological-Origin
  - State-Wwtp-Normal
- Bacteria (superkingdom)
  - Firmicutes(gram-positive bacteria)
    * Actinomycetes(high G+C gram-positive bacteria) (class) (slots: Qualitative-Abundance, Object-Found-In-Location) (axioms: '*Actinomycetes* Cause-Of *State-Abnormal-Biological-Origin*'
      - Nocardia (slots: Qualitative-Abundance, Image, Object-Found-In-Location) (axioms: )
    * Bacillus/Clostridium group(low G+C gram-positive bacteria) (class) (slots: Qualitative-Abundance, Object-Found-In-Location) (axioms: )
  - Fibrobacter/Acidobacteria group

### 4.6 Development environment

After an exhaustive bibliographic research and taking into account the advises of other ontology engineers of Edinburgh University, we decided to adopt the Ontolingua formalism and development environment to build the WaWO ontology. Ontolingua is a quasi standard in ontology construction. It is a complete work environment, which includes partial methodological guidelines, a specific language and an on-line editor (accessed through an HTML interface at http://WWW-KSL-SVC.stanford.edu:5915/&service=frame-editor).

The meta-ontology used by Ontolingua is the Frame Ontology which is used by default by all the ontologies being built on the server. The object of the wastewater domain are represented as classes and individuals. In more general terms, these classes and individuals, as well as functions, relations and axioms, are all frames (in the context of the meta-ontology).

At the end of its basic construction, WaWO will be translated to Lisp to be integrated with DAI-DEPUR+. The majority of axioms are added in this phase, directly in Lisp. However we consider fundamental the translation of the axioms from Lisp to KIF, the completion of the on-line version of WaWO, and its final publication in the Ontolingua ontology library.

### 4.7 Reuse of existing ontologies

In WaWO we (re)use the Frame Ontology, as it is the default meta-ontology in the Ontolingua environment.

We also reuse the UpperCyc ontology, the Ontolingua translation of the most general part of Cyc ontology. UpperCyc, at the moment, includes 2180 classes and covers most of the top concepts of WaWO, such as *body of water*, *wastewater* or *town*. Unfortunately, in UpperCyc many basic concepts of wastewater domain, such as *bacterium*, *descriptor* or *effector*, are not included.

### 4.8 Axioms and dynamics

Ontology dynamics is defined by axioms. They are the elements which permit the detailed modeling of the domain. In WaWO, axioms definition is still at the prototype level. As shown in the paragraph on diagnosis (section 4.3.2), the explicitation of axioms about cause-effect relations between microbe presence and state of the plant and between state of the plant and actuation can generate a certain degree of dynamicity. Inferences are started under petition of other elements of DAI-DEPUR+ (the KBSs); if inferences are successful, these petitions are positively answered, and the answers explained.

But in wastewater domain the relations among concepts are not completely well-known. Sometimes the experts do not agree on the structure of the domain. Although various mathematical models have been introduced to describe the relation between microorganisms and substrate and between microorganisms and states of the WWTP, they do not provide an entirely satisfactory description of the cause-effect links existing within the treatment plant. For this reason, the inferences of the ontology on cause-effect axioms are limited by the actual knowledge available to the experts at the moment of axiom encoding.

### 4.9 Advantages coming from WaWO integration into DAI-DEPUR+

The advantages which come from the integration of WaWO ontology into DAI-DEPUR+ decision support system are:

- More reliable management:
  - solution of impasses,
  - use of microbiological knowledge,
  - qualitative simulation and prediction (once temporal integration is done),
  - cause-effect reasoning.
- Semi-automatic generation of a static knowledge base to be used as a rule-based expert system prototype.
- Reuse / portability towards similar domains.
- Potential knowledge-sharing among various WWTP.

## 5 CONCLUSION AND WORK PROGRESS

In collaboration with biologists and chemical engineers, we will complete the evolution in WaWO from a superficial knowledge to a deeper knowledge of the microbiological component of the wastewater treatment process. This should allow a better management of wastewater treatment plants.

A process of adaptation, update and refinement of current AI reasoning methods (rule-based and case-based reasoning) will take place for the integration of the ontology into the resultant environmental decision-support system.

In details, future work includes:

- Full development of the ontology for wastewater treatment management (WaWO), through a deeper study of the *microbiological component* of the wastewater treatment process and the addition of axioms about supervision.

- Modeling of static cause-effect relations within the ontology.
- Integration of the WaWO ontology with some temporal reasoning. For this, a primary task which we will address is that of temporal projection in wastewater treatment. This induces the following set of requirements on the ontology:
  - Temporal projection requires the evaluation of the truth value of a proposition at some point in time in the future. We therefore need to define axioms that express how the truth of a proposition changes over time. In particular, we need to address the frame problem and express the properties and relations that change or do not change as the result of an activity.
  - We must redefine the notion of a state of the WWTP, that is, define what is true of the WWTP before and after performing different activities. This is necessary to express the causal relationship between the preconditions and the effects of an activity (dynamic cause-effect relations).
  - The time interval over which the state has a certain status is bounded by the times at which the appropriate actions that change status occur. This interval defines the duration of a state. This is essential for the construction of schedules.
  - We want a uniform hierarchical representation for activities (aggregation). Plans and processes are constructed by combining activities. We must precisely define how activities are combined to form new ones. The representation of these combined activities should be the same as the representation of the sub-activities. Thus, aggregate activities (sets of activities or processes) can themselves be represented as activities.
  - The causal and temporal structure of states and sub-activities of an activity should be explicit in the representation of the activity.
- Construction of a sub-ontology for knowledge management: this sub-ontology will connect *impasse situations* to the agents (people, systems) with the knowledge to solve them.
- Publication of the WaWO ontology on the KSL Ontology Server.
- Creation of a local graphic interface to edit the translated version of the WaWO ontology.
- Revision of DAI-DEPUR CBR system and rule-based system to adapt to the new ontology: case redesign, library differentiation, temporal evolution, new similarity measures.
- **Evaluation**: we need to prove results about the properties of WaWO microtheory in order to provide a characterization and justification for our approach; this will enable us to understand the scope and limitations of the approach.

## Acknowledgments

This research is supported by the Generalitat de Catalunya (grant 1997FI 00630 UPC PG) and by the Spanish CICYT project AMB97-0889.

This paper integrates with the "Generation, construction and integration of knowledge bases for knowledge-based systems" research project, carried out by people of the KEML group at UPC, Barcelona, Spain. Previous work of the KEML group includes a research line about wastewater treatment plant management, from which some information has been reused for this paper.

The WaWO ontology has been developed with the collaboration of the Software Systems and Processes Group (Edinburgh University) headed by professor Dave Robertson.

## REFERENCES

[1] Beccari, M., 1991. *Principi dei trattamenti biologici*, in Canziani, R. (editor) *Trattamento delle acque di rifiuto*, pp. 99-113, Istituto per l'Ambiente, Milano, Italy 1991.

[2] Borst, P., Akkermans, H. and Top, J., 1997. *Engineering Ontologies*, in International Journal of Human-Computer Studies 46, pp. 365-406, 1997.

[3] Comas, Q., R-Roda, I., Ceccaroni, L. and Sànchez-Marrè, M., 1999. *Semi-automatic learning with quantitative and qualitative features*, in Proceedings of the VIII Conference of the Spanish Association for Artificial Intelligence (CAEPIA-99), November 16-19, 1999, Murcia, Spain.

[4] Cortés, U. and Sànchez-Marrè, Ceccaroni, L., R-Roda, I. and Poch, M., 2000. *Artificial intelligence and environmental decision support systems*, accepted by Applied Intelligence (2000).

[5] Gómez-Pérez, A. and Benjamins, R., 1999. *Overview of Knowledge Sharing and Reuse Components: Ontologies and Problem-Solving Methods*, in Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods(KRR5), Stockholm, Sweden, August 1999.

[6] Gruber, T.R., 1993. *A Translation Approach to Portable Ontologies*, in Knowledge Acquisition 5(2), pp. 199-220, 1993.

[7] Jones, D.M., Bench-Capon, T.J.M. and Visser, P.R.S., 1998. *Methodologies for Ontology Development*, IT&KNOWS - Information Technology and Knowledge Systems, 15th IFIP World Computer Congress, Vienna (Austria) and Budapest (Bulgaria).

[8] Lenat, D. and Guha, R.V., 1990. *Building Large Knowledge Based Systems: Representation and Inference in the CYC Project*, Addison Wesley Pub. Co., 1990.

[9] Monod, J., 1942. *Recherches sur la croissance des cultures bactériennes*, Hermann and Cie, Paris 1942.

[10] Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T. and Swartout, W.R., 1991. *Enabling Technology for Knowledge Sharing*, in AI Magazine (winter 1991), pp. 36-56.

[11] Sànchez-Marrè, M., Cortés, U., Lafuente, J., R-Roda, I. and Poch, M., 1996. *DAI-DEPUR: an integrated and distributed architecture for wastewater treatment plants supervision*, in Artificial Intelligence in Engineering 10 (1996), pp. 275-85.

[12] Tandoi, V., 1991. *Trattamento a fanghi attivi - Gestione del processo e aspetti microbiologici*, in Canziani, R. (editor) *Trattamento delle acque di rifiuto*, pp. 165-84, Istituto per l'Ambiente, Milano, Italy 1991.

[13] Uschold, M., 1998. *Knowledge level modelling: concepts and terminology*, in The Knowledge Engineering Review, 13(1), pp. 5-29, 1998.

[14] Uschold, M. and Grüninger, M., 1996. *Ontologies: principles, methods and applications*, in The Knowledge Engineering Review 11(2) (1996), pp. 93-136.

[15] Wen, C-H. and Vassiliadis, C.A., 1998. *Applying hybrid artificial intelligence techniques in wastewater treatment*, in Engineering Applications of Artificial Intelligence 11 (1998), pp. 685-705.