A knowledge-based system for the diagnosis of waste-water treatment plants § *

Lluís Belanche
Miquel Sànchez
Ulises Cortés
Universitat Politècnica de Catalunya
Departament de Llenguatges i Sistemes Informàtics
Pau Gargallo, 5. Barcelona. 08028. Catalonia, **SPAIN**
belanche@lsi.upc.es, sanchez@lsi.upc.es
Pau Serra
Universitat Autònoma de Barcelona
Departament de Química
Edifici C, 08193 Bellaterra (Barcelona). Catalonia, **SPAIN**
iqte2@cc.uab.es

Abstract

In this work we discuss the development of an expert system with approximate reasoning which resorts to a new methodology for attribute selection in **knowledge-based systems**. *First, we make a survey of the purifying process and its problems, as well as those of conventional automatic control methods applied to industrial processes. Next, we establish a definition of the relevance concept for a given set of attributes, which includes the special case of non-relevant attributes or* **nought attributes**. *A new heuristic is here proposed in such a way that it finds out the more relevant attributes from those initially selected by the expert, reducing the cost of the formation & validation of decision rules and helping to clarify the underlying structure of a non well-structured domain as are waste-water treatment plants.*

1. Introduction

A malfunction of a waste water treatment plant is a social and biological hard problem. A poorly treated waste water outside the plant can achieve dangerous consequences for human beings as well as the nature itself.

In this paper we show the application of Knowledge Based Systems to the diagnosis and control of waste water treatment plants. The work has been developed in a wide cooperation among the Chemical Engineering group of the UAB, the Functional Languages and Artificial Intelligence of the UPC and experts from real plants. The use of an AI tool as the Expert Systems, as a first approach, gives a new vision to control and diagnose industrial processes.

During the last years the AI research effort in the field of industrial processes control has been increasing quite a lot ([**MAED89**], [**FLOR89**], [**SANZ89**].) The conventional automatic process control systems (feed-back control, feed-forward control, optimal control, adaptive control, ...) have to cope with some difficulties:

- ⋆ The problem's complexity.
- ⋆ A non well-structured domain.
- ⋆ Most information is neither numeric nor quantified.
- ⋆ Uncertainty or approximate knowledge.
- ⋆ A dynamic system.

Finally, we can say that the classical methods work on the "normal" state of the plant but not in other "extra-normal" states of working. Bringing to mind the main advantages of expert systems and the above difficulties, one could think that match very well. Hence, we have developed an expert system to the diagnosis with uncertainty reasoning [**SANC91**], [**SANC90**].

§ Partially supported by GRANT TIC-90 801/C02. CICyT. ESPAÑA
* Partially supported by GRANT ROB-89 0479-C03-02. CICyT. ESPAÑA

2. Goals

The main goal, already acquired, is to develop an expert system with approximate reasoning applied to diagnosis in waste water treatment plants in order to make easier the plant control and supervision, to increase the water quality outside the plant and to reduce the cost of the operations. The plants taken as models are based on the main technic usually applied: the activated-sludge process **[ROBU90]**. In the future, we shall extend the system to build a Supervisor Knowledge Based System (see Fig. 1) applied to diagnosis and control, integrating both heuristic knowledge and numerical processing algorithms (statistical estimations, ...).

Fig. 1: A Supervisor Knowledge-Based System

3. Development and Design

We have built up an expert system called DEPUR, over the shell MILORD, in order to support the approximate reasoning. The main problem on a waste water treatment plant is the balance's break between biomass and substrate. This problem may be originated in one of the three treatments that experiments the water until it exits the plant: primary, secondary and tertiary. Each one carries out the elimination of a concrete pollutant. Thus, the diagnosis process tries to determine the causes of the problems detected. An expert system is composed by the Data Base, the Knowledge Base, the Meta-knowledge Module, and other general modules independent of the application domain. Hence, to develop the system it is required to build the Data Base, the Knowledge Base as well as the Meta-knowledge Module (the control strategies). The design process of an intelligent system can be expressed commonly as shown in Fig. 2.

Fig. 2: Design process of an intelligent system. Fig. 3: Diagnostic process in an Expert System

A waste water treatment plant is composed usually by two independent but interacting subsystems: sludge line and water line. We have focused on the water line subsystem domain. The *domain modelling phase* includes relevant attribute selection and attribute structuring. The attribute selection process has been developed twofold: firstly, in a typical manner (expert attribute selection) and, secondly, by means of a new automatic attribute selection approach based on the relevance and potential utility concepts **[BELA91]**. The knowledge engineering process (attribute structuring) relies upon two approaches:

⋆ *Decision trees*: the classical approach derived from the interactions between the expert and the engineer. This approach is governed by the cause-effect criteria. We have used multi-valued decision trees from which inference rules are derived. The rules point out to discover the causes of a certain problem or situation previously detected.

⋆ *Automatic classification techniques*: more recent approach derived from automatic learning in AI. We have used the LINNEO software in this phase, based on fuzzy non well structured domain classifications. From the characterization of the classes obtained more general inference rules are derived. Them lead to know the most probable working situation of the plant.

Decision trees have been used in the causes identification phase and automatic classification techniques in the problem identification process as shown in Fig. 3.

4. Attribute selection

4.1. Introduction

The nature of classification has been widely studied in **ML**. In special those processes that can be performed automatically by a machine. Many successful first-generation expert systems were in practice classifiers. This kind of systems uses a set of rules, implemented as decision trees, to determine the class of a given input on the basis of a set of its characteristics. In the classical approach, the human expert is responsible for deciding which are the *relevant attributes* for the classification and formulation of rules. The limitations of this approach, pointed out by several authors (*e.g.*: [**HAYE84**], [**BAIM88**], etc.), have stimulated the development of systems that could take the responsibility of deciding whether an attribute is potentially useful or not, by using methods of automatic inductive inference for discovering the patterns or relations between data that can be useful in the formation rules of classifications. In **ML**, several *heuristic* methods have been developed for attribute evaluation in terms of their potential utility (that has been identified with relevance) in order to decide the class to which the input belongs. Almost all of them are based upon the classical information theory (*e.g.*: [**QUIN79**], [**QUIN86**], [**SCHL86**], [**BAIM88**], [**LOPE90**], etc.).

4.2. The attribute relevance problem

In this section a characterization of the attribute relevance problem is made, and a formal definition of relevance and non-relevant attributes or **nought** attributes is proposed.

4.2.1 Characterization of relevance

Let $\mathcal{U}$ be a universe and $\mathcal{C} = \{C_1, C_2, \ldots, C_m\}$ a classification (partition) of $\mathcal{U}$. Let us assume that the elements of $\mathcal{U}$ can be represented by the ordered n-tuple $(A_1(u), \ldots, A_n(u))$ of attributes (*i.e.* measurable characteristics) of the objects in $\mathcal{U}$, with ranges $X_1, X_2, \ldots, X_n$, respectively. Formally speaking, every A_i is a function with domain $\mathcal{U}$ and range X_i. Notice that, in general, $\{(A_1(u), \ldots, A_n(u)) \mid u \in \mathcal{U}\}$ is a subset of $X_1 \times \ldots \times X_n$ but not necessarily equal. Let us designate r_i as the cardinality of X_i.

Intuitively expressed, the more *information* an attribute supplies to a classification, the more relevant to the classification it is. The natural way to *measure* the information an attribute supplies is by means of its capacity to discriminate the elements of $\mathcal{U}$. The relevance of an attribute is not an absolute property, but is relative to a classification or interpretation. An attribute can be very important for a classification but *irrelevant* in other. For instance, people's color of skin is not important to determine (discriminate) the kind of cardiovascular diseases they have, but it is important to determine their ethnic classification. In practice, the potentially useful initial set of attributes should be selected by the expert [**BAIM88**].

We state here that the relevance of an attribute depends not only on the proportion of objects that this attribute discriminates, but in *how* it interacts with the rest of the attributes. Any formal definition of relevance must then capture these two concepts. The following definitions are taken from [**NUÑE91**], and will serve as a theoretical approach to the relevance problem.

Definition

The attribute A_i is more relevant than the attribute A_j, if and only if the minimum number of attributes to be added to A_i, to obtain a fully discriminating set, is less than the number required to A_j. If this number is the same for both attributes, then we conclude that they are equally relevant.

This definition could be expressed as:

Let $N = \{1, 2, \ldots, n\}$, and, for every $S \subset N$, let

$$\begin{aligned} \mathcal{A}_S &= \{A_s \mid s \in S\}, \\ \mathcal{S}_i &= \{S \subset N \setminus \{i\} \mid \{A_i\} \cup \mathcal{A}_S \text{ is fully discriminating}\} \\ \mathcal{S}_j &= \{S \subset N \setminus \{j\} \mid \{A_j\} \cup \mathcal{A}_S \text{ is fully discriminating}\} \end{aligned}$$

Then A_i is more relevant than A_j if, and only if,

$$min\{\#S | S \in \mathcal{S}_i\} < min\{\#S | S \in \mathcal{S}_j\}.$$

A_i is equally relevant than A_j if, and only if,

$$min\{\#S | S \in S_i\} = min\{\#S | S \in S_j\}.$$

Definition

Let S_1 and S_2 be subsets of N, and let

$$\mathcal{S}_{S_1} = \{S \subset (N \setminus S_1) \mid \mathcal{A}_{S_1} \cup \mathcal{A}_S \text{ is fully discriminating}\}$$
$$\mathcal{S}_{S_2} = \{S \subset (N \setminus S_2) \mid \mathcal{A}_{S_2} \cup \mathcal{A}_S \text{ is fully discriminating}\}$$

$\mathcal{A}_{S_1}$ is more relevant than $\mathcal{A}_{S_2}$ if, and only if,

$$min\{\#S|S \in S_{S_1}\} < min\{\#S|S \in S_{S_2}\}.$$

The sets of attributes A_{S_1} and A_{S_2} have the same relevance $\mathcal{R}$ if, and only if,

$$min\{\#S|S \in S_{S_1}\} = min\{\#S|S \in S_{S_2}\}.$$

The intuitive idea that underlies the definition is that two sets of attributes, regardless of their cardinality, are equally relevant if they have the same discrimination capacity, or, equivalently, if *the amount of information that is left for them to be fully discriminating is the same.*

Hence, in order to determine if a given set of attributes is more relevant than another, we should generate, at least, all the fully discriminating (of maximum relevance) sets that contain any of the two given sets. But this is a non-practical and intractable method from the point of view of complexity. Thus the proposed solution has been to establish heuristics to evaluate the potential utility of an attribute, in such a way that they pick up, successively, those which will lead to build decision trees close to the optimum one. The solutions proposed so far have some elements in common, and its discussion lead us to to the heuristic that we propose later on. For a complete survey of these measures and their characteristics, see **[NUÑE91]**.

The definition of relevance for a given set of attributes $\mathcal{A}$ introduces an equivalence relation $\mathcal{R}$ in the power set $\mathcal{P}(\mathcal{A})$, where the equivalence classes are formed by equally relevant sets. Joined with the previous definition, this allows us to introduce a *total order,* $\leq$, in the quotient set of equivalence classes $\mathcal{P}(\mathcal{A})/\mathcal{R}$.

Definition
Let $\mathcal{A}_{S_1}$, $\mathcal{A}_{S_2}$ be two attribute sets and $[\mathcal{A}_{S_1}]$, $[\mathcal{A}_{S_2}]$ their equivalence classes. Then:

$$[\mathcal{A}_{S_1}] \;\leq\; [\mathcal{A}_{S_2}]$$

if, and only if,

$$min\{\#S|S \in S_{S_1}\} \leq min\{\#S|S \in S_{S_2}\}.$$

Note that the classes $[\mathcal{A}]$ and $[\emptyset]$ are, respectively, the maximum, and minimum with respect to this order, *i.e.* $[\emptyset] \leq [\mathcal{A}_S] \;\leq\; [\mathcal{A}], \quad \forall\, \mathcal{A}_S \subset \mathcal{A}$.

Another important remark is that if $\mathcal{A}_{S_1}$ is more relevant than $\mathcal{A}_{S_2}$, then any subset of $\mathcal{A}_{S_1}$ is more relevant than any other subset of $\mathcal{A}_{S_2}$. The problem, in practice, is to obtain the *optimum representative of $[\mathcal{A}]$, which classifies the original set $\mathcal{U}$ with the minimum effort.* This is the aim of the above mentioned algorithms.

2.2 The nought attributes

It has been pointed out so far that the relevance of a set of attributes is not inherent to the attributes, but it depends on the classification. Frequently we have to deal with situations in which an attribute, or in general, a set of attributes, are not important to a given classification process (though they could be relevant to a different one), i.e., *they have no relevance for a given classification.* This null relevance will be referred to as **nought** [**SANC89**] relevance. Hereafter, we will refer to non-relevant attributes as **nought** attributes. In such cases these attributes are yet given and cannot be ignored. What we might want is that such attributes do not disturb the classification considering only the *non*-**nought** ones.

In accordance to the relevance definition, **nought** attributes are those without discrimination capacity. Therefore, any **nought** set has null relevance and is an element of $[\emptyset]$, the empty set.

Definition

A set of attributes $\mathcal{A}_{\mathbf{n}}$ is **nought** if, and only if, $[\mathcal{A}_{\mathbf{n}}] = [\emptyset]$.

Note that any **nought** set of attributes $\mathcal{A}_{\mathbf{n}}$, when added to a given set of attributes $\mathcal{A}_S$, does not alter its relevance, that is, the discriminating capacity of class $[\mathcal{A}_S]$ is the same than that of class $[\mathcal{A}_{\mathbf{n}} \cup A_S], \; \forall \mathcal{A}_S \subset \mathcal{A}$.

Theorem

Let $\mathcal{A}_{\mathbf{n}}$ be a **nought** set of attributes, then $[\mathcal{A}_{\mathbf{n}} \cup \mathcal{A}_S] = [\mathcal{A}_S] \;, \quad \forall \mathcal{A}_S \in \mathcal{P}(\mathcal{A})$

4.3. A measure of the potential utility of an attribute

Given $i \in N$, for every a_{ij} $(j = 1, 2, \ldots, r_i)$, let $p_{jk} = \#(A_i^{-1}(a_{ij}) \cap C_k), where k = 1, \ldots m$, be the number of objects taking the value a_{ij}, and placed in the class C_k. Since $C_1, C_2, \ldots, C_m$ is a partition of the universe, we have that

$$p_j = \sum_{k=1}^{m} p_{jk} = \#A_i^{-1}(a_{ij}).$$

Let $\overline{p_j}$ be the arithmetic average of objects with value a_{ij} in the classes, that is, $\overline{p_j} = \frac{\#A_i^{-1}(a_{ij})}{m}$.
Let $D = \{(p_{j1}, p_{j2}, \ldots, p_{jm}) \mid j = 1, 2, \ldots, r_i\}, \quad i \in N$. Then, the potential utility of each attribute A_i may be expressed as a function from set D to $R^+ \cup \{0\}$.

Definition

Let $u : \mathcal{A} \to \mathcal{R}^+ \cup \{0\}$ be the function which evaluates the potential utility of an attribute, defined as follows:

$$u(A_i) = \sum_{j=1}^{r_i} \sum_{k=1}^{m} (p_{jk} - \overline{p_j})^2.$$

Note that such a function u is non-negative, and that $u(A_i) = \sum_{j=1}^{r_i} (m-1) S_j^2$, where S_j^2 is the statistical variance of the values $\{p_{jk} \mid k = 1, \ldots, m\}$.

Definition

Given a pair of attributes, A_i, A_l, attribute A_i is potentially more useful than attribute A_l if, and only if, $u(A_l) < u(A_i)$.

That is, attribute A_i is more relevant than attribute A_l, if and only if,

$$\sum_{j=1}^{r_l} \sum_{k=1}^{m} (\#(A_l^{-1}(a_{lj}) \cap C_k) - \frac{\#A_l^{-1}(a_{lj})}{m})^2$$
$$< \sum_{j=1}^{r_i} \sum_{k=1}^{m} (\#(A_i^{-1}(a_{ij}) \cap C_k) - \frac{\#A_i^{-1}(a_{ij})}{m})^2.$$

For every $j = 1, 2, \ldots, r_i$, let $u_j\big((p_{j1}, p_{j2}, \ldots, p_{jm})\big) = (m-1) S_j^2$. The sum $(m-1) S_j^2$, for every j, stands for the contribution of value a_{ij} to the potential utility of attribute A_i, and can be interpreted as the *potential utility of value* a_{ij}. This assumed, the potential utility of an attribute is the sum of the potential utility of its values.

A useful feature of the function u_j is that, the farther from the average $\overline{p_j}$ the values $p_{j1}, p_{j2}, \ldots, p_{jm}$ are, the greater the function's value is. An increment to some values makes a decrement to some others. In particular, if distribution is altered in such a way that the proportion of objects in one of the classes is incremented —being decremented in the same amount in any other class—, resulting in a distribution not as random as before, then the potential utility of the value is greater. This can be formally expressed as:

If for some positive integer z, $z \le p_{js}$ we consider $(p_{js} - z)$ and $(p_{jt} + z)$, $s, t \in M = \{1, \ldots, m\}$ and

$$(p_{js} - \overline{p})^2 + (p_{jt} - \overline{p})^2 < \big((p_{js} - z) - \overline{p}\big)^2 + \big((p_{jt} + z) - \overline{p}\big)^2,$$

then

$$u_j\left((p_{j1}, p_{j2}, \ldots, p_{jm})\right) < u_j\left((p_{j1}, \ldots, p_{js} - z, \ldots, p_{jt} + z, p_{t+1}, \ldots p_m)\right).$$

Another useful feature of the estimation of an attribute's potential utility is the fact that u_j is very *sensitive* to the number of objects that each value classifies: the greater the number of objects is, the greater u_j is, being *very* significative the difference between a value that classifies a lot of objects and one not classifying many. As a result of this *sensitivity*, the evaluation of the potential utility of each attribute's value —and that of the attribute itself— is better in much cases, since it takes into account the classification ability of each value and, hence, of the attribute.

The function $u(A)$ is a sum of the $u_j\left((p_{j1}, p_{j2}, \ldots, p_{jm})\right)$, $j = 1, 2, \ldots, r_i$. From the characterization of functions u_j is possible to figure out the features of the function u. The following are properties of the functions u_j, which are next spread out for the function u. Such properties are useful to estimate the potential utility of an attribute. For simplicity, we do not include the subindex j for p_{jk}. Let $(p_1, p_2, \ldots, p_m) \in D$, arbitrary.

Property 1 (Symmetry)

$$u_j\left((p_{\sigma(1)}, p_{\sigma(2)}, \ldots, p_{\sigma(m)})\right) = u_j\left((p_1, p_2, \ldots, p_m)\right),$$

for every permutation σ of degree m.

Property 2 (Maximum of the function u_j)

$$u_j\left((p_1, p_2, \ldots, p_m)\right) \leq u_j\left((0, \ldots, 0, p_1 + p_2 + \ldots + p_m, 0, \ldots, 0)\right),$$

for every position k, $1 \leq k \leq m$.

Property 3 (Minimum of the function u_j)

$$u_j\left((\overline{p_j}, \overline{p_j}, \ldots, \overline{p_j})\right) \leq u_j(p_1, p_2, \ldots, p_m).$$

(In fact, $u_j\left((\overline{p_j}, \overline{p_j}, \ldots, \overline{p_j})\right) = 0$).

The first property says that what actually matters is the number of objects classified in a given class, rather than which class it is. Moreover, in what concerns to function $u(A)$, no matters which is the attribute's value assigning the biggest number of objects to a given class, but that there exists such a value. In particular (property 2), u_j takes its maximum value when all the objects are concentrated in the same class. Finally, u_j is minimum when we have a uniform distribution on the classes (property 3). Then we say that value a_{ij} has its minimum potential utility.

Another remarkable property of function u is that, if the number of objects that a value assigns to one class equals the sum of the number of objects that two values assign to that class, then the potential utility of the single value is greater or equal than that of the two values. This means that the function u does *not* enhance attributes having a lot of values, fundamental for the sake of efficiency.

In order to obtain an operational characterization of the **nought** attributes, those having potential utility zero could be regarded as potentially nought, or we could even force an attribute to have potential utility zero if we wish to *neutralize* it for a classification.

4.4. Validation of the measure

4.4.1 A model for fuzzy classification: LINNEO$^+$

LINNEO$^+$ ([**MART91**]) is a software for fuzzy classification and rule generation which resorts to analytical and empirical knowledge acquisition techniques. The human expert abstracts a collection of observations and a set of attributes he thinks of as relevant. At the same time, the expert is also able to express what he already knows about the domain. This knowledge is represented as constraints and/or examples on classes already known to exist.

Starting with this knowledge and data, LINNEO$^+$ uses induction over the observations to generate a classification. The expert then evaluates the results and eventually modifies some attributes or some aspects of his domain theory in order to start the classification again. When a correct classification is obtained, fuzzy rules are derived from the generated classes.

4.4.2 The relevance module

A relevance module has been attached to LINNEO in order to guide it in the classification process. This is accomplished in a cooperative way: firstly LINNEO produces a classification in accordance with the expert. Next, using the initial domain theory and the resultant classification, the relevance module produces a list of the relative relevance of the attributes to that classification. Now, this list is subject of an eventual modification by the expert. In addition, a threshold is established in such a way that the module generates a final list in which just a few of the initial set of attributes (25%, say), are present. The classifier then restarts using this new list as the final set of attributes and a new classification is obtained. This is expected to be very close to that obtained resorting to the entire initial list. The attributes not considered are then taken as to be nought.

A further approach has been developed in order to improve the final relevance list. Given a classification, we can apply the relevance module to each single generated class, then considering only those objects belonging to that class. What we obtain is the relative relevance of the attributes when considered as to characterize a subclass of the classification. An interaction with the expert is also possible here for each of the lists of relevance, and local thresholds can be derived. Next, a global list of relevance is formed, by means of merging all the subclasses' relevance lists. How must one interpret these local lists? The first attributes of each list are those more relevant to discriminate between the objects of the class it refers to. The last ones (in order of relevance) will serve to distinguish that class as a whole from the rest. Not surprisingly, these attributes appear at the top of the original global list, and the last attributes of this list appear as very relevant in some (often just one) local list.

Fig. 4: General architecture of DEPUR.

5. Architecture

The architecture of the prototype system DEPUR (see Fig. 4) is a classic one of an expert system, but we have considered some important aspects as the microbiological and temporal ones. The dictionary of the data base is composed by 167 attributes, wit a 15The application groups 79 rules and 12 modules. Also we have obtained 15 metarules, oriented to control and lead the inference process. The modularity of the application, the self-explanation and the user friendly interface proportioned by the shell MILORD are other outstanding properties.

6. Results

We have obtained 10 classes, each one representing good related prototypical situations, that show the most common situations and problems in the working state of the waste water treatment plants. These classes lead the inference process in the diagnosis. The contraposition of system diagnostics against the answers of human experts, has yielded very good results. It has been tested with several prototypical situations in the working of a plant (bulking, foaming, ...). A diagnostic tool based on Knowledge Based Systems for a waste water treatment plant it is a useful approach in the field of intelligent systems when applied to industrial processes.

The new fuzzy techniques for automatic classification represent another important aspect in order to structure the knowledge domain. The treatment of uncertainty in a Knowledge Based System applied to industrial processes control is not very usual. In doing so, DEPUR supposes a main step towards the building of Ecology Preserving Intelligent Systems.

Another interesting point is the possibility to adapt the system to the specific characteristics of the waste water treatment plant that is under control. All in all, we can say that DEPUR is like a fresh air to the Conventional Automatic Control Processes.

In the future we have to reach a second phase: the development of a Knowledge Based Supervisor System, that incorporates classical control techniques with the diagnosis process. We have to extend the domain to sludge line, investigate in a more accurate way the microbiological aspects and make a better temporal information treatment.

7. Conclusions

Doubtless, the construction of a diagnosis tool for controlling a waste-water plant is an interesting contribution to the field of intelligent systems when applied to industrial processes, and can be of great help to the plant manager. The use of new attribute selection methods are also significant in what concerns to the extraction of the expert's knowledge, together with the treatment of uncertain knowledge, expressed and dealt by means of a fuzzy classifier and a module for attribute relevance. The system represents one step towards the construction of systems capable of preserving a natural resource as is water.

8. Suggested future work

Nowadays, research concentrates on the relationship between the notions of *potential utility* and *relevance*. It has been concluded that an attribute having zero as potential utility does not have to be a nought attribute. What is more, reciprocally, it is not necessary for nought attributes to have null potential utility. Moreover, an operational characterization of the *nought* attributes is needed, based on the proposed heuristic of potential utility, all in all leading to a classification process able to clarify & simplify the domain it is being applied to, as is –in our current development– waste-water treatment. In the future, the work has to advance towards the construction, as mentioned above, of an Supervisor Knowledge Based System including diagnosis and control, and integrating numerical control algorithms and heuristical and qualitative knowledge.

9. References

[BAIM88] Baim, P. W. *A method for attribute selection in inductive learning systems.* IEEE Transaction on pattern analysis and machine intelligence 10 (6), (1988) pp. 888–896.

[BELA91] Belanche, Ll. *To be or nought to be. An irrelevant question?.* Master Thesis. Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya (1991).

[FLOR89] Floris, V., *Artificial Intelligence in the operation and management of water resources in South Florida.* Personal Comm. (1989).

[HAYE84] Hayes-Roth, F. *The knowledge-based expert system: A tutorial.* Computer 17 (9), (1984).

[LÓPE90] López de Mántaras, R., Crespo, J. J., *El problema de la seleción de atributos en el aprendizaje inductivo: nueva propuesta y estudio experimental.* In Memorias IBERAMIA 90, 2º Congreso Iberoamericano de Inteligencia Artificial. Ed. Limusa-Noriega. México, (1990) pp. 259–271.

[MAED89] Maeda, K. *A knowledge-based system for the wastewater treatment plant.* A future generation computer systems 5, pp. 29-32. North Holland (1989)

[MART91] Martín, M., Sangüesa, R., Cortés, U. (1991) *Knowledge acquisition combining analytical and empirical techniques.* In Birnbaum, L., and Collins, G. (eds.) Machine Learning: Proceedings of the Eighth International Workshop. San Mateo, CA: Morgan Kaufmann.

[NUÑE91] Núñez, G., Alvarado, M., Cortés, U., Belanche, Ll. *About the attribute relevance's nature.* In Proceedings of the TECCOM-91. México, D.F.

[QUIN79] Quinlan, J. R. *Learning efficient classification procedures and their application to chess end games* In Michalski, R.S., Carbonell, J.G. and Mitchell, T.M. (Eds): *Machine Learning: An Artificial Intelligence Approach.* Tioga, PA, California, (1984) pp. 463–482.

[QUIN86] Quinlan, J.R. *Induction of decision trees* In Machine Learning 1(1). Kluwer Academic Publishers, Boston MA (1989), p81-106.

[ROBU90] Robusté, J. *Modelització i identificació del procés de fangs activats.* Ph. D. Thesis. Departament de Química. Universitat Autònoma de Barecelona (1990).

[SÁNC89] Sánchez, E. *Importance in knowledge systems.* Information System 14 (6) (1989) pp. 455-464.

[SÁNC90] Sánchez, M. *DEPUR: una aplicació de MILORD al tractament de les aigües residuals en una depuradora.* Research report. (LSI 90-12). Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya (1990).

[SÁNC91] Sánchez, M. *DEPUR: Aplicació del sistemes basats en el coneixement al diagnòstic en plantes de tractament d'aigües residuals.* Master Thesis. Departament de Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya (1991).

[SANZ89] Sanz, R., Aguilar J., Sierra C., Godó Ll., Ollero A. *Adaptative control with a supervisor level using a rule-based inference system with approximate reasoning.*. In A.I. in scientific computation: towards secong generation systems. IMACS 1989.

[SCHL86] Schlimmer, J. C., Fisher, D. *A case study of incremental concept induction.* In Proc. of the fifth national conference on artificial intelligence. Ed. Morgan Kaufmann, (1986) pp. 496–501.