T-Air: A Case-Based Reasoning System for Designing Chemical Absorption Plants

Josep Lluís Arcos

IIIA, Artificial Intelligence Research Institute
CSIC, Spanish Council for Scientific Research
Campus UAB, 08193 Bellaterra, Catalonia, Spain.
arcos@iiia.csic.es, http://www.iiia.csic.es

Abstract. In this paper we describe a case-based reasoning application developed for aiding engineers in the design of chemical absorption plants. Based on the notion of flow sheet, the paper describes how the application uses a highly structured representation of cases and similarity criteria based on chemical knowledge for designing solutions following an interactive case-based reasoning approach.

1 Introduction

TECNIUM is an engineering company that designs and sells installations and equipments for gas treatment around the world. The gas treatment is required in many and diverse industrial processes such as:

– *Control of the atmospheric pollution* due to corrosive residual gases which contain vapours, mists, and dusts of industrial origin;
– *Industrial gas purification*, as a stage of a production process;
– *Recovery* of product manufacture, starting from gaseous sources;
– *Degassing* of liquids which contain vapours or gases in solution;
– *Dust removal* of particles mixed in a gas.

Examples of gas treatments are the absorption of gases and vapours such as SO_2, CLH, or CL_2; the absorption of NO_x with recovering of HNO_3; the absorption of drops and fogs such as PO_4H_3 or $ClNH_4$; dust removal in metalic oxides; and elimination of odours from organic origin.

The main problem in designing gas treatment plants is that the diversity of possible problems is as high as the diversity of industrial processes while the experimental models about them is small. The knowledge acquired by engineers with their practical experience is the main tool used for solving new problems. In TECNIUM, the most important presentation card is its more than thirty five years of experience.

An additional issue with gas treatment is that there are many customers sending consultations for analyzing their specific problem but not many of them finally make an order. This issue forces TECNIUM to spend many efforts in designing proposals without profit. Moreover, the proposals have to be analyzed in

D.W. Aha and I. Watson (Eds.): ICCBR 2001, LNAI 2080, pp. 576–588, 2001.

detail because the economical assesments of the offer are binding. For instance, during the last year the amount of offers exceeded thirty millions of euros while the order ratio was around the twenty per cent. This issue motivated the first contacts for providing computer tools for helping in the design of gas treatment plants.

Because the experience accumulated along the years is the main knowledge used in the design of gas treatment plants, the chances for using a case-based reasoning approach were very high. Moreover, the different existing case-based reasoning commercial applications have demonstrated the viability of the CBR approach [9] for solving complex problems and, specifically, for design support [6,7].

Given these initial hypotheses and our previous expertise in developing case-based reasoning systems [2,4] with the use of knowledge intensive methods for modeling similarities in complex structured cases, we decided to start a project of six month duration for developing and testing a CBR application.

For focusing the problem only the chemical absorption (most than 80 %) was covered by the *T-Air* application and we excluded other currently used technologies such as oxi-redox reactions or biofilters. Moreover, the plastic material used for constructing the equipments—they are using more than eight different plastic combinations as well as some metal materials—is determined by the expert except in some predefined problems.

For the design of gas treatment plants about forty different types of main equipments have been covered. We also started with a database of a thousand of solved problems involving, each of them, from two to twenty different equipments.

The organization of this paper is as follows. In Section 2 we briefly present the working principles of chemical gas treatment for illustrating the design process and the design decisions where the use of case-based reasoning techniques can help to the experts. Section 3 describes the *T-Air* application and the case-based reasoning techniques used. Finally, in Section 4 we present the conclusions and future directions of the work.

2 Working Principles of Chemical Gas Treatment

The chemical gas treatment is based on applying to a gas current (the contaminant) another liquid current (the solvent) in a parallel or opposite direction for obtaining a resulting clear gas plus a residual liquid and/or dust useful for another industrial process—or, at least, a liquid that can be eliminated more easily. There are different equipments for achieving this goal each of them based in a different principle such as the mass transfer, the venturi effect, or cyclones. Any chemical absorption plant is designed by determining a *flow sheet* (see Figure 1) composed of a collection of washing elements, tanks for storing the solvents, fans for aspirating the polluted gas, and pumps for circulating the washing liquids. Moreover, for each equipment a specific model and the values of its working pa-

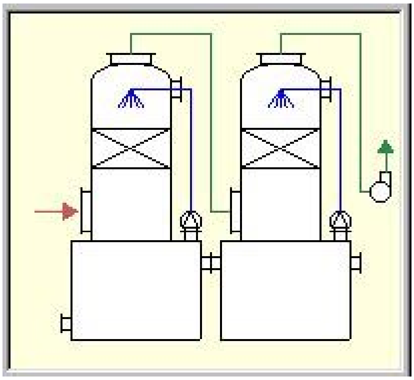

Fig. 1. An example of a simple flow sheet generated by *T-Air* with two scrubbers, each of them on top of a tank with a pump that sucks the washing liquid from the tank to the scrubber, and one fan at the end of the process.

rameters have to be determined. Below we briefly summarize the two main used principles.

Removal of gaseous pollutants by counter-current *Flow Packed Scrubbers* is based on the principle of mass transfer. This is defined as the transfer, in mass, of gaseous polluting molecules from the air stream into the scrubbing liquid. Transfer is achieved by a combination of diffusion, physical absorption, and/or chemical reaction. Highly soluble gases are transferred from an air stream into water by diffusion and physical absorption. Medium or low solubility gases are often scrubbed by absorption, followed by chemical reaction with the treated scrubbing liquid.

The selection of the required equipment and dimensions is determined by two parameters: the number of mass transfer units N_{OG} and the height of transfer unit H_{OG}. The number of mass transfer units is determined given the concentration of the incoming polluted gas and the desired pollution concentration of the outcome gas. The height of transfer unit H_{OG} is determined experimentally depending on factors such as the solubility, concentration, and temperature of the polluting gas. There are only reported H_{OG} values for some prototypical gas-liquid configurations—such as $NH_3 - air - H_2O$, or $Cl_2 - air - NaOH$. For the other cases the value of H_{OG} is determined using the chemical similarity among gases and the experience of previously built installations.

In washers using the *Ventury* effect, the absorption of gas and the collected of particles or vesicles is performed in a gas-liquid parallel current with high energy and turbulence. In the zone of the ventury throat, the liquid is injected in the form of fine droplets which presents a large surface to moisten and facilitates the diffusion of the liquid into the gas. At the same time, an agglomeration and increasing of the mass of the solid particles is produced, being afterwards separated in a cyclonic separator or in a liquid-gas impingement separator, both

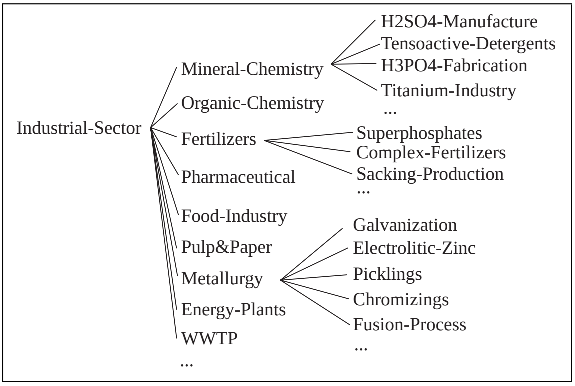

Fig. 2. A portion of the hierarchy for Industrial Sectors.

provided with a high efficiency mist-eliminators of low pressure drop, so avoiding an eventual droplets emission to the atmosphere.

The selection of the adequate configuration of basic equipments for each problem is made bearing chiefly in mind the following data:

- The industrial sector of the customer (see Figure 2).
- Provenance of gas (the industrial process that originates the gas);
- Composition of gas;
- Gas flow to be treated (a range from 100 to 100.000 m^3/h. can be washed with an efficiency of 99 %);
- Temperature (from -10 to 1.000 $°C$);
- Concentration and solubility of the polluting gas;
- Solids and/or vesicles contents;
- Granulometry;
- Pressure drop to be admitted;
- Efficiency or maximum emission desired.

As a summary of this brief description of the working principles of chemical gas treating, we want to emphasize that there are many experimental parameters involved. These parameters are mainly determined from the past experience using the knowledge about the common issues present in different industrial processes and the similarity among the chemical properties of different gases. This was the original working hypothesis for developing an application based on case-based reasoning techniques.

3 System Description

T-Air is implemented in *Noos* [2,5], an object-centered representation language designed to support knowledge modeling of problem solving and learning. *Noos* was developed using the notion of *episodic memory* and provides a collection

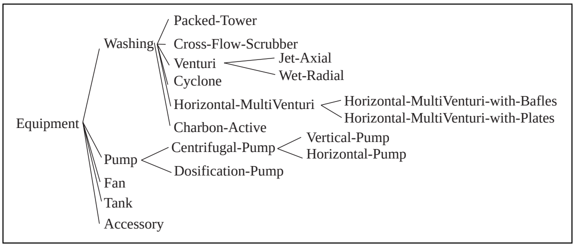

Fig. 3. A portion of the hierarchy for equipments used in Absorption systems. There are four main equipments: gas-washing units, pumps, fans, and tanks. Examples of accessories are spray-nozzles, valves, or demisters.

of components for case-based reasoning such as retrieval perspectives [3] and adaptors [8]. *T-Air* is currently running in Windows computers in the TECNIUM intranet.

In the development of the system two engineers from TECNIUM were involved as domain experts and main users of the system, an additional engineer was consulted for technical issues, and a AI-CBR researcher was the project manager and programmer. The development has taken six months and the utility of the system has exceeded the initial expectatives of TECNIUM. An extension covering the aspects excluded in this initial phase is planed to start in the next months.

We have covered forty different types of equipments organized in five categories: gas-washing units, pumps, fans, tanks, and accessories (see Figure 3). Each equipment has in turn around twenty different models.

3.1 Modeling Chemical Knowledge

In *T-Air*, we have modeled the chemical knowledge that is the basis for determining the similarity between a gas composition in a new problem and those treated previously and stored in the case base. Moreover, since we are mixing a polluted gas with a washing liquid, we also have modeled the properties of the washing liquids, the compatibility among gases and liquids, and some chemical reactions involved.

First of all, the basic substances contained in the periodical table were codified with their main properties such as molar mass, critical temperature, and critical pressure, the groups they belong (metals, non-metals, halogens, ...) and the main properties of these groups.

The second component modeled was gas compounds with their properties such as critical temperatures and pressures or solubilities. The gas compounds are organized in a hierarchy of compounds grouped by chemical notions such as organic compounds, inorganic-compounds, acids, oxides, alcohols, etc. Because the number of possible gas compounds is extremately high, only those present in the case base have been modeled initially. The design decision was that the knowledge about new gas compounds will be acquired when a new problem requires it. The approach we have adopted is close to work of capturing expert design knowledge through "concept mapping" [7]. As we will describe later, when the engineer is analyzing a problem involving a gas not previously covered, the engineer uses the navigation interactive capabilities of *T-Air* for determining the possible characteristics of the gas and improves the chemical knowledge of the system by means of completing the information.

The third component modeled was liquid compounds (solvents). We have followed the same approach taken for gas compounds: only those present in the case base have been modeled. As in gases the navigation capabilities of *T-Air* are used for capturing new knowledge during the analysis of new problems.

The last component modeled was the chemical reactions such as oxidation or reduction. An example of property modeled is the exothermic or endothermic process associated to a reaction.

3.2 Case Representation

Solved problems are represented as structured cases embodying four different kinds of knowledge;

- The *Input Knowledge* embodies data about the customer such as the industrial sector it belongs or the industrial process that originates the polluting gas; data about the working conditions of the installation such as temperature or gas flow; data about the composition and concentration of input polluted gas; and data about the desired concentration in the output emission. Additionally, the customer may constrain the solution desired—for instance proposing a washing liquid that results from another process of the customer.
- The *Chemical Case-Model* embodies a graph structure, generated by *T-Air* using the chemical knowledge, modeling the main characteristics of the input gas. The chemical case-model extends the input knowledge and fixes the thresholds for the working conditions, the main issues to be analyzed, and the kind of gas treatment required. The typology of gas treatment covered by *T-Air* is absorption, dust removal, elimination of fogs and mists, degassing, and elimination of odours. The solution for a specific problem can involve several of them.
- The *Flow Sheet* describes the solution designed for cleaning a polluted gas. As shown in Figure 1, the flow sheet specifies a collection of required equipment (mainly scrubbers, pumps, tanks, and fans), the design and working parameters for each equipment, and the topology of the installation (the gas circuit and the liquid circuits). The flow sheet is also represented as a graph structure.
- *Annotations* are meta-information that describes the reasons for a given decision. Examples of annotations are the design decisions forced by the user requirements (such as the washing liquid), over-dimensionated parameters either because security or unknown input data, different laws in different countries, or spatial requirements. The use of these annotations is important for discriminating different solutions given the similar input information, or for focusing an initial flow sheet when there is unknown data.

An important remark about the cases is the different level of detailed information given in the input knowledge and in the chemical model. This fact is motivated because the different nature of gases. For instance, in gases with highly known solubility with water the exact concentration is not essential for determining a solution. In the opposite side, gases with low solubility or risking parameters (such as originators of exothermic reactions) have to be analyzed in more detail before taking decisions about the flow sheet.

3.3 The Case Base

The Case-Base is composed of three different kinds of cases: performed installations, proposed installations, and pilot experiments.

Performed installations are the collection of designed and constructed installations developed by TECNIUM during the last ten years. The information was obtained from an existing database used by TECNIUM for reporting performed installations. The information contained in the database was incomplete and a the generation of cases involved the participation of the engineers. After, this process, the database was not longer used and *T-Air* is now used for storing all the information about performed installations. Because we are storing all the constructed installations developed during the last ten years, there are cases—such as those involving the use of biological solutions—out of the scope of the current system. Nevertheless, this cases have an special annotation and are used just for presenting the solutions to the engineer and the engineer has to decide all the parameters (only price lists have been included in *T-Air*).

Proposed installations are customer consultations that have been analyzed for offering a solution where there is no customer order for effectively construct them—either because the customer is studying the proposal or because is not interested in making an order. All the proposed installations are stored in the system but only those considered as relevant are used in case retrieval (see subsections 3.6 and 3.7). When a customer decides to order a previously proposed installation, the engineer can retrieve the proposed solution and complete it using the *T-Air* in the same way.

Pilot Experiments are tests performed either by TECNIUM or by universities with some prototypical gas-liquid configurations. The goal of these experiments is twofold: i) collect experimental information by means of systematically studying different range of gas features (pressure, flow, and concentration) in different gas-liquid configurations, and ii) optimize the design parameters of equipments for minimizing dimensions and costs of equipments. This optimization knowledge is mainly used by *T-Air* in the adaptation phase (see subsection 3.5).

Cases are stored in an external database and accessed by SQL queries. The application started with a thousand of performed installations, zero proposed installations, and 200 initial pilot experiments. TECNIUM expects to incorporate five hundred proposed installations in a year where at least one hundred will become performed installations.

3.4 Case Retrieval

T-Air fulfills two different tasks: the analysis of customer consultations and the detailed design of an installation. In the first task the main goal is to determine the cost of a proposal minimizing the risk with the unknown information. When we have to effectively construct an installation, all the parameters in the flow sheet have to be determined even though some of them do not have cost implications.

Because sometimes there is unknown information or some data is only required in some problems, we decide to follow the interactive case-based reasoning approach [1,4]. First of all, an initial form is presented to the engineer and he decides the information initially provided. Given that information the *T-Air* proposes a first solution or presents the alternatives supported by the different retrieved cases and the chemical case-model. Then, the user can introduce more data either by directly filling them or navigating through of the cases or the chemical case-model. At this second stage, the system is working in the background monitoring the decisions of the user and preventing possible issues—like conflicting parameters, risk working conditions of equipments, or uncovered problems—by means of warning messages. An important goal in this stage is to monitor the decisions of the user for acquiring meta-information in the form of annotations.

Like in [10], case retrieval is a two stage process. First, the input data is relaxed to ensure that a minimum number of relevant cases are retrieved from the database. In a second stage, additional similarity criteria are used for ranking the precedents. The process of relaxing the input data uses the hierarchy of industrial sectors or chemical concept hierarchies. For the numerical features we also use the notion of range [10]. Examples of numerical features are the temperature and the gas flow.

The second stage is a process where an initial flow sheet is incrementally refined by using domain criteria, the retrieved cases, and the interaction with the user. This stage is based on the use of *perspectives* [3]—a powerful mechanism for describing declarative biases for case retrieval in structured representations of cases—and on the notion of support sets. Excluding the trivial problem where the user is just designing a new problem that has been previously solved by the system, the usual circumstance is that a new problem is solved by using parts of solutions coming from different cases. In this context the procedure of *T-Air* is to present a flow sheet with different cases supporting different design parameters. For instance, the presence of dust supports the decision of including a ventury in the installation, the kind of industry can support the use of different washing liquids, and the amount of gas flow can support the use of different fan models. After achieving an initial flow sheet with their support sets, the adaptation phase starts.

Let us present a simplified trace of the retrieval phase: the engineer enters a new problem into the system by defining a gas compound generated in a “Fine Chemistry” industrial process with the presence of fogs and a high concentration of sulfates. The first retrieval step, is the identification of the important features of the new problem and the construction of the initial *chemical case-model*. Given the chemical case-model and the input knowledge, *T-Air* performs three initial

SQL queries: i) a first query for retrieving cases about similar industrial sectors using the hierarchy of industrial sectors for generalizing "Fine Chemistry" and ranking the obtained cases using this hierarchy; ii) a second query regarding cases involving sulfates using the hierarchy of gas compounds in a similar way; and iii) a third query is performed using the chemical case-model about the washing processes involved. The final step of the retrieval phase is to build an initial flow sheet with their support sets (the cases supporting design decisions and their similarity with the current problem). It is important to notice that we are performing different SQL queries because there are features of the problems that influence different parts of the design. The goal of the adaptation phase is to check for possible incompatibilities or for overestimated parameters due to the presence of equipments that can help each other.

3.5 Case Adaptation

The goal of the case adaptation is to determine precise models and parameter values for all the equipments involved in the flow sheet. Parameters only known by experience are determined using a collection of domain adaptors developed with the experts that specified when criteria such as mean, maximum or minimum can be used for adapting the values from cases. Moreover we have an equational model for each equipment that allows to determine non-experimental values using the values from cases. Finally, the adaptation knowledge is also charged of determining possible conflicts between different support sets.

The construction of a new solution for a given problem by combining parts of the solutions of different cases is performed in the following way: First of all we have to determine the parts of the solution in each case relevant for our problem. For instance, if one of the tasks in a given problem is the problem of dust removal, using the chemical knowledge and the models about the equipments, we can determine the equipments relevant for this task. Then, given the collection of partial solutions to our problem, and our adaptation knowledge about the general principles of design of absorption plants, we connect all the partial solutions with a high support set, each of them solving a part of the problem, and link each of them to the possible alternatives (those partial solutions with a low support set). After that, the process of refinement of the solution starts by determining possible conflicts between different support sets or the presence of equipments that can help each other. For instance, increasing the dimensions of a given scrubber we can eliminate another one.

An important remark about partial solutions is that *T-Air* presents only one flow sheet with bullets pointing to the engineer the places where there are alternatives and the engineer can inspect and select some of them. The final decision is performed by the engineer and the goal of the adaptation phase in *T-Air* is to present a flow sheet and the possible alternatives and provide an easy mechanism for allowing the final design decisions.

Cases retrieved from pilot experiments are mainly used in the adaptation phase. Pilot experiments allow to optimize the values of equipment parameters—especially in experimental parameters—for minimizing dimensions and costs. For instance, when there is a five years old case supporting a value of 0.6 for H_{OG}

for the same gas in the problem, and a two years old case pilot experiment supporting a value of 0.4 for H_{OG} with a similar gas, taking into account the degree of similarity, we can decrease the final value to 0.5.

3.6 Case Storage

All the installations designed by the system have to be stored because *T-Air* is also the repository of designs proposed/performed by the company. Moreover, when a customer makes a consultation not covered by the system, the engineers want to explore the existing case base in an exhaustive way using a collection of criteria such as navigating through the chemical hierarchy an inspecting to the solutions designed in the cases.

The current storage policies in *T-Air* are very simple: i) preliminary consultations are not used in retrieval until the experts marks them as fully analyzed; and ii) when there are several equal cases solved with an equivalent design, only the latest is indexed for retrieval with an additional annotation that marks the number of equal cases. The experience with the use of the system also confirms our initial hypothesis that cases representing proposed installations with many unknown values are rarely retrieved.

The case storage is an open issue in *T-Air* that requires further analysis. Nevertheless, the main issue now is not an efficiency problem. Discussing with the experts, the main problem is how to minimize the amount of information that the engineer has to manage for identifying the important issues in each problem.

3.7 User Interface

T-Air has been developed in *Noos* for Windows computers running in an Intranet. An important issue with the graphical interface was to find the most appropriate way of displaying the design parameters because of the high amount of data and taking into account that the expert was interested in modifying any of them.

Another important issue was to provide useful navigation tools for inspecting the chemical knowledge and the case base. This requirement is crucial when the expert is trying to design the solution for a problem that is not covered with the current knowledge of the system. In those cases, during the testing phase we had to tune some interactive mechanisms of acquiring new chemical information.

An example of the procedure followed by the expert with a not covered problem is the following: the expert starts navigating through the hierarchy of industrial sectors, then follows the hierarchy of industrial processes involved and then, the kind of gases previously treated. Suppose that in our hypothetical problem we have acids from organic origin, the important properties of these gases are listed and the expert consults to his bibliography about our specific problem. This information is introduced in the system and the retrieval starts. After the

[1] The system interface has been developed in Catalan but for helping to the comprehensibility of the system the labels have been translated to English.

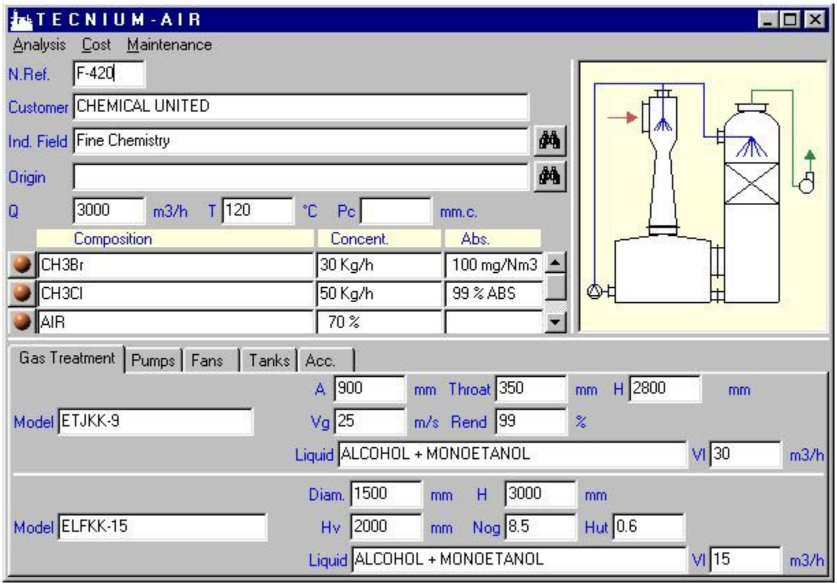

Fig. 4. A snapshot of the T-Air system[1].

retrieval, either there is an initial proposal or the expert has to determine an initial flow sheet. At this point the interaction between *T-Air* and the expert is the same as in covered problems. A sample screen of the *T-Air* application is shown in Figure 4.

4 Conclusions

The utility of the *T-Air* system has exceeded the initial expectatives of TECNIUM. The first indicator of success was that we advanced the delivery of a first prototype of the system because just with some preliminary similarity criteria and the possibility of manipulating the design parameters the experts were able to decrease the time for analyzing customer consultations. A second measure of success is the request of connecting the information provided by *T-Air* with other existing applications used for construct the installations. For instance, since the engineers have determined the elements and many dimensional features of an installation, this information can be used by analysts that elaborate the detailed list of required materials and can be also used by the draftsman responsible of the technical drawing. The last successful criterion is that we are planning an extension of the application for covering the aspects not addressed in this initial initial phase.

A side effect of the use of *T-Air* is that now all the proposed installations are stored in the TECNIUM server and not only the information about performed installations as before. This is important for the performance of *T-Air* but is also very important for the company.

Another important aspect discovered during the field testing phase is the facility that the system provides to the engineer of changing any design parameter. Moreover, this changing facility has increased the reliance on the use of system and has demonstrated that for many problems the expert is not really changing them. Moreover this feature allows engineers to test more than one design alternative for a given problem, capability that was impractical without the use of the system because of the high amount of information and calculus required for designing each plant.

Related with the use of meta-information as annotations, the initial approach was to ask an explanation to the engineer when some decisions monitored by *T-Air* were not supported by the current knowledge. Nevertheless, this was demostrated a wrong strategy because we are trying to minimize the design time and the engineer is usually busy in many issues at the same time. The current approach we are exploring is how to periodically analyze the knowledge of the system for detecting these unsupported decisions and develop a specific tool for this purpose.

Acknowledgements. The research reported in this paper was supported by the *T-Air* project financed by the TECNIUMcompany. Any possible mistake or error in the description of the working principles of chemical gas treatment is the sole responsability of the author.

References

1. David W. Aha, Leonard A. Breslow, and Héctor Mu noz Avila. Conversational case-based reasoning. *Applied Intelligence*, 14(1):9–32, 2001.
2. Josep Lluís Arcos. *The Noos representation language*. PhD thesis, Universitat Politècnica de Catalunya, 1997. online at `www.iiia.csic.es/~arcos/Phd.html`.
3. Josep Lluís Arcos and Ramon López de Mántaras. Perspectives: a declarative bias mechanism for case retrieval. In David Leake and Enric Plaza, editors, *Case-Based Reasoning. Research and Development*, number 1266 in Lecture Notes in Artificial Intelligence, pages 279–290. Springer-Verlag, 1997.
4. Josep Lluís Arcos and Ramon López de Mántaras. An interactive case-based reasoning approach for generating expressive music. *Applied Intelligence*, 14(1):115–129, 2001.
5. Josep Lluís Arcos and Enric Plaza. Inference and reflection in the object-centered representation language Noos. *Journal of Future Generation Computer Systems*, 12:173–188, 1996.
6. D. Hinkle and C. Toomey. Applying case-based reasoning to manufacturing. *AI Magazine*, 16(1):65–73, 1995.
7. David B. Leake and David C. Wilson. Combining cbr with interactive knowledge acquisition, manipulation and reuse. In *Proceedings of the Third International Conference on Case-Based Reasoning (ICCBR-99)*, pages "218–232", Berlin, 1999. Springer-Verlag.
8. Enric Plaza and Josep Lluís Arcos. Towards a software architecture for case-based reasoning systems. In Z. Ras and S. Ohsuga, editors, *Foundations of Intelligent Systems, 12th International Symposium, ISMIS 2000*, number 1932 in Lecture Notes in Artificial Intelligence, pages 601–609. Springer-Verlag, 2000.

9. Ian Watson. *Applying Case-Based Reasoning: Techniques for enterprise systems.* Morgan Kaufmann Publishers Inc. San Francisco, CA, 1997.
10. Ian Watson and Dan Gardingen. A distributed case-based reasoning application for engineering sales support. In *Proc. 16th Int. Joint Conf. on Artificial Intelligence (IJCAI-99)*, pages 600–605, 1999.