Journal of Chromatography, 485 (1989) 341–347
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 22 068

Note

Cathie: expert interpretation of chromatographic data

ROBERT MILNE
Intelligent Applications Ltd., Kirkton Business Centre, Kirk Lane, Livingston Village, West Lothian EH54 7AY (U.K.)

Gas chromatographs are now applied in almost all areas of manufacturing and process control and their use represents substantial benefits to the end user in the form of both product quality monitoring and overall safety monitoring. Many essential tasks in process monitoring can only be accomplished by using a gas chromatograph to verify the results. Currently, however, these instruments are usable only by trained chemists. This greatly restricts the number of companies that can use these instruments and the uses to which they are put.

In order to enable these instruments to spread more widely throughout industry and benefit a larger group of companies in a much wider way, a method is needed to make the monitoring and interpretation of the instrument results easier. An expert system has been developed that is capable of serving as a tool kit for interpreting the results from a gas chromatograph for a particular process, in addition to providing an expert analysis of the state of the instrument in order to detect possible failures or deteriorations. The system, Cathie, was developed based on the earlier Violet system (Intelligent Applications, Livingston, U.K.). Violet provides for expert interpretation of rotating machinery information. An early prototype of Cathie was developed by modifying the existing Violet system.

In this paper, first the basis of expert systems and how they may be relevant to the interpretation of chromatographic systems is discussed. Expert systems provide a means of using knowledge to make interpretive decisions. By abstracting the data, one can write interpretive rules in the same simple language that a chemist would use. The functionality of Cathie is then discussed in more detail, describing the basic steps that it needs to achieve. A real application is described to illustrate how these techniques are applied.

THE USE OF EXPERT SYSTEMS

With the integration of instruments and computers comes the desire to have the computer perform data reduction and processing and to participate in the decision process concerning the status of the machine in question[1–3]. The use of artificial intelligence (AI)[4] techniques is a natural extension of that trend. Within the field of AI, many systems have been built that perform sophisticated diagnosis and status monitoring tasks[5].

AI involves the use of knowledge to solve problems. In order to solve many

0021-9673/89/$03.50

problems, a diverse set of knowledge is needed. This may be simple information such as alarm thresholds, or complex information such as the causal sequence of how the main process works. The difficulty with traditional methods of programming is representing and manipulating that information. AI provides tools to represent and manipulate that knowledge more easily and effectively.

To date, systems have been built around a rule-based framework on fairly expensive computers[6]. There are two important trends that are now emerging: the availability of knowledge-based systems and knowledge-based system tool kits on smaller and cheaper machines, and a more direct coupling of knowledge-based systems with test and measurement instruments.

Today, most knowledge-based systems for fault diagnosis or data interpretation require the user to read values from various test instruments and manually type them into the knowledge base. When the knowledge-based system needs more information, it requests the user to measure a particular set of values. The user then turns to the device, makes some measurements on the instrument and types the values into the machine. With current computer and interfacing technology, the requirement to enter these data manually will soon be superseded.

Users of test and measurement instruments represent a wide range of sophistication, from those that desire simple turn-key systems to those that are capable of using very complex instruments and developing their own complex diagnosis and status monitoring programmes. However, whatever the level of experience of the user, they all require the same foundation, that is, knowledge-based system properly interfaced with a test and measurement instrument, which allows the users to write their own rule bases.

Once such a system has been developed, several types of rule bases can be written. For example, a rule base could be supplied to perform quality checks on one specific product such as coffee, another can be supplied which advises the user on how best to use the instrument to measure properly and effectively whatever the application is.

AN EXAMPLE

When the chemist uses a complex test instrument, he (or she) thinks in a certain "language" to describe the display that he is examining. As a simple example, he can consider the chromatogram from a sample of coffee. He is most concerned with some of the key compounds and their concentration. A chemist may think:

> "IF: the concentration of nitrogen is *high*
> and the concentration of helium is *high*
> and the concentration of hydrogen is *low*
>
> THEN: there is a possible breach in the pressure vessel".

Although this is a simplified example, it helps to illustrate a case where knowledge-based systems could be used. The primary task is to write the rules for the knowledge base which can detect this condition automatically. There are two primary steps which must be conducted. First, the rules or knowledge base must be built up. Although this is never simple, it only requires traditional AI knowledge engineering.

The second and more unique step is that the knowledge-based system program needs automatically to be able to sense where the compounds are and how they change when the system changes. For our simple example, we would need the interface software to detect the location of the peaks, in addition to the system state[7].

To make the connection complete, it is necessary to provide the software to allow the chemist to describe his problem in terms of his language. That is, the ability to compute the commonly required parameters and values must be provided. Each common term will have a direct way in which it is acquired and updated.

From the engineer's viewpoint, then, he can describe to the software his problem and how the parameters he measures lead to various specific problems, and the system will then automatically measure these[8].

BENEFITS OF THE EXPERT SYSTEMS

Before a new technology is introduced, it must provide clear benefits to address needs in the existing process. Process monitoring systems directly address many limitations inherent in the requirement for a human to monitor the system. Although the process parameters are relatively simple to see at a glance, it is rare that a human operator can stand continuously over the system, monitoring all the parameters. Very small systems normally do not justify constant human attention. Very large systems require constant human attention, but often contain thousands rather than hundreds of parameters[9].

Software has been developed to provide monitoring before. This software often requires an experienced person to develop and is very inflexible. A process control chemist does not want to spend time learning how to develop complex software, rather he wants to be able to adjust set points quickly and easily, obtain an overall picture of the operation of the system, or provide changes in the interpretation of the process. Many of the monitoring situations require the knowledge of the process that the chemist has to use in conjunction with the process parameters. An expert system for process monitoring addresses these points directly. Rather than a person being required to monitor, the expert system has the capabilities to imbed his intelligence into evaluating situations and recommending courses of action[10].

In order to develop a monitoring system, the knowledge used by an experienced chemist is often necessary. The problem is how to represent this knowledge easily in a computer and how to manipulate it. This is precisely what the expert system rule language provides. There is also a very strong need to be able to adjust the software rapidly and easily. Conventional computer programming languages make this far too difficult to be practical on a wide basis. Expert systems, with their interactive editors and English-like rule languages, provide a very simple means to meet this objective. Although we do not pretend that the programming changes are trivial at present, it is certainly as easy as possible to build up rapidly rules of interpretation or to modify the way the system is working[11]. The expert system provides a high-level abstraction of the process which provides the chemist with a much easier means of controlling the interpretation and monitoring program.

CATHIE

The problem for the end user is to be able to identify rapidly whether a display represents the proper combination of components and to predict how much time is left until failure of specific components. It can be a very difficult and time-consuming task to identify each particular component, compare it with its expected value and then provide an overall interpretation of the pattern of high levels and normal levels.

Typically, this routine task requires highly skilled people. Many companies desire to automate this so that it can be done autonomously or with low levels of training. Previous work on rotating machinery bears many similarities to the output of a gas chromatograph, and this experience is now being applied to analytical chemistry.

In order to describe a chromatogram to another person, one speaks in terms of peaks at elution times and their area or concentration. There are specific problems associated with where these peaks might occur, based on the state of columns and the method being used. In order to develop an interpretation of the software, one can think in terms of rules of interpretation based on descriptions in this high-level form. That is, one might say that: "IF the acetone is at normal concentration and the C_8H_{16} is at a high concentration, THEN not enough of stabilizer was added to the initial product". The goal of the system we are developing is to allow the user to write rules of this form and let the software extract the peak values and make comparisons for high or low values. The basic approach is to start with a list of expected compounds, their elution times and concentrations and produce an output to the expert system stating what is actually found. Finally, the expert system will run the rules to produce an interpretation.

The primary needs of current and future users of gas chromatography fall into two categories: Interpreting the output of the instrument to verify whether the sample is as expected, and verifying that the instrument is working properly. The interpretation of each sample is dependent on the end process and, as a result, the product needs to provide a "tool-kit" capability to each end user. The verification that the instrument is working properly, however, is not dependent on the variation of the exact process. This can be provided as a single expert system for all end users. The output of a gas chromatograph is a series of peaks, the height of each peak representing the concentration of the sample and the time at which the peak elutes representing the actual compound. A particular compound should always appear at the same retention time assuming the instrument is working properly. In order to verify that, for example, the whisky produced in the current batch is the same as it should normally be, it is necessary for the operator to compare the retention time and concentration of each peak with a standard manually. There are a number of software packages available today that will identify the names of the peaks under normal operating conditions. The comparison with the standard is still largely a manual process, this can be very time consuming and introduce errors. In order to make the system much easier to use, lower the skill level and be performed much faster, it is necessary to develop software that is capable of checking for the presence of the desired peaks at the expected retention time and concentration. The chemist has not only to identify whether the peaks are as expected, but also to determine what actions to take if they are not correct.

We have developed a subsystem that has the ability to locate and identify the peaks and provide an output describing whether each peak is above or below the expected concentration[12]. All of the peak outputs for an individual chromatograph can then be processed by the expert system. The current Violet system does precisely this for the fast Fourier transform data produced from rotating machinery. A menu-driven facility is available to identify each peak, its frequency and the expected amplitude of the vibration, and the software is able to match all the peaks expected with the current spectrum, then produce a summary of any deviations. The summary is then used by the expert system to produce a more detailed interpretation. The initial application of Violet was in helicopter rotor blade analysis. It is currently being used by British Petroleum to examine the vibration of pumps and compressors in the North Sea[13].

The peak finding must be robust, as the retention time may vary slightly according to the state to the instrument and the concentration may vary within acceptable tolerance limits. The menu-driven capability to describe the peaks must also be able to specify a variance in the retention time as a percentage of the expected time, as well as the variance in the concentration. In general there are more peaks in a sample than a chemist will actually use for interpretation, so it will be necessary for the software to ignore specific components. If major peaks appear when not expected, this could have very significant implications, and as a result it will be necessary for the software to detect any large unexpected peaks and the expert system to be able to deal with those directly.

Earlier the Violet software was modified to work with a Perkin-Elmer gas chromatograph instead of a spectrum analyser, with limited success of the above task. In order to enhance the system, the peak matching and the ability to deal with tolerances were improved. Currently the vibrations are of a fixed frequency range whereas chromatographic peaks can vary over a long time period. Further development was needed to handle variable length chromatograms. Better facilities were also added to provide basic or training examples, minimizing the operator development time. A considerable library of chromatograms has been collected that include a number of situations where ambiguity will arise in determining which peak is associated with which compound. This is largely because of the very close retention times for a number of peaks. A software module was developed that can bring into the expert system the descriptions of peaks in a broad region and then develop a set of expert system rules for dealing with the common cases of ambiguity. This greater accuracy in means of peak matching will greatly enhance the utility to the end user and represents a significant advance over existing software. Not only is knowledge used in the interpretation, but also in the peak identification.

Although there are a range of problems that can occur in a gas chromatograph, the most common is the deterioration of the column. When the column deteriorates, two major changes take place: the shapes of the peaks themselves will change, and the retention times of the peaks will be reduced. Although this reduction in retention times is roughly uniform, it is non-linear, and very difficult to predict. A gross deterioration of the system prevents most peak finders from being able to identify any of the compounds. However, an expert chemist can rapidly identify that a deterioration has taken place and still locate the compounds. By providing a robust means for the expert system to manipulate the data, capturing this experience and expertise is relatively straightforward and process independent.

In order to provide full functionality to the end user, additional support functions have been developed. The system was required to have the capability to analyse deterioration trends over time periods of weeks, days and months. This required a disk-based trending system. The system needed to have the ability to cope with irregular time measurements and missing data points for periods when an analysis was not conducted. The ability to detect the deteriorations was developed around logarithmic changes rather than linear changes.

It was necessary to develop software to look for uniform deviations. Techniques from pattern recognition and operations research can be used to understand how a proposed deviation may explain the changes in the retention times. In some systems, the changes in the rise time of the peak will be detected. In other outputs the resolution will not be adequate to discover this clue.

This problem has been attacked at two levels, providing software that is able to detect a uniform shift in the data and using pattern-matching techniques to identify the best fit of the retention times and the expected values. This is also being attacked by using the expert system to capture the expertise of senior chemists in detecting how the deterioration takes place.

APPLICATION

The South of Scotland Electricity Board is very concerned about the difficulties of constantly interpreting chromatographic data and required a system for their Torness Nuclear Power Plant. Lessons learned in that development have indicated how best to proceed to a much broader general-purpose system.

The Torness Nuclear Power Station has five on-line chromatographs, each collecting a small selection of gases. These gases are combined and passed through a communications link to a PC, where the expert system resides. The expert system constantly acquires the data from each instrument and combines these into a summary of the current state of the instruments. It then uses the knowledge-based peak matching to identify the current retention time and concentration for each of the compounds of interest. A rule base then examines this combination of compounds to detect faults such as ingress of air into the pressure vessel, a breach of the pressure vessel, oil creeping into the pressure vessel, excess of water or other changes.

The system must interact with the operators at times, *i.e.*, to detect when the power level of the reactor has changed or helium has been injected into the pressure vessel. Not only does the resulting system provide for continuous interpretation, but also provides a framework to assist the chemist in understanding the gas chemistry within the pressure vessel.

The expert system was developed through a number of carefully structured discussions focusing on the available compounds and what they indicate. The resulting system is currently undergoing field trials.

BENEFITS TO THE USER

The benefits to the end user fall in two broad categories: increased speed and lower skill levels required. A typical speed-up will reduce a 20-min task to only 1 min.

For a company having a high throughput of analyses per day this can lead to considerable personnel savings. The system will also enable many people to use gas chromatographic systems who have not been able to before because of the level of training and experience needed. Although it is more difficult to quantify this benefit, it will be substantial in many companies. Organizations using the resulting system will normally produce a higher quality product and be able to maintain much higher standards of quality control.

CONCLUSION

There are two main aspects to the system Cathie, which uses expert system technology to provide for the automatic interpretation of chromatographic data: providing a more flexible means to identify and match peaks in a normal chromatogram so that their pattern can be interpreted by the expert system, and providing the mechanism to detect deteriorations in the instrument. The peak identification in the first step needs to be able to take account of any deteriorations by the second step. This capability is not present in any commercial systems today. It is envisaged that the extra reasoning power available from the expert system will make this straightforward.

By providing an effective means to allow the chemist to express his knowledge about matching the peaks, detecting instrument problems and providing an interpretation, a very powerful system has been developed. Although still in its infancy, this technology has the potential to make a considerable impact on the future of chromatographic interpretation.

REFERENCES

1 R. Hanna, paper presented at *the Eastern Analytical Symposium, New York October 2–7, 1988.*
2 G. C. Levy, paper presented at *the Eastern Analytical Symposium, New York, October 2–7, 1988.*
3 B. Wythoff, paper presented at *the Eastern Analytical Symposium, New York, October 2–7, 1988.*
4 P. Winston, *Artificial Intelligence*, Addison-Wesley, Reading, MA, 1984.
5 P. Harmon, *Expert Systems Strategies*, Newsletter of Cutter Information Corp., Vol. 2, No. 8, August 1986.
6 A. Barr and E. Feigenbaum, *The Handbook of Artificial Intelligence*, William Kaufmann Press, Los Altos, CA, 1989.
7 R. Milne, paper presented at *the 7th International Workshop on Expert Systems and Their Applications, Avignon, May 13–15, 1987.*
8 R. Milne, *IEE Proc., Part D*, 134 (1987) pp. 238–244.
9 R. Moore, paper presented at *the 1st International Conference on Applications of Artificial Intelligence to Engineering Problems, Southampton, April 15–18, 1986.*
10 W. Clocksin and A. Morgan, in *The 7th European Conference on Artificial Intelligence, Brighton, July 21–25, 1986*, European Conference of Artificial Intelligence, pp. 350–356.
11 K. Forbus, *Qualitive Process Theory*, MIT Technical Report, No. 789, 1984.
12 R. Milne, paper presented at *the Eastern Analytical Symposium, New York, October 2–7, 1988.*
13 R. Milne, *Chart. Mech. Eng.*, 33 (1986) 45–46.