Bioprocess Engineering 7 (1992) 331 - 334

Bioprocess Engineering

A knowledge based system for diagnosing microbial activities during a fermentation process

M. Pokkinen, Z. R. Flores Bustamante, H. Asama, I. Endo, Wako-shi, Japan, R. Aarts, and P. Linko, Espoo, Finland

Abstract. A knowledge based system, LAexpert, was developed to diagnose microbial activities during a fermentation process on the basis of specific rates determined on-line. The LAexpert is a supervisor for a process control system and assists operators in fault diagnosis. The LAexpert was implemented using a fuzzy expert system shell based on the object oriented programming tool Smalltalk /Mac running in a Macintosh II computer. The shell can handle uncertainties both in the measurements and knowledge by fuzzy reasoning.

List of symbols

X	g/l	biomass dry weight (g/l)
S	g/l	substrate concentration (g/l)
P	g/l	product concentration (g/l)
μ_c, ν_c, π_c	1/h	specific rates calculated from on-line measured data of X, S and P (1/h)
μ_d, ν_d, π_d	1/h	specific rates read from database of BIOACS (1/h)

1 Introduction

To measure and control the microbial physiology inside a bioreactor has for many years been a goal in fermentation research. On-line monitoring and control of the cell mass, and product and substrate concentrations is the classical approach to this problem. Recently, Endo et al. [1, 2] developed the Bio Advanced Control System (BIOACS) for automatic monitoring and control of fermentation processes, and Pokkinen et al. [3] applied the BIOACS to diagnose microbial activities during lactic acid fermentation based on specific rate equations.

It is only a part of the solution to enable the measurements of the key parameters concerned with microbial metabolism. To understand and diagnose the complex microbial activities in a fermenter based only on some physical and chemical measurements needs special expertise. As a matter of fact, operators need to make fast decisions based on a few and sometimes contradictory measurements during the fault diagnosis in the fermentation industry. Expert systems that assist operators have proven to be useful in such applications. Aarts et al. [4] developed a fuzzy expert system shell, BIOTALK, based on object oriented programming environment, Smalltalk, and applied it for enzyme production control. In this study, the shell was modified and a knowledge based system for diagnosing biological faults of a lactic acid fermentation process based on on-line measurements by using the BIOACS was developed.

2 The expert system shell

An expert system consists of a shell and the knowledge base. The shell includes the user interface, reasoning control mechanism, mechanism for automatic fault detection, and a tool to construct the knowledge base. The shell components can be designed to be used in several different applications, but the specific knowledge base has to be formulated for each case. This applies especially to bioprocesses in which the characteristics of microorganisms and process conditions vary widely.

In the present work the expert system was implemented using a shell developed by Aarts et al. [4] with object oriented Smalltalk/V286 on a WYSE 211, IBM PC-AT compatible computer. The shell was extended and transferred to Smalltalk/V Mac on a Macintosh II computer with 5 MB of main memory and 100 MB hard disk. The object oriented programming makes it easy to modify the system and to add new functions to the shell.

The shell can handle uncertainties by fuzzy reasoning both in the measurements and in the knowledge. The advantages of applying fuzzy reasoning in bioprocesses have been discussed by Linko [5]. Fuzzy theory can deal with process variables in linguistic terms such as high, very high or low, and facilitates the transfer of subjective human expert knowledge and the handling of inprecise data. When the shell is used for a new application the process variables are characterized by fuzzy membership functions, ideal profiles, set point values and tolerances from set points [4]. In this work fuzzy membership functions, e.g. "pH is low" were defined as trapezoidal membership functions for each pro-

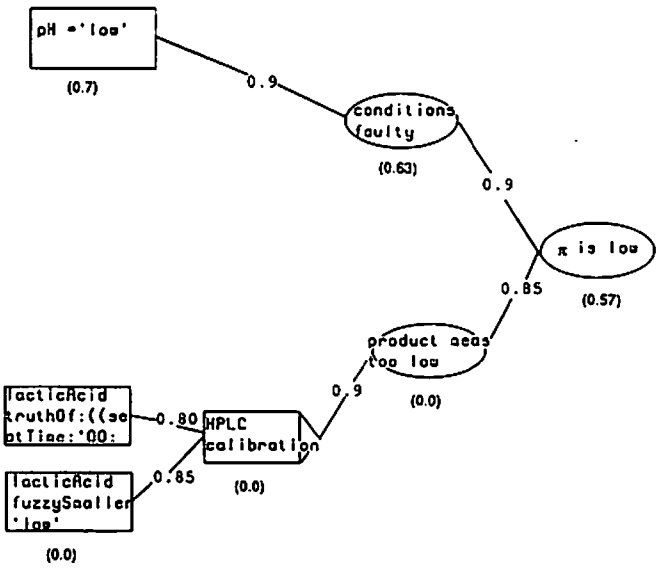

Fig. 1. A part of the knowledge network of the LAexpert. The truth value of each node is shown in brackets. When a fault "π is low" is found the LAexpert checks all the chains connected to this end node. The chain that has the highest total truth value gives a reason for the fault. In this case start node "pH is low" has a truth value of 0.7. Multiplying the truth value (0.7) and certaintly factor of the connection between nodes (0.9) the truth of the fact "conditions are faulty inside the fermentor" (0.63) is obtained

cess variable, and so the truth of the fact "pH is low" can be a value between 0 to 1. The definition of the fuzzy set for the variables with the ideal profile was done by using the variable's deviation from the ideal profile instead of the actual values of the variable. For the fuzzy values – operators not, and, or and probability-or – are defined in the shell according to Aarts et al. [4].

Defining the fuzzy sets and characteristics for each variable yields a large part of the knowledge in the system. Main parts of the knowledge were represented as a network of nodes as illustrated for an example chain in Fig. 1. The nodes stand for a fact in the fermentation. Different types of nodes referring to fuzzy operators – and, or, probability-or – are available for knowledge implementation. Each pair of connected nodes represents an "if-then" rule. For each connection between the nodes a certainty factor to express the degree of relational strength was defined when the network was built. The two example chains from the network shown in Fig. 1 could be written as follows:

1) If pH is low, then conditions inside the fermentor are faulty and due to this π is low.
2) If both the initial and the current lactic acid concentrations are lower than 'low', then HPLC calibration is 'low', and due to this the product measured is 'low' and thus π is 'low'.

In Fig. 1, the rectangular nodes on the left are called start nodes. The oval nodes are "or nodes" and the pointed rectangulars are "and nodes". The right most nodes are defined to be "end nodes". They represent detected faults. For each "end node" a fact, e.g., sentence "π is low" is written in normal language. The causes for a fault form a chain that is connected to left side of the "end node". The facts of other nodes in the chain are also written in normal language. The last node in this chain is a "start node". The fact of a "start node" is written in Smalltalk syntax using messages to automatically communicate and to continuously evaluate the fact in context with the knowledge-based system. The shell has a user friendly graphical interface that can be operated by mouse. The knowledge network also can be easily edited due to the graphical representation (Fig. 2).

3 The lactic acid fermentation expert system (LAexpert)

A batch lactic acid fermentation of Lactobacillus casei [3] was selected as a case study for the development of the LAexpert. The diagnosis of microbial activities was based on the specific rates of cell growth, substrate consumption and product formation.

During a fermentation run the LAexpert received from the BIOACS calculated (μ_c, v_c, π_c) and standard (μ_d, v_d, π_d) specific rate values. The latter values were estimated from the database curves according to each new substrate measurement. LAexpert first determined the phase of the fermentation for each specific rate. Secondly, LAexpert compared the two input values of each specific rate. If the values calculated from the on-line data (μ_c, v_c, π_c) differed from the standard values (μ_d, v_d, π_d), the fault diagnosing engine in the LAexpert was activated. The fault diagnosis is also activated if the phases of the calculated specific rates differ from that in the database. In the fault diagnosis, all of the "end nodes" are assumed as candidates of root causes of the fault in the knowledge network. When the most reasonable "end node" is found, the diagnosing system checks all the chains of nodes connected to this "end node" backward to find the chain that has the highest total truth value. The truth value of a chain is calculated by multiplying the truth value of the node with the certainty factor of the connection between nodes. The system displays the fault observed, its truth value, and its causes to the operator. Further, explanations and instructions about the fault detection process are given.

An example of the fault diagnosis report is shown in Fig. 3. In this fermentation a fault "π is low" was found at time 6 hours. The LAexpert then examined all the chains of nodes connected to this "end node" (Fig. 1). In this case, the reason for the fault was "pH is low" with a truth value of 70%. In other words, the pH at time 6 hours belonged to the fuzzy class "low" with the truth value of 70% according to current pH measurement. The fact "conditions inside the fermentor are faulty" was 63% true (70 · 0.9), and fault "π is low" was 57% true (63 · 0.9). Because the two other "start nodes" in Fig. 1 had truth values of 0.0 according to the current lactic acid concentration measurement, the first chain had the highest truth value.

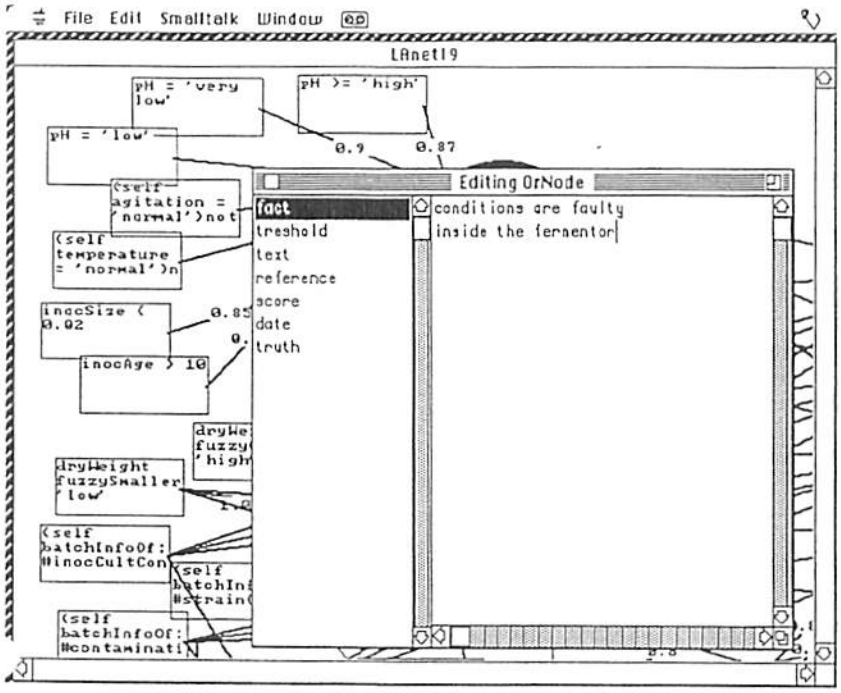

Fig. 2. The user interface of the LAexpert when editing the knowledge base. The statement fact for each node can be edited through a window

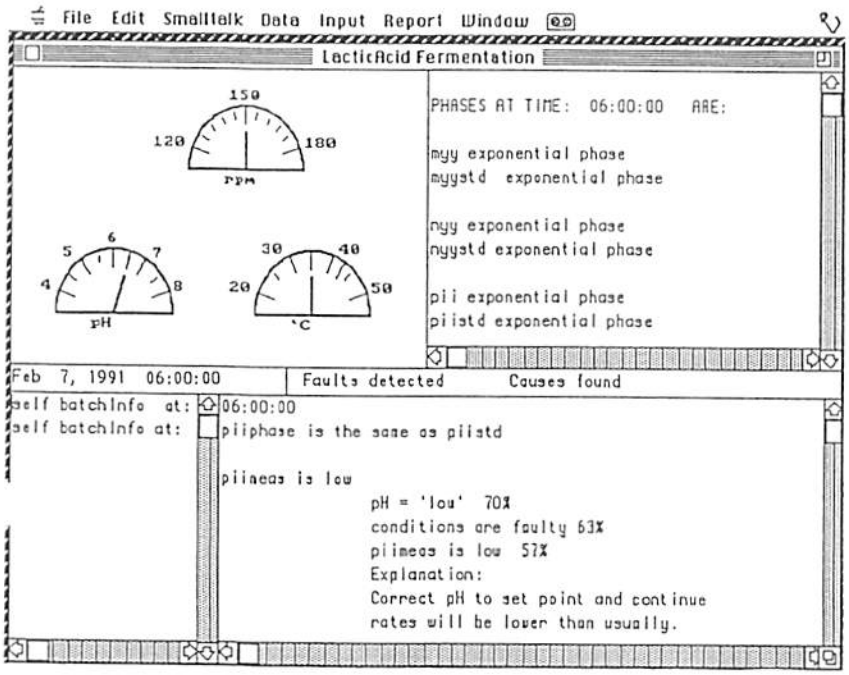

Fig. 3. The user interface of the LAexpert during the fault diagnosis

4 Conclusions

The knowledge based system, LAexpert, was developed to diagnose lactic acid fermentation on the basis of the specific rates calculated on-line by the BIOACS. In the hierarchy, the LAexpert is considered to be a supervisory system of the BIOACS. The LAexpert was tested with real fermentation data. It proved the ability to determine the phases of the fermentation and to detect faults that were not deducible from off-line measured analytical data. With the help of the LAexpert, faults could be detected and corrected in an early stage.

The knowledge base is the most important part of the expert system. The knowledge acquisition is continued to improve the capabilities of the system in explaining the microbial activities.

Acknowledgements

The authors are grateful to the Science and Technology Agency (Japan), Ministry of Trade and Industry (Finland), Foundation for Biotechnical and Industrial Fermentation Research (Finland), Foundation for Technical and Commercial Sciences (Finland) and The Finnish Cultural Foundation for financial support.

References

1. Endo, I.; Asama, H.; Nagamune, T.: A database system and an expert system for realizing factory automation in the bioindustries. In: A. Fiechter, H. Okada, R. D. Tanner (Eds.): Bioproducts and bioprocesses, pp. 337–345. Heidelberg: Springer-Verlag Berlin 1989
2. Endo, I.; Nagamune, T.: A database system for fermentation processes, Biopr. Eng. 2 (1987) 111–114
3. Pokkinen, M.; Flores Bustamante, Z. R.; Asama, H.; Endo, I.; Linko, P.: Diagnosing lactic acid fermentation based on specific rates of growth, substrate consumption and product formation. Biopr. Eng. 7 (1992) 319–323
4. Aarts, R.; Suviranta, A.-M.; Rauman-Aalto, P.; Linko, P.: An expert system in enzyme production control. Food Biotechnol. 4 (1990) 301–315
5. Linko, P.: Uncertainties, fuzzy reasoning and expert systems in bioengineering: Ann. New York Acad. Sci. 542 (1988) 83–101

Received May 27, 1991

M. Pokkinen (visiting scientist at Riken)
Soft-Point Oy
Valkjärrentie 1
SF-02150 Espoo
Finland

Z. R. Flores Bustamante
H. Asama
I. Endo
Chemical Engineering Laboratory
The Institute of Physical and Chemical Research
Wako-shi, 351-01
Saitama
Japan

R. Aarts
P. Linko (corresponding author)
Laboratory of Biotechnology and Food Engineering
Helsinki University of Technology
SF-02150 Espoo
Finland