

**Entwicklung von Expertensystemen zur Prozeß-
überwachung mit der Entwicklungsumgebung G2
unter besonderer Berücksichtigung
der Wissensakquisition und –implementierung**

Dem Fachbereich Chemie
der Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften

- Dr. rer. nat. -

genehmigte Dissertation

von

Dipl.-Chem. Russell Gomersall

geb. am 10. Juni 1968 in Shornecliff (GB)

Hannover 1998

For Dad

Referent: PD Dr. B. Hitzmann

Korreferent: Prof. Dr. T. Scheper

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter von PD Dr. rer. nat. Bernd Hitzmann in der Zeit vom Juli 1994 bis Februar 1998 im Institut für Technische Chemie der Universität Hannover. Diese Arbeit war eingebunden in ein Kooperationsprojekt der Volkswagenstiftung mit dem Titel: „Untersuchungen zur Integration von Expertenwissen in ein System zur Modellierung, Überwachung und Regelung von bioverfahrenstechnischen Prozessen“. Diese Arbeit wurde initiiert und geleitet von Herrn Prof. Dr.-Ing. A. Munack, Herrn Prof. Dr.-Ing. K.-H. Bellgardt und Herrn PD Dr. rer. nat. B. Hitzmann, bei denen ich mich für die Diskussionen bedanken möchte.

Ich bedanke mich bei Herrn Priv.-Doz. Dr. rer. nat. B. Hitzmann für die Leitung, Betreuung, die Diskussionen, seine hilfreichen Anmerkungen und seine Kritik.

Herrn Univ.-Prof. Dr. rer. nat. Thomas Scheper danke ich für die Übernahme des Korreferates. Außerdem gilt mein Dank den Mitarbeitern des Projektes Dipl.-Inf. S. Grumann, Dipl. Ing. D. Ludewig und Dipl.-Ing. M. Seeger.

Bei Herrn PD Dr. rer. nat. R. Guthke möchte ich mich für seine Freundschaft und anregenden Diskussionen während meines Gastaufenthaltes am HKI-Jena bedanken. Des weiteren für die Möglichkeit über die Ergebnisse der Zusammenarbeit zu berichten. Bei der Firma PSI-AG Berlin möchte ich mich für die Einladung zur Mitarbeit am Preussag-Projekt sowie für die Erlaubnis einige Ergebnisse in dieser Arbeit vorstellen zu dürfen bedanken.

Die vorliegende Arbeit wäre nicht möglich gewesen ohne die ausgezeichnete Unterstützung durch die Mitarbeiter des Instituts. In diesem Rahmen möchte ich mich bei Dr. rer. nat. Claus Tollnick, Dipl.-Chem. Guido Seidel und Dipl.-Chem. Uli Kirchner für die Überlassung experimenteller Daten bedanken. Bei den Mitgliedern des „Computerclubs“ Dipl.-Chem. Michael Arndt, Dr. rer. nat. Michael Hoff, Dr. rer. nat. Axel Löhn, Dipl.-Chem. Thomas Pekeler, Dr. rer. nat. Karsten Schöngarth und Dipl.-Chem. Michael Stenns möchte ich mich für die gute Zusammenarbeit bedanken.

Some words of grateful appreciation to my family Helen, Sabine, Jakob and Anatole: I have never met anybody who could compete with your wit, cooking skills (that's for you Mum!) and sense of humour. In times of great despair it was always a great help to know that you would always accept me as I am, no matter what happens.

Last but not least möchte ich meiner lieben Freundin Regine für ihr Verständnis und Hilfe während meiner Promotion danken. Ich weiß, daß es bestimmt nicht leicht war und hoffe, daß ich ihr während ihrer Promotion genauso eine Stütze sein kann.

Abstract

Russell C. Gomersall

Entwicklung von Expertensystemen zur Prozeßüberwachung mit der Entwicklungsumgebung G2 unter besonderen Berücksichtigung der Wissensakquisition und –implementierung

Key words: Bioprozesse, G2, Expertensystem, Prozessüberwachung, On-line Störfalldiagnose, Wissensakquisition, Wissensimplementierung, Clusteranalyse

Für die zuverlässige Überwachung, Störfalldiagnose und Zustandsermittlung von Bioprocessen wird viel Wissen zahlreicher Fachgebiete benötigt. Dieses Wissen wird üblicherweise von erfahrenen Prozeßbediener zur Verfügung gestellt. Um die Überwachung von Bioprocessen mit dem Computer zu automatisieren muß das Wissen zunächst gesammelt und in einem Computersystem implementiert werden. In dieser Arbeit wurde ein Expertensystem entwickelt, mit dem Bioprozesse überwacht werden können. Das Expertensystem ist so konzipiert, daß es flexibel und schnell für die Überwachung unterschiedlicher Prozesse angepaßt werden kann. Um dies zu erreichen, wurde eine leistungsfähige, graphische Benutzerschnittstelle implementiert, die den Bediener in allen Schritten der Wissensakquisition interaktiv unterstützt. Die vorliegende Arbeit bietet Lösungen, um die wichtigsten Schritte der Wissensakquisition erheblich zu vereinfachen. So ermöglicht die Analyse historischer Daten eine induktiven Erhebung prozeßspezifischen Faktenwissens. Darüber hinaus zeigen Untersuchungen mit Ergebnissen der Clusteranalyse, daß sie direkt zur On-line-Klassifizierung, Fehlererkennung und Prozeßvorhersage genutzt werden können. Ein weiterer wichtiger Schritt der Wissensakquisition stellt die Implementierung von Prozeß- und Fehlermodelle durch den Benutzer dar. Hierfür wurde eine vollständig dialoggeführte Regelerstellung entwickelt, welche die einfache Erstellung von Regeln ermöglicht. Beim Einsatz des Expertensystems wird der Prozeßzustand und zahlreiche Störungen sowohl bei einer Kultivierung von *Acremonium chrysogenum* als auch bei einer Hefe-Kultivierung zuverlässig erkannt. Durch eine auf dem Internet basierende Visualisierung wird eine wesentlich höhere Verfügbarkeit der Ergebnisse des Expertensystems erreicht.

Das Problem der fehlenden Übersichtlichkeit großer Regelmengen wird durch die Entwicklung einer speziellen Erklärungskomponente gelöst. Des Weiteren wird ein graphisches Modul zur Wissensimplementierung und –visualisierung entwickelt und diskutiert, welches die Transparenz einer Wissensbasis und somit die Akzeptanz von Expertensystemen erheblich vergrößert. Dieses Modul befindet sich bereits in einer industriellen Anwendung.

Abstract

Russell C. Gomersall

Development of an Expertsystem for Process Supervision with the Expertsystemshell G2 under special consideration of the Akquisition and Implementation of Knowledge

Key words: Bioprocess, G2, expert system, on line process supervision, fault diagnosis, knowledge akquisition, knowledge implementation, cluster analysis

For the reliable supervision, fault diagnosis and state estimation of bioprocesses knowledge of various subject areas is required. This knowledge is normally provided by experienced process operators. To be able to automate the supervision of bioprocesses with computers, the specific knowledge first of all needs to be akquired and implemented in a computer. In this thesis an expert system was developed with which bioprocesses can be supervised, faults detected and the state of the process estimated. The expert system is conceived to allow a flexible and fast adaption to various different processes. To achieve this a powerful graphical user interface was implemented which supports the operator interactively. This contribution provides solutions which considerably simplify the main steps of knowledge akquisition. For example the analysis of historical data allows an induction of process specific knowledge. Furthermore investigations with results of cluster analysis show that they can directly be deployed for on line classification, fault detection and process prediction. Another important step of knowledge akquisition is the implementation of process- and fault models. To support the operator during this task a dialog oriented rule generator was implemented. During the systems deployment for a cultivation of *Acremonium chrysogenum* and a yeast cultivation the process state as well as several faults where reliably estimated. By implementing an internet based visualisation a considerable higher availability of the expert systems conclusions was achieved. The lack of transparency of a large amount of rules was solved by developing a special explanation module. Furthermore a graphical module for knowledge implementation and visualisation was developed. This module greatly enlarges the transparency of knowledge bases and helps to enlarge the acceptance of expert systems. The module is in industrial use.

1	EINLEITUNG.....	1
1.1	ZIELSETZUNG DIESER ARBEIT.....	2
2	GRUNDLAGEN VON EXPERTENSYSTEMEN	4
2.1	AUFBAU VON EXPERTENSYSTEMEN	5
2.2	WISSENSREPRÄSENTATION.....	6
2.3	INFERENZSTRATEGIEN	6
2.4	WISSENSBASIERTER DIAGNOSE	7
2.5	ECHTZEIT-SYSTEME	8
2.6	AUSWAHL EINER EXPERTENSYSTEMSHELL.....	9
2.7	STAND DER TECHNIK.....	13
3	MATERIAL UND METHODEN	19
3.1	KULTIVIERUNG VON <i>ACREMONIUM CHRYSOGENUM</i> IM RÜHRKESSEL	19
3.2	PROZEBLEITSYSTEM.....	20
3.3	NETZWERK	21
3.4	EXPERTENSYSTEM ENTWICKLUNGSUMGEBUNG.....	21
4	EIN EXPERTENSYSTEM ZUR ÜBERWACHUNG VON BIOPROZESSEN	23
4.1	WISSENSAKQUISITION.....	23
4.1.1	<i>Heuristische Prozeßmodelle</i>	<i>27</i>
4.1.1.1	<i>Ablauf einer Kultivierung und Phasenunterteilung.....</i>	<i>28</i>
4.1.1.2	<i>Klassifikation des Prozeßzustandes</i>	<i>29</i>
4.2	DATENANALYSE	30
4.2.1	<i>Charakterisierung der Meßsignale</i>	<i>31</i>
4.2.2	<i>Berechnung von Richtwerten und Intervallen zur qualitativen Beurteilung und Klassifizierung von numerischen Merkmalen</i>	<i>34</i>
4.2.3	<i>Mustererkennung zur Signalunterteilung.....</i>	<i>41</i>
4.3	KONZEPTION DES EXPERTENSYSTEMS	46
4.4	SYSTEMBESCHREIBUNG	47
4.4.1	<i>Architektur des Gesamtsystems</i>	<i>47</i>
4.4.2	<i>Datenschnittstelle.....</i>	<i>48</i>
4.4.3	<i>Externe Datenanalyse</i>	<i>49</i>
4.4.4	<i>Überwachungssystem</i>	<i>50</i>
4.4.4.1	<i>Wissensrepräsentation.....</i>	<i>51</i>
4.4.4.1.1	<i>Objekte.....</i>	<i>51</i>
4.4.4.1.2	<i>Relationen.....</i>	<i>57</i>
4.4.4.1.3	<i>Regeln</i>	<i>58</i>
4.4.4.1.4	<i>Prozeduren.....</i>	<i>61</i>

4.4.4.2	Arbeitsablauf der Überwachung	62
4.4.5	<i>Benutzerschnittstelle</i>	64
4.4.5.1	Protokollfunktion	64
4.4.5.2	Erklärungskomponente.....	66
4.4.5.3	Zusätzliche Visualisierung über das Internet.....	68
4.4.5.4	Wissenserwerbskomponente.....	71
4.4.5.4.1	Prozeßspezifisches Faktenwissen.....	72
4.4.5.4.2	Methodisches Wissen.....	74
4.5	ANWENDUNGSBEISPIELE	77
4.5.1	<i>Überwachung einer Fed-batch-Kultivierung von <i>Acremonium chrysogenum</i></i>	78
4.5.2	<i>Anpassung des Expertensystems</i>	91
4.6	ZUSAMMENFASSUNG.....	102
5	ANWENDUNG DER ERGEBNISSE DER CLUSTERANALYSE ZUR ÜBERWACHUNG VON BIOPROZESSEN	105
5.1	GRUNDLAGEN DER CLUSTERANALYSE	105
5.2	ON-LINE-KLASSIFIZIERUNG.....	106
5.3	MODELLVORHERSAGE	107
5.4	IMPLEMENTIERUNG	109
5.5	ERGEBNISSE.....	109
5.6	ZUSAMMENFASSUNG.....	114
6	EIN EXPERTENSYSTEM ZUR WISSENSVERARBEITUNG IN DER STAHLERZEUGUNG .	116
6.1	PROBLEMSTELLUNG	116
6.2	MATERIAL UND METHODEN.....	117
6.3	PROBLEMLÖSUNG.....	118
6.3.1	<i>Inferenzkomponente</i>	119
6.3.2	<i>Implementierung von neuen Regeln</i>	120
6.3.3	<i>Testen einer neuen Wissensbasis</i>	124
6.3.4	<i>Aktivieren einer neuen Wissensbasis</i>	126
6.4	ZUSAMMENFASSUNG.....	127
7	SCHLUßBETRACHTUNG	128
8	ABKÜRZUNGSVERZEICHNIS.....	131
9	LITERATURVERZEICHNIS.....	132

1 Einleitung

Die Biotechnologie wird heute weltweit als eines der Schlüsselfelder künftiger technologischer Innovationen angesehen. Bei der Nutzung biologischer Verfahren im industriellen Maßstab stehen Produktivität und Betriebssicherheit im Vordergrund. Eine hohe Betriebssicherheit kann durch angepaßte Anlagenkonzepte und durch verbesserte Meß- und automatisierte Überwachungstechniken erreicht werden. Die meisten modernen regelungstechnischen Verfahren zur Prozeßkontrolle basieren auf mathematischen Prozeßmodellen. Die Ermittlung solcher Modelle für biotechnische Prozesse ist jedoch aus mehreren Gründen problematisch:

- Die zellinternen Stoffwechselfvorgänge und deren Regulationsmechanismen sind kompliziert und oftmals nur qualitativ bekannt, weshalb eine sachgerechte Modellbildung vielfach nicht möglich ist.
- Soweit Modellentwicklungen sinnvoll sind, spiegeln sie die hohe Komplexität der Prozesse wieder. Die Modelle sind deshalb sehr aufwendig zu erstellen und durch viele prozeßspezifische Parameter gekennzeichnet.
- Wesentliche, insbesondere intrazelluläre Zustände sind durch fehlende Sensoren nicht meßbar, d.h. die detaillierten Modelle sind – sofern sie existieren – vielfach nicht verifizierbar.
- Reaktionskinetiken und fluiddynamische Gesetzmäßigkeiten führen zu stark nichtlinearem Prozeßverhalten.
- undefinierte Medienzusammensetzung und Abweichungen bei der Vorbehandlung der Organismen führen zu einem nicht exakt reproduzierbaren Verfahrensablauf.

In der industriellen Praxis werden Kultivierungen deshalb heute meist noch von Prozeßbedienern durchgeführt, die aufgrund ihrer Erfahrung in der Lage sind, den Prozeßverlauf zu analysieren, eventuelle Fehler zu erkennen und Störungen zu beheben.

Durch den Einsatz ‚intelligenter‘ Automatisierungssysteme, wie z. B. Expertensysteme, wird versucht die Führung von komplexen, teilweise nur qualitativ bekannten Prozessen zu verbessern. Der Begriff des Expertensystems wurde Mitte der 50‘er Jahre geprägt und rief wie auch andere Teilgebiete der künstlichen Intelligenz hohe Erwartungen in vielen Bereichen hervor. Viele dieser Erwartungen an die künstliche Intelligenz sind eindeutig zu hoch gesteckt und nicht erfüllt worden. Schuld daran dürfte vor allem die falsche Auslegung des Begriffs ‚intelli-

gent‘ sein. Die Vorstellung ‚intelligente‘ Anwendungen zu programmieren, die dem Menschen die Arbeit vollständig abnehmen ist sehr reizvoll. Viele der in diesem Zusammenhang vorgestellten Systeme enthalten aber nur wenige Merkmale, die den Begriff ‚intelligent‘ rechtfertigen. Dennoch haben sich aus dem Bereich der künstlichen Intelligenz verschiedene Anwendungen, dabei insbesondere Expertensysteme, in der biotechnischen Praxis bewährt und sind dort nicht mehr wegzudenken.

Ein Nachteil vieler Expertensysteme, die für den Bereich der Biotechnologie entwickelt wurden, ist die eingeschränkte Verwendbarkeit dieser Systeme. So ist häufig ein Expertensystem speziell für einen Bioprozeß entwickelt worden. Es besteht von daher ein hoher Bedarf an flexiblen Expertensystemen für die Biotechnologie, die sich ohne großen Aufwand an die Gegebenheiten eines Bioprozesses adaptieren lassen. Das Ziel dieser Arbeit war von daher die Entwicklung eines verallgemeinerbaren Expertensystems zur Überwachung von Bioprocessen.

1.1 Zielsetzung dieser Arbeit

Innerhalb des Projekts "*Untersuchungen zur Integration von Expertenwissen in ein System zur Modellierung, Regelung und Überwachung von Bioprocessen*" beschäftigt sich diese Arbeit mit der Konzeptionierung und Entwicklung eines Expertensystems zur Überwachung von Bioprocessen. Die verfolgten Ziele dieser Arbeit sind unter anderem eine Erhöhung des Automatisierungsgrads von biotechnischen Prozessen sowie eine Optimierung der Prozeßführung vor allem in Hinblick auf Betriebsstabilität und -kosten. Ein weiteres Ziel bei der Entwicklung eines solchen Systems ist die Akquisition, Strukturierung und Speicherung des Wissens erfahrener Prozeßbediener. Um die Anwendungsvielfalt des Expertensystems zu steigern, soll das System mit Hilfe der verwendeten Methoden nicht für einen speziellen Prozeß sondern möglichst allgemein formuliert werden, so daß das System für unterschiedliche Prozesse eingesetzt werden kann.

Zur Entwicklung eines Expertensystems werden in der Regel ein Wissensingenieur und mindestens ein Experte des betroffenen Fachgebiets benötigt. Probleme ergeben sich bei dieser Vorgehensweise, wenn z. B. der Fachexperte sein Fachwissen nicht deutlich vermitteln kann. Idealerweise besitzt der Experte daher auch die Fähigkeit des Wissensingenieurs und kann somit sein Wissen selbständig implementieren. Eine Alternative stellt die Entwicklung geeigneter

Werkzeuge zur Unterstützung des Experten bei der Wissensakquisition und Anpassung der Wissensbasis an neue Prozesse dar. Dies soll im Rahmen dieser Arbeit mitentwickelt werden. Zunächst soll ein Überwachungssystem entwickelt und umgesetzt werden, mit dem on line Bioprozesse überwacht werden können. Insbesondere die Ermittlung des Prozeßzustands sowie die Erkennung und Diagnose von Störfällen stehen hierbei im Vordergrund. Zu diesem Zweck müssen im wesentlichen heuristische Prozeß- und Fehlermodelle – die sich der erfahrene Bediener angeeignet hat – in dem Expertensystem implementiert werden. Da das Expertensystem on line – möglichst unabhängig vom Benutzer – den Bioprozeß überwachen soll, ist es auf die On-line-Daten des Prozeßleitsystems als wesentliche Informationsquelle angewiesen. Es muß deshalb ein Konzept umgesetzt werden, welches ermöglicht die On-line-Daten mit den heuristischen Prozeßmodellen zu verarbeiten.

Um dem Anspruch gerecht zu werden das Expertensystem flexibel an neue Prozesse anzupassen. werden im Rahmen dieser Arbeit entsprechende Konzepte entwickelt. Dabei sollen Antworten auf die folgenden Fragen gefunden werden: Welches Wissen kann allgemein formuliert und bei einer Anpassung übernommen werden? Welches Wissen kann automatisch, beispielsweise durch eine Datenanalyse, gewonnen werden? Welches Wissen muß von dem Benutzer beigetragen werden und wie kann man eine optimale Unterstützung gewährleisten?

Die Umsetzung dieses Konzepts soll in einer leistungsfähigen Benutzerschnittstelle resultieren, mit der in Off-line-Sitzungen ein erfahrener Benutzer eigenständig eine Anpassung des Systems vornehmen kann.

Je mehr Daten einer Kultivierung vorliegen, desto mehr bieten sich induktive, datengetriebene Ansätze zur Wissenserhebung an. Die Clusteranalyse stellt solch einen Ansatz zur automatischen Datenanalyse dar. In einer Untersuchung soll analysiert werden, inwieweit die Ergebnisse einer Clusteranalyse – ohne zusätzliches heuristisches Wissen – direkt in einem Expertensystem für die Überwachung und Vorhersage eines Bioprozesses verwendet werden können.

Im Gegensatz zu diesem datengetriebenen Ansatz sollen des Weiteren neue Möglichkeiten der Wissenserhebung und –repräsentation entwickelt werden. Im Vordergrund steht hierbei die Entwicklung einer selbsterklärenden, graphisch-unterstützten Möglichkeit Expertenwissen zu implementieren und darzustellen. Hierfür soll ein Expertensystem entwickelt werden, das in einen industriellen Produktionsprozeß getestet wird.

2 Grundlagen von Expertensystemen

Die Entwicklung von Software erfordert technische Kenntnisse über die verwendete Hardware, das eingesetzte Betriebssystem und die Programmiersprache sowie beispielsweise über Hilfssoftware wie Datenbanken und Datenübertragungssysteme. Außerdem sind Fertigkeiten im Bereich der Analyse und des Design von Programmen und der Fehlererkennung und Fehlerbehebung erforderlich.

Zur Planung und Realisierung von Expertensystemen reichen DV-technische Kenntnisse allein nicht aus. Vielmehr werden ebenfalls Expertenwissen über den zu unterstützenden Fachbereich verarbeitet. Wissensverarbeitung als Teil der automatisierten Datenverarbeitung erlaubt die DV-gestützte Umsetzung dieser – dem menschlichen Problemlösungsverhalten entsprechenden – Techniken.

Bei der konventionellen Datenverarbeitung muß der Systementwickler den Programmablauf festlegen, d.h. die Reihenfolge der Verarbeitungsschritte wird für jede Eingabekonstellation im Voraus betrachtet und anschließend in den Programmcode übertragen. Daraus folgt, daß Daten, Wissen über die Problemlösung und Wissen über die Steuerung der Anwendung dieses Wissens im Algorithmus miteinander vermischt werden.

Bei der Wissensverarbeitung mit Expertensystemen wird diese Vermischung vermieden. Die einzelnen Wissens Elemente werden in bausteinhafter Form formuliert, beispielsweise in einfachen Regeln. Im Idealfall können diese Wissens Elemente in beliebiger Reihenfolge aufgeschrieben werden. Die Abarbeitung der Wissens Elemente führt ein unabhängig entwickelter Interpreter (Inferenzmaschine) selbständig aus. Die Schlußfolgerungskomponente eines wissensbasierten Systems ist meist konventionell programmiert. Der Lösungsweg wird also von einem Inferenzmechanismus, der mit bestimmten Konfliktlösungsmechanismen ausgestattet ist, in eigener Verantwortung gesucht. Dadurch ergeben sich gegenüber der herkömmlichen Programmierweise folgende Vorteile:

- Ein Experte kann ohne spezielle DV-Kenntnisse das System mit Wissen füllen.
- Das System wird überschaubarer und das Wissen kann ständig erweitert werden.
- Es gibt Möglichkeiten 'unscharfes' Wissen zu verarbeiten.

2.1 Aufbau von Expertensystemen

Abbildung 2-1 zeigt die Architektur eines Expertensystems und der wichtigsten Komponenten. Im allgemeinen wird zwischen dem Steuersystem und der Wissensbasis unterschieden [Puppe 1991].

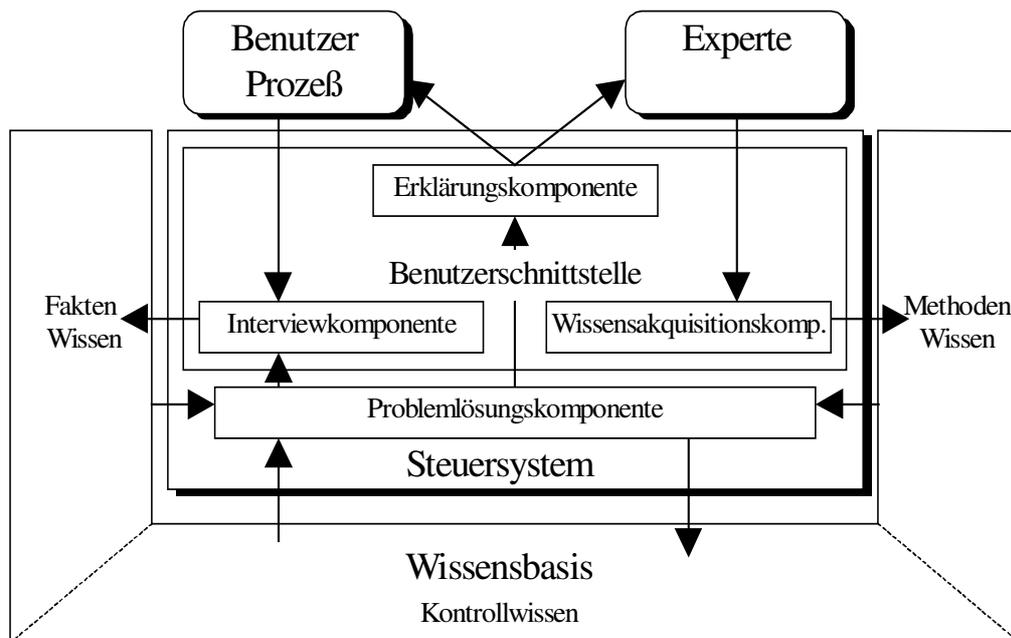


Abbildung 2-1 Aufbau eines Expertensystems

Die Wissensbasis stellt den Kern eines Expertensystems dar. In ihr ist das Wissen und das Problemlösungsverhalten einer oder mehrerer Fachexperten des Anwendungsfeldes abgebildet. Das Wissen besteht im Allgemeinen aus Wissen über konkrete Objekte des abzubildenden Systems und deren Eigenschaften - dem Faktenwissen - und dem Wissen über kausale und zeitliche Zusammenhänge dieser Objekte - dem sogenannten Methodenwissen.

In dem Steuersystem sind sowohl die Problemlösungskomponente (Inferenzmaschine) als auch eine Benutzerschnittstelle enthalten. Die Benutzerschnittstelle besteht aus drei eigenständigen Komponenten, die unterschiedliche Aufgaben erfüllen: Die Wissensakquisitionskomponente ermöglicht dem Entwickler, das Wissen der Wissensbasis in einem interaktiven Dialog zu implementieren, zu erweitern und zu verändern. Die Erklärungskomponente ist die Schnittstelle mit dem Benutzer, welche die Schlußfolgerungen des Systems darstellt und erläutert. Hierbei bieten die meisten Expertensysteme Antworten wie das Expertensystem zu den Schlußfolgerungen gekommen ist und warum das Expertensystem eine bestimmte Information benötigt (bei

einem interaktivem Dialog). Bei interaktiven Systemen ist die Interviewkomponente für die Kommunikation mit dem Benutzer während einer Problemlösung zuständig. So erfragt sie z. B. fehlendes Wissen, daß zur Lösung des Problems benötigt wird.

Die Problemlösungskomponente beinhaltet im wesentlichen die Inferenzmaschine, die von einer Kontrollstrategie geführt nach vorgegebenen Mustern in den Fakten sucht und neues, vorher implizites Wissen aus den Fakten generiert, die explizit in der Wissensbasis abgelegt werden [Sundermeyer 1991].

2.2 Wissensrepräsentation

Die Möglichkeit, Wissen in einer für das Expertensystem verständlichen Form zu repräsentieren, ist eine Grundvoraussetzung für den Bau 'intelligenter' Systeme. Die Wissensrepräsentation ist somit als formale Sprache zu sehen, in der das Expertenwissen ausgedrückt wird. Zu den Basiskonzepten der Wissensrepräsentation gehören Semantische Netze, Frames, Prädikatenlogik und Regeln. Regeln werden in Expertensystemen sehr häufig eingesetzt und bestehen aus einer Vorbedingung (*Prämisse*) und einer Aktion (*Konklusion*). Nach Puppe [1988] wird üblicherweise zwischen zwei Aktionsformen bei Regeln unterschieden: Mit der *Implikation* und *Deduktion* wird der Wahrheitsgehalt einer Prämisse festgelegt und mit der *Handlung* wird ein Zustand manipuliert.

Regelbasierte Systeme – auch Produktionssysteme genannt – zeichnen sich durch Modularität und ihre Adaptionsfähigkeit aus. Erreicht die Wissensbasis eine bestimmten Größe, werden reine Produktionssysteme ineffizient und es mangelt ihnen häufig an Übersichtlichkeit, Konsistenz und Transparenz [Sundermeyer 1992].

2.3 Inferenzstrategien

In Expertensystemen werden üblicherweise zwei Inferenzstrategien verfolgt. Eine zielgesteuerte Problemlösungsstrategie stellt das Rückwärtsschließen – auch Backward-Chaining genannt – dar. In einem Produktionssystem wird die Rückwärtsschließen dadurch erreicht, daß ein Ziel in Form einer Hypothese vorgegeben wird und die Regeln der Wissensbasis rückwärts durch-

gearbeitet werden, bis ein Satz zufriedenstellender Prämissen gefunden wird, unter denen die Hypothese zutrifft.

Das Vorwärtsschließen – Forward-Chaining genannt – ist ereignisgesteuert. Sie bezieht sich auf das Schlußfolgern aus erfüllten Prämissen und dem Ausführen der Konklusionen. Dadurch wird die Datenbasis verändert, so daß weitere Schlußfolgerungen gezogen werden können. Dieser Prozeß wiederholt sich bis keine Prämisse mehr erfüllt ist. Ein Nachteil des Vorwärtsschließens besteht darin, daß vielfach Daten betrachtet werden, die irrelevant für die betrachtete Sachlage sind.

Eine speziell für diagnostische Anwendungen konzipierte Kontrollstrategie ist die *Hypothesize & Test* Strategie, bei der aufgrund von Symptomen durch Vorwärtsverkettung Verdachtshypothesen generiert werden. Diese werden in der Testphase gezielt durch Rückwärtsverkettung untersucht [Sticher 1990b, Gehlen 1993]. Die Qualität dieser Kontrollstrategie wird durch die Effizienz und Korrektheit der Verdachtsgenerierung bestimmt.

2.4 Wissensbasierte Diagnose

Nach Isermann [1992] unterteilt sich die wissensbasierte Überwachung in die eigentlichen Aufgaben der Fehlererkennung durch Erzeugung analytischer und heuristischer Symptome und der Fehlerdiagnose mit Klassifikations- und Inferenzmethoden.

Bei der analytischen Symptomerzeugung wird quantifizierbares Wissen über den Prozeß verwendet, um analytische Information zu erzeugen. Hierzu wird mittels gemessener Variablen eine Datenverarbeitung durchgeführt, um Merkmale bzw. Kenngrößen zu bilden durch:

- Grenzwert-Überwachung (Merkmale sind überschrittene Signaltoleranzen),
- Signalanalyse (Merkmale sind z. B. Varianzen, Amplituden und Trends) und
- Prozeßanalyse (Merkmale sind Parameter, Zustandsgrößen oder Residuen).

Diese Merkmale werden mit den normalen Merkmalen des fehlerlosen Prozesses verglichen. Wobei Methoden der Erkennung signifikanter Änderungen eingesetzt werden. Die resultierenden Änderungen der Merkmale bilden dann die analytischen Symptome.

Als Quelle heuristischer Symptome gehören neben qualitativen Informationen über die der Bediener verfügt auch die bisherige Prozeßgeschichte in Form von Wartung, Reparatur, früheren Fehlern, Standzeiten ebenso wie statistische Daten über denselben oder ähnliche Prozesse.

Heuristische Symptome werden als linguistische Variablen (z. B. niedrig, normal, hoch) oder als unscharfe Zahlenwerte angegeben.

Die Aufgabe der Fehlerdiagnose besteht aus der Bestimmung des Typs, der Größe und dem Ort des Fehlers sowie dem Zeitpunkt seiner Erkennung auf der Grundlage der analytischen und heuristischen Symptome.

2.5 Echtzeit-Systeme

Ein Großteil der entwickelten Expertensysteme arbeitet interaktiv im Dialog mit dem Benutzer, um eine Aufgabe zu lösen. Da keine zeitliche Beschränkung in der Antwortzeit besteht und der Benutzer immer in den Vorgang der Diagnose mit einbezogen ist, sind solche Systeme zur Automatisierung von Diagnoseaufgaben wenig geeignet. Das Ergebnis einer Diagnose muß vorliegen, bevor sich der Zustand des Prozesses soweit verändert hat, daß das Ergebnis veraltet ist. Die direkte Ankopplung des Systems an den Prozeß erfordert ein zeitgerechtes Antwortverhalten. Hierzu wird die Dialogkomponente interaktiver Systeme bei Realzeit-Expertensystemen um ein Datenaufnahmesystem erweitert, welches Meßdaten zuführt. Folgende Anforderungen werden an ein Realzeit-Expertensystem gestellt [Griesmer et al. 1984, Laffey et al. 1988]

- Sicherstellung eines kontinuierlichen Betriebs,
- eine effiziente Integration von numerischer und symbolischer Datenverarbeitung,
- die Möglichkeit einer Fokussierung der Arbeit,
- die Möglichkeit einer zeitabhängigen Schlußfolgerung und
- ein vorhersagbares zeitliches Verhalten.

Wichtig ist die Sicherstellung des kontinuierlichen Betriebs und die damit verbundene Speicher- und Datenorganisation, da ein Realzeit-Expertensystem zum Beispiel für mehrere Tage im Einsatz sein muß.

Die Anwendungsgebiete von Realzeit-Expertensystemen finden sich hauptsächlich in der Prozeßkontrolle. Im Gegensatz zu interaktiven Systemen sind Realzeitsysteme durch eine hohe Ausführungsgeschwindigkeit, eine relativ kleine Wissensbasis, einen hohen Anteil an numerischer Datenverarbeitung und einer wenig funktionalen Erklärungskomponente gekennzeichnet [Linkens und Chen 1995].

2.6 Auswahl einer Expertensystemshell

Zur Erstellung eines Expertensystems bedient man sich heute in der Regel einer Expertensystemshell. Der Name beruht auf der Tatsache, daß diese Werkzeuge häufig schon entwickelte Expertensysteme darstellen, von denen die Wissensbasis gelöscht wurde. Diese sogenannte ‚Schale‘ (Shell) beinhaltet alle für ein Expertensystem wichtigen Komponenten. Davis [1993] beschreibt wie Expertensystemshells zu der Entwicklung von Überwachungssystemen beitragen können. Die Auswahl einer angebrachten Entwicklungsumgebung trägt wesentlich zum Erfolg einer Neuentwicklung bei und stellt eine große Herausforderung dar. Kuesten und McLellan [1994] sowie Harmon [1993] greifen diese Problematik auf und geben Richtlinien und weiterführende Literatur für die Auswahl einer geeigneten Expertensystemshell an. Bei der Auswahl sollten die folgende Punkte besondere Beachtung finden:

Datenschnittstelle

Ein noch so mächtiges Werkzeug erfordert trotzdem eine Schnittstelle zu anderen Programmen, die mit herkömmlichen Programmen wie C++, C oder FORTRAN geschrieben sind. Bei datenintensiven Anwendungen ist eine Schnittstelle zu herkömmlichen Datenbanken unerlässlich.

Entwicklungsumgebung

Bei der Beurteilung der Entwicklungsumgebung müssen zahlreiche Fragen beantwortet werden. Bietet die Umgebung die Möglichkeit prozedurale Funktionen zu implementieren oder anzubinden? Werden objekt-orientierte Programmiermöglichkeiten unterstützt? Wird die Wissensbasis kompiliert oder interpretiert? Erfordern Änderungen eine Neucompilierung oder wird die Wissensbasis inkrementel kompiliert? Inwieweit existieren Debugger für die Wissensbasis? Wie groß sind die graphischen Fähigkeiten des Werkzeugs? Kann das Expertensystem auf unterschiedlichen Plattformen eingesetzt werden? Neben diesen Fragen, welche die direkten Fähigkeiten des Werkzeugs beurteilen, stellen sich Fragen bezüglich der Unterstützung durch Dokumentation und Service des Anbieters.

Inferenzmaschine

Zu den wichtigsten Inferenzstrategien, die von der Inferenzmaschine zur Navigation der Wissensbasis und Schlußfolgerung Verwendung finden, sind unter anderem die Vorwärts-, Rückwärts- und Bidirektionale-Verkettung sowie die Breiten-, Tiefen- und die Erschöpfende-Suche zu nennen. Zusätzlich bieten Inferenzmaschinen unterschiedliche Konfliktlösungsmechanismen an, die aus einer Regelmenge, deren Bedingungssteile erfüllt sind, eine Regel aussucht und deren Aktionsteil ausführt. Zu dem Bekanntesten gehört der RETE-Algorithmus. Häufig kommen einfachere Mechanismen zum Einsatz wie z. B. die Konfliktlösung nach der Priorität, Komplexität, Spezifität der Regeln oder einfach nach dem Zufallsprinzip. Eine weitere wichtige Eigenschaft einer Inferenzmaschine ist die Fähigkeit mit unscharfen Werten umgehen zu können, wie sie in Fuzzy-Variablen häufig zum Einsatz kommen.

Wissensrepräsentation

Es gibt unterschiedliche Möglichkeiten zur Wissensrepräsentation in Expertensystemen. Zu den Wichtigsten gehören Regeln, Entscheidungstabellen, Frames, Scripts, semantische Netze, Induktionen und Blackboards. Regeln stellen einen modularen, uniformen Ansatz dar; Entscheidungstabellen hingegen unterstützen eine strukturierte Formulierung des Wissens, die besonders von Vorteil ist, wenn eine gegebene Auswahl an Lösungsalternativen vorliegt. Frames stellen einen objekt-orientierten Ansatz dar, bei dem einem Objekt unterschiedliche Eigenschaften zugesprochen werden. Zu diesen Eigenschaften gehören z. B. Fakten, Regeln, Methoden, Startwerte und Aktivwerte. Scripts ähneln den Frames, geben aber das Wissen über Sequenzen von Ereignissen wieder, während Frames eher das Wissen von statischen Objekten symbolisieren. Expertensystemshells bieten häufig mehrere Formen der Wissensrepräsentation an. Im Vorfeld ist es unerlässlich, das Wissen der Domain zu strukturieren und festzulegen, welche Repräsentationsart am Besten für das Problem geeignet ist. Neben der Art der Wissensrepräsentation stellt sich hierbei auch die Frage nach der Kapazität der Wissensbasis, wobei diese häufig nur durch den Speicher des Systems limitiert wird. Bei Echtzeit-Systemen ist die Ausführungsgeschwindigkeit von großer Bedeutung, da sich bei großen Produktionssystemen schnell zu lange Antwortzeiten ergeben.

Implementierung

Die heutigen Expertensystemshells sind in unterschiedlichen Programmiersprachen verwirklicht und werden für zahlreiche Hardware und Betriebssysteme angeboten. Viele frühere Entwicklungen in LISP sind mittlerweile in C portiert worden. Die Vorteile solcher Systeme sind

- eine höhere Ausführungsgeschwindigkeit,
- eine größere Portierbarkeit (die Fähigkeit auf unterschiedlichen Hardware/ Betriebssystemen zu laufen),
- die Einbettung in bestehende Anwendungen und
- die Kompatibilität (Kommunikation mit anderen Anwendungen, Datenbanken und Netzwerken).

Benutzerschnittstelle

Die Benutzerschnittstelle dient der Kommunikation mit dem Anwender. Bei der Entwicklung dieser Schnittstelle sollte diese den Anforderungen und Bedürfnissen des Endnutzers angepaßt werden. Einige Shells bieten besondere graphische Werkzeuge, mit denen Benutzerschnittstellen entwickelt werden können. Andere Shells hingegen erfordern, daß die Benutzerschnittstelle mit herkömmlichen Werkzeugen erstellt wird, z. B. mit VisualBasic oder VisualC.

Kantrowitz [1997] bietet eine ständig aktualisierte Übersicht verfügbarer Expertensystemshells im Internet an. Es werden ca. 80 verschiedene Entwicklungsumgebungen mit ihren Fähigkeiten beschrieben. In der Tabelle 2-1 ist eine Auswahl verschiedener Expertensystemshells und ihre wesentlichen Eigenschaften aufgelistet. Es muß hervorgehoben werden, daß die Auswahl sehr stark von der geplanten Anwendung abhängt und daß es kein optimales Werkzeug gibt. Daß die Entscheidung für eine Shell auch durch finanzielle Aspekte beeinflusst wird, machen die Lizenzkosten der Werkzeuge deutlich, die zwischen 0,- DM und 100.000,- DM liegen.

Tabelle 2-1 Eigenschaften einiger Expertensystementwicklungsumgebungen

Name	Wissensverarbeitung	Bemerkung
FOCL	Horn Klauseln, nur Rückwärtsverkettung	
OPS5/OPS83	Regeln, Objekte, RETE-Alg.	VMS, UNIX, DOS
BABYLON	Frames, Constraints	
MIKE	Regeln, Frames, RETE-Alg.	DOS
ES	Fuzzy-Regeln	
WindExS	Regeln	Windows
RT-Expert	Regeln	DOS, Windows und UNIX; Integration von Regeln in C, C++.
CLIPS 6.0	Regeln, Fuzzy-Regeln, Blackboards, Objekte	Integration von Regeln in C,C++; Zahlreiche Plattformen und Erweiterungen des Systems durch DYNA-, Agent-, Fuzzy-und WxClips.
JESS	Siehe CLIPS	Java-Portierung von CLIPS
ACQUIRE	Regeln, Entscheidungstabellen	Ausgereifte kommerzielle Entwicklungsumgebung
CAM	Entscheidungsbäume	
CBR Express	Fallbasierte Entscheidungen	
COMDALE /C/X		Echtzeitfähigkeiten, speziell für die Prozeßüberwachung
C-PRS	Prozedurale Schlußfolgerungen in C	UNIX, Echtzeitfähigkeiten, speziell in der Prozeß Regelung und Überwachung
Easy Reasoner	Fallbasierte Entscheidungen und Diagnosebäume	
Esteem	Fallbasierte Entscheidungen und Regeln	
EXSYS	Fuzzy-Logik, neuronale Netze	SQL-Schnittstelle
FLEX	Frames, Prozeduren und Regeln	Zahlreiche Plattformen und kommerzielle Anwendungen.
G2	Regeln, Objekte, Prozeduren, Relationen, Connections, Vorw.- und Rückwärtsverkettung	Schnittstellen zu Datenbanken und externen Programme, graphische Client-Server Entwicklungsumgebung, weltweiter Vertrieb und Service, zahlreiche kommerzielle Anwendungen
GBB	Generische Blackboard	
HUGIN	Modellbasiertes Wissen	
KDS	Fallbasiertes Wissen	
NEXPERT OBJECT	Regeln, Formulare	Für 30 Plattformen erhältlich, graphische Entwicklungsumgebung
RAL	Regeln, Objekte	Integration in C
ReMind	Fallbasiert	
Rete++	Regeln, Objekte	Integration in C, C++, graphische Entwicklungsumgebung
RT-works	Regeln, Objekte und Prozeduren	Zahlreiche Module für Echtzeitanwendungen, graphische Fähigkeiten
VBXpert	Regeln, Objekte, RETE-Alg.	Integration in Visual Basic
YAPS	Regeln	Modularisierung der Wissensbasis möglich

Die in dieser Arbeit vorgestellten Expertensysteme wurden alle mit der Expertensystemshell G2 von GENSYM erstellt. G2 stellt bezogen auf ihre Funktionalität von allen zur Zeit gewerblich erhältlichen Shells die Leistungsfähigste dar. Durch die häufige Verwendung von G2 bei der Entwicklung industrieller Expertensysteme wird deutlich, daß mit dieser Shell ein Werkzeug zur Entwicklung von Expertensystemen vorhanden ist, die sich im industriellen Einsatz bewähren können.

2.7 Stand der Technik

Methoden der Datenanalyse

Viele verschiedene Expertensysteme sind für die Optimierung und Regelung von biologischen Prozessen entwickelt worden. Expertensysteme sind aber nur so leistungsfähig, wie das Wissen ihrer Wissensbasis [Aynsley et al. 1993]. In der Biotechnologie basiert ein wesentlicher Anteil des Wissens, das für die Führung biotechnischer Prozesse benötigt wird auf den Heuristiken erfahrener Bediener. Die Akquisition dieses Wissens und die Implementierung in einer für den Computer verständlichen Form stellt die größte Schwierigkeit bei der Entwicklung von Expertensystemen dar. Die häufigsten Ansätze, um diese Problematik zu überwinden, basieren auf der Datenanalyse. Sonnleitner und Locher [1992] betonen den Nutzen von Datenanalyse und Mustererkennung zur Identifizierung von physiologischen Zuständen in biotechnischen Daten. Isermann [1996] beschreibt den Einsatz konventioneller, mathematischer Modelle zur Prozeßüberwachung. Zahlreiche Veröffentlichungen berichten von der Anwendung neuronaler Netze auf Kultivierungsdaten [Montague und Morris 1994, Simutis et al. 1994]. Golmer und Posten [1995] verwenden das ‚time warping‘ und die Formant Analyse für die Phasenerkennung in Bioprozessen. Locher et al. [1995] setzen sowohl Methoden der Mustererkennung als auch Entscheidungsbäume und Wavelet-Filter ein, um Kultivierungsdaten zu verwerten. In einem anderen Ansatz von Saner und Stephanopulous [1992] wurde die Hauptkomponentenanalyse zur Mustererkennung in Kultivierungsdaten eingesetzt. Die Ergebnisse dieser Analyse wurden für die On-line-Überwachung eines Bioprozesses verwendet. King [1992] verwendet multiple, nicht lineare Filter, um Störfälle in Bioprozessen zu entdecken. Guthke und Roßmann [1991] gehen das Problem der automatischen Erstellung von Wissen in zwei Schritten an. Im ersten Schritt werden die Daten anhand einer hierarchischen Klassifikation analysiert, und die einzel-

nen Meßdaten in Klassen großer Ähnlichkeit unterteilt. Eine Klasse ist durch den zeitlichen Verlauf des Mittelwerts und der Standardabweichung definiert. In dem zweiten Schritt werden Regeln aus Zugehörigkeitstabellen unter Verwendung des Fischer Tests generiert. Das so gewonnene Wissen wird vornehmlich für die Optimierung von Bioprozessen eingesetzt [Guthke und Ludewig 1994].

Expertensysteme in der Prozeßtechnik

Trotz der großen Anzahl entwickelter Expertensystemprototypen und der überwiegend positiven Erfahrungen mit deren Nutzen, wird der Expertensystemtechnologie für den praktischen Einsatz eher mit Skepsis begegnet. Mögliche Gründe hierfür sind, daß Aufwand und Nutzen schwierig abzuschätzen und nur ein geringer Anteil der Projekte über das Stadium des Prototypen hinaus gekommen sind. Eine allgemeine Übersicht der Vorgehensweise bei der Entwicklung von Expertensystemen wird von Buchanan [1983] gegeben. Zöllner und Mänzel [1996] untersuchen Möglichkeiten, frühzeitige Entscheidungen über die Wirtschaftlichkeit von Expertensystemen finden zu können. Treffen bei einem geplanten Projekt bestimmte K.-o.-Kriterien zu, so bietet sich der Einsatz von Expertensystemen nicht an. Ein K.-o.-Kriterium liegt dann vor, wenn

- kurze, zeitkritische Reaktionen des Systems gefordert werden,
- umfangreiche arithmetische Berechnungen notwendig sind,
- kein Experte des geplanten Einsatzgebiets verfügbar ist,
- die Verantwortung für die Wissensakquisition ungewiß ist,
- Zeitdruck besteht,
- kritische Pfade des Wissensgebietes besritten werden,
- das Wissensgebiet nicht verbalisierbar ist,
- die Problemstellung undynamisch ist.

Trotz dieser vielen K.-o.-Kriterien gibt es eine Vielzahl von Einsatzgebieten für Expertensysteme in der Prozeßtechnik. Viele Expertensysteme wurden zunächst interaktiv ausgelegt. So sind z. B. zahlreiche Systeme beschrieben, die für die Fehlerdiagnose von komplexen Analyseräten eingesetzt werden. Tsuji und Jenkins [1989] stellen ein wissensbasiertes interaktives System zur Fehlerdiagnose von HPLC-Anlagen dar. Durch die interaktive Arbeitsweise werden viele Möglichkeiten – die Computer bieten – nicht ausgenutzt, weshalb sie mehr und mehr direkt an Prozesse angebunden werden. Über Probleme beim Echtzeit-Einsatz von Expertensy-

stemmen berichten Jones und Rodd [1993]. Durch die On-line-Überwachung wird der Benutzer zusätzlich entlastet. So stellen Hitzmann und Brandt [1994] sowie Hitzmann und Lohn [1997] regelbasierte Expertensysteme zur On-line-Störfalldiagnose bzw. Kalibrationsvalidierung von FIA-Systemen vor. Hoff [1997] verwendet ein ähnliches Konzept, um on line eine HPLC-Anlage zu überwachen.

Poletykin [1994] stellt ein Expertensystem vor, das anhand von Fehlersimulationen Regeln ableitet, mit denen der Fehler erkannt und der Prozeß wieder in den Normalzustand geführt werden kann. Das System findet bei einer Gas-Kraftwerk Anwendung. Ein regelbasiertes Expertensystem zur Diagnose und Regelung der Energieverteilung in einem Stahlwerk wird von Lee et al. [1992] vorgestellt. Das System wurde mit der Expertensystemshell Nexpert-Objekt entwickelt. Ziel des Systems ist es eine optimale Nutzung der Heizgase zu erreichen, die in bestimmten Anlagen als Nebenprodukt anfallen.

Boelow et al. [1996] kombinieren kognitive Methoden zur Ermittlung von zeitvarianten Prozessen mit Fuzzy-Regel, um fallbasierte Automaten zu entwickeln. Die Anwendungsbeispiele dieser Arbeiten reichen von der Steuerung einer Brennkammer bis hin zur Auge-Hand-Koordination eines Bedieners. Kaldorf [1992] beschreibt mögliche Einsatzbereiche für Echtzeit-Expertensysteme in der chemischen Industrie. Als Beispiel wird die Realisierung eines Expertensystems beschrieben, welches an das Prozeßleitsystem gekoppelt ist und ca. 200 Fehler in der Anlage diagnostizieren kann. Über Einsatzgebiete für Expertensystemtechnologie in der Lebensmittelbranche gibt Davidson [1994] Aufschluß.

Die Anwendung von neuronalen Netzen ist in vielen Bereichen der Überwachung verbreitet. Rojas-Guzmann und Kramer [1993] sowie Fogarty [1994] vergleichen die unterschiedlichen Eigenschaften und Möglichkeiten von neuronalen Netzen mit regelbasierten Systemen bei der Überwachung von chemischen Prozessen. Peters [1996] beschreibt unterschiedliche Einsatzmöglichkeiten, wobei ein hybrider Ansatz mit neuronalen Netzen und regelbasiertem Expertensystem verwendet wird. Die Einsatzgebiete sind u. a. die Überwachung von Hochspannungsnetzwerken, Hochöfen und Kernkraftwerke.

Eine Beschreibung unterschiedlicher Expertensysteme, die mit G2 realisiert wurden, gibt Mason [1993]. So stellt ‚Linkman‘ ein System zur Qualitätssicherung und zum Energiemanagement in der Zementfabrikation dar, das von ABB vertrieben wird. Die Firma Blue Circle gibt an mit dem System, welches auf sechs ihrer Anlagen eingesetzt wird, 10 % ihrer Energiekosten (12 Mill. DM) zu sparen. Die Firma ICI setzt G2 als Beratungssystem zur Fehlerdia-

gnose und Unterstützung der Fehlerbehebung bei einer Beschichtungsanlage ein. Die durch das System geschätzten Einsparungen belaufen sich auf ca. 350.000,-DM/Jahr (1993) wodurch sich die Investitionen innerhalb eines halben Jahres amortisiert haben. Die Firma Norsk Hydro [Saelid 1992] setzt hingegen G2 beim Umweltmanagement einer Düngemittelfabrik ein.

Ein Expertensystem zur Überwachung und Regelung eines kontinuierlichen α -Butanol Reaktors in der Petrochemischen Industrie wird von Cauvin et al. [1993] vorgestellt. Primäres Ziel ist es anhand von implementierten Wirkungszusammenhängen einzelner Prozeßgrößen den Reaktor stetig in einem stationären Zustand zu halten.

Expertensysteme in der Biotechnik

Ein Produktionssystem zur On-line-Überwachung einer *E. coli*-Kultivierung wird von Hitzmann [1989] beschrieben. Dieses SUPERVISOR genannte Expertensystem wurde auf der Basis von OPS5 entwickelt und ist für die Überwachung, Regelung und Steuerung prozeßrechnergekoppelter Bioreaktoren konzipiert. Von Hitzmann [1992] als auch von Clapp [1991] wurde eine Übersicht über die Entwicklung von Expertensystemen in der Biotechnik gegeben. Gehlen [1993] stellt ein wissensbasiertes System vor, daß eine lernende Regelung beinhaltet, die zur Optimierung der Ausbeute eingesetzt wird. In einem eigenständigen, regelbasierten Modul wird der Prozeß auf Abweichungen vom Normalverlauf überwacht. Im Falle einer Abweichung wird die Lernstrategie ausgeschaltet. Das System wurde bei der Fed-batch-Kultivierung von *B. subtilis* zur Produktion von α -Amylase eingesetzt.

Chynoweth et al. [1994] stellen ein Echtzeit-Expertensystem zur Regelung einer kontinuierlichen anaeroben Kultivierung vor. Über die on line gemessene Methanausbeute stellt das Expertensystem fest, ob der Fermenter über- oder unterladen ist oder ob eine Inhibierung vorliegt. Roca et al. [1996] setzen ein Expertensystem zur Regelung und Überwachung einer kontinuierlichen Fermentation von *S. cerevisiae* zur Produktion von Ethanol ein. Das Expertensystem ist hybrid aufgebaut, wobei das numerische Modul in C und das wissensbasierte Modul in OPS83 programmiert wurde. Zur Überwachung des Prozeßzustands wurde ausschließlich der CO_2 Gehalt des Abgases betrachtet. Ebenfalls mit dem Produktionssystem OPS83 wurde ein Expertensystem entwickelt, das von Halme und Karim [1989] beschrieben wird. Das System wurde bei der Batch-Kultivierung von *B. subtilis* zur Produktion von α -Amylase eingesetzt. Das System erhält von dem Prozeßleitsystem BIOSTAT On-line-Daten, die für die Ermittlung des Prozeßzustands sowie der Störfallerkennung herangezogen werden. Die Wissensbasis des

Systems ist dynamisch ausgelegt, wobei die Regeln mit jeder zusätzlichen ‚Erfahrung‘ angepaßt werden. Das Laufzeitverhalten des Systems wird dadurch verbessert, daß nur Regeln der einzelnen Phasen aktiviert werden. Nakajima et al. [1994] sowie Numers et al. [1994] beschreiben ein wissensbasiertes System zur Fuzzy-basierten überwachten Regelung einer *lactobacillus* Kultivierung zur Produktion von Lactat. In diesem System, welches auf einem MAC II mit Hilfe der Programmiersprache Smalltalk/V realisiert wurde, werden die on line gemessenen Daten Temperatur, pH, Agitation, Substratzugabe, Produktkonzentration sowie die Bio-trockenmasse von dem Prozeßleitsystem eingelesen. Dem speziellen zeitlichen Verlauf der Kultivierung wird mit zeitlich variablen Zugehörigkeitsfunktionen Rechnung getragen, die aus Kultivierungen berechnet werden, die unter Standardbedingungen verliefen. Einen Redesign dieses Systems unter Verwendung von G2 beschreiben Nakajima et al. [1997]. Dabei wird ein Schwerpunkt der Arbeiten auf die Entwicklung einer leistungsfähigen Mensch-Maschine-Schnittstelle gelegt.

Aarts et al. [1990] beschrieben ebenfalls ein mit Smalltalk realisiertes Expertensystem, welches zur Störfalldiagnose eines Bioprozesses eingesetzt wird. In dem System finden Fuzzy Mengen, Regeln und Objekte bei einer Batch-Kultivierung von *Aspergillus niger* zur Produktion von Glucoamylase Anwendung.

Steyer et al. [1996] unterstreichen den Nutzen von qualitativen Regeln zur Beurteilung Biotechnischer Prozesse. In dem von Ihnen beschriebenen Realzeit-Expertensystem BIOTECH werden industrielle Kultiverungen von *Saccharomyces cerevisiae* und *Escherichia coli* unter Verwendung qualitativer Prozeßmodelle überwacht. Zu den Aufgaben des Expertensystems zählen unter anderem die Beurteilung des Prozeßzustands, die Diagnose von Sensorstörungen und die Ermittlung der optimalen Fermentationsdauer.

Giron-Sierra et al. [1994] verwenden die Parameter eines Differentialgleichungsmodell in Kombination mit einem objekt-orientierten, wissensbasierten Expertensystem zur Überwachung einer Hefekultivierung. Die Berechnung der Modellparameter erfolgt mittels eines in C realisierten numerischen Moduls und das Expertensystem wurde unter Verwendung der Expertensystemshell KAPPA verwirklicht.

Fraleigh und Lauer [1994] verwenden eine Kombination von Fuzzy-Logik und neuronalen Netzen bei der Qualitätskontrolle einer Chromatographiekolonie zur Proteinseparation. Das System ist in der Lage anhand von Druckprofilen, Störungsursachen zu diagnostizieren. Dabei überwacht ein autoassoziatives Netz die nicht linearen Korrelationen von Druck, Durchfluß

und Temperatur. Die Ausgaben des neuronalen Netzes werden mit Fuzzy-Regeln verarbeitet. Mit der gleichen Methoden-Kombination haben Ignova et al. [1996] ein Expertensystem zur intelligenten Prozeßüberwachung einer industriellen Penizillin Produktion mit G2 entwickelt. Dieses System untersucht, die On-line-Meßgrößen zuerst mit Regeln auf Plausibilität und Störfälle. Ein nachgeschaltetes Modul verarbeitet die validierten, normierten Daten mit neuronalen Netzen. Das Expertensystem generiert Vorhersagen über die Produktivität anhand der Kohlendioxidproduktionsrate und der Sauerstoffaufnahme. Darüber hinaus wird eine zukünftige Vorhersage der Viskosität anhand der On-line-Größen ermittelt. Aldridge [1994] beschreibt den Einsatz von G2 und dem Genzym Diagnostic Assistant (GDA) zur Entwicklung eines Systems zur Überwachung von Durchflußzytometern, die bei Tierzellkultivierungen Einsatz finden. Das System wurde für den industriellen Einsatz bei Novo Nordisk entwickelt. Eine Beschreibung der graphischen Möglichkeiten des G2-Moduls GDA zur schnellen Entwicklung von Überwachungsmodulen geben Larsson [1992] und Fraleigh et al. [1992]. Haferkamp [1995] berichtet von dem schrittweisen Ersatzes eines Prozeßleitsystems durch ein G2-basiertes System in einem industriellen Fermentationsbetrieb der Firma NSC Technologies. Durch die stufenweise Einführung von eigenständigen Wissensbasen, die spezielle Aufgaben übernehmen, konnte so nach und nach das alte Prozeßleitsystem, ohne Unterbrechung des Betriebs, ersetzt werden. Von dem Einsatz von G2 zur Entwicklung von intelligenten Anwendungen zur On-line-Analyse von industriellen Fermentationsdaten der Firma Eli Lilly berichtet Alford et al. [1992]. Zu den implementierten Anwendungen gehören unter anderem die regelbasierte Vorhersage des Animpfzeitpunkts anhand der Kohlendioxidproduktionsrate. Hierbei wurde eine erhöhte Zuverlässigkeit des Systems gegenüber der herkömmlichen DCS basierten Lösung erreicht.

3 Material und Methoden

3.1 Kultivierung von *Acremonium chrysogenum* im Rührkessel

Bioreaktor und Peripherie

Es wird eine Pilotfermentieranlage Biostat UD30 (Firma BBI, Melsungen) mit 30 l Arbeitsvolumen eingesetzt. Das Kulturgefäß besteht aus poliertem Edelstahl und ist mit drei 6-Blatt-Scheibenrührern ausgestattet. Die Rührerwelle wird durch den Reaktorboden geführt und ist mit einer mediengeschmierten Gleitringdichtung versehen. Das Temperiersystem basiert auf einem Doppelmantel und einem externen Rohrbündelwärmetauscher, der mit Heißdampf und Kühlsole oder Betriebswasser gespeist wird und ermöglicht ein Aufheizen von Raumtemperatur auf Sterilisationstemperatur (121°C) innerhalb von 30 Minuten. Die Begasung erfolgt über ein Gas-Ratio-Control-System mit zwei thermischen Massflow-Controllern für Sauerstoff und Preßluft. Dies gestattet eine Regelung des Sauerstoffgehalts der Zuluft und damit des pO_2 bei konstanter Begasungsrate.

Die Regelung von pH, pO_2 , Temperatur und Drehzahl erfolgt über die integrierte digitale Kontrolleinheit (DCU). Abbildung 3-1 zeigt den schematischen Aufbau der Anlage und Tabelle 3-1 zeigt die Sonden/Armaturen, mit denen das Kulturgefäß ausgestattet ist:

Tabelle 3-1 Ausstattung des Kulturgefäßes

Größe	Gerätebezeichnung	Meßbereich	Hersteller
Manteltemperatur	PT100	0-150°C	Braun, Melsungen
Reaktorinnentemperatur	PT100	0-150°C	Braun, Melsungen
Kopfdruck	Manometer, mechanisch	0-3bar	Braun, Melsungen
PH-Wert	Gelelektrode	2-12	Broadley-James
Gelöstsauerstoff	pO_2 -Sonde	0-150 %	Mettler-Toledo
Kulturfluoreszenz (1-D)	Fluorosensor	0-10V	Ingold
Kulturfluoreszenz (2-D)	Fluoreszenzspektrophotometer	-	Hitachi F4500
CO ₂ im Abgas	UNOR 6N	0-100 %	Maihak, Hamburg
O ₂ im Abgas	OXOR 610	0-100 %	Maihak, Hamburg

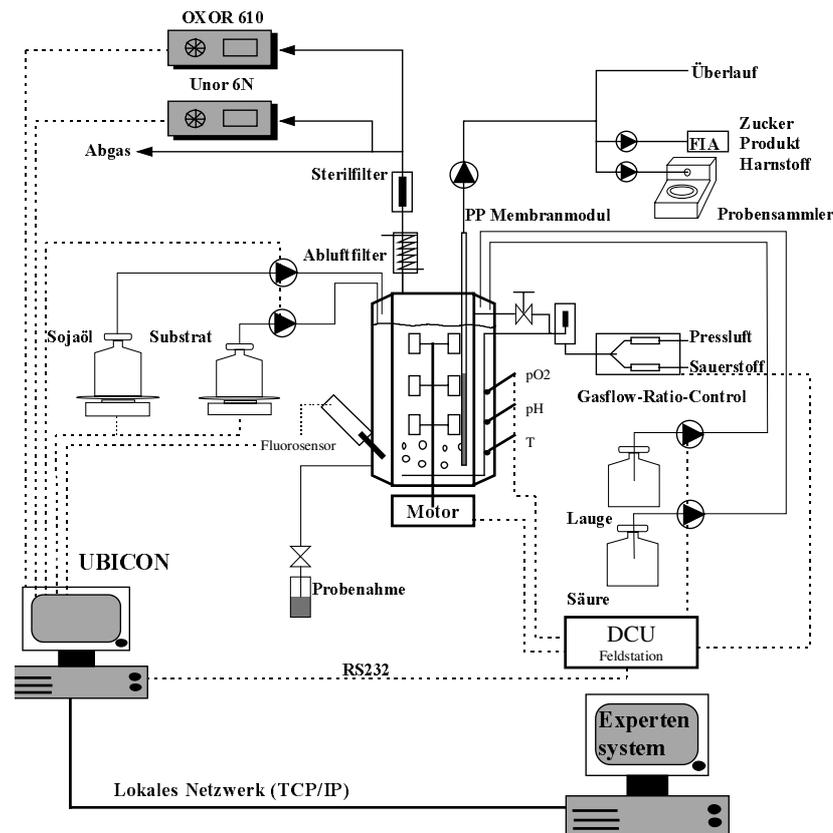


Abbildung 3-1 Schema der Kultivierungsanlage mit Peripherie

3.2 Prozeßleitsystem

Die Datenerfassung und Speicherung erfolgt über das Prozeßleitsystem UBICON (Fa. ESD, Hannover). Es basiert auf einem Industrierechner 68030 mit VME-Bus-Architektur und ist mit einer ASIO-16- (16 RS232-Schnittstellen), einer DIOC48- (36 digitale Ein- und 12 Ausgänge), einer DAC-812- (8 D/A-Kanäle) und einer ADTH-16-Karte (16 A/D-Kanäle) ausgestattet. Das Echtzeit-Multitasking-Betriebssystem RTOS-UH und die Programmiersprache Pearl finden hierfür Verwendung. Über einen leistungsfähigen Formelinterpreter können virtuelle Größen wie Kohlenstoffproduktionsrate (CPR), Sauerstofftransferrate (OTR) und Respirationsquotient (RQ) in Echtzeit berechnet, sowie Regelalgorithmen und Fütterungsstrategien realisiert werden. Die Datenspeicherung erfolgt üblicherweise über Minutenmittelwerte, kann jedoch bei Bedarf mit einer Frequenz von bis zu 1 s^{-1} erfolgen. Für alle Prozeßgrößen können Warn- und Alarmgrenzen gesetzt werden, bei deren Überschreiten bzw. Unterschreiten der Bediener eine entsprechende Mitteilung erhält. Operatoreingaben sind jederzeit möglich.

3.3 Netzwerk

Über einen Netzwerkadapter und einem TCP-Server ist es möglich, extern auf das System zuzugreifen. Die digitale Kontrolleinheit (DCU) des Reaktors ist über eine serielle RS232-Schnittstelle an das Prozeßleitsystem UBICON gekoppelt. Über einen Softwaretreiber werden mit einer Frequenz von 1 s^{-1} die aktuellen Daten der DCU eingelesen und ggf. Regler-Sollwerte an die DCU übermittelt.

3.4 Expertensystem Entwicklungsumgebung

Das Expertensystem wurde auf einer Digital AlphaVAX-Station DEC 3000 mit 96 MB Hauptspeicher und 3 GB Plattenspeicher unter dem Betriebssystem OpenVMS (64 Bit) mit der Echtzeit-Expertensystemshell G2 (GENSYM) entwickelt.

G2 ist ein Werkzeug für die Entwicklung von Echtzeit-Expertensystemen komplexer Anwendungen, die eine kontinuierliche und intelligente Überwachung, Diagnose und Regelung benötigen [G2 User Manual]. G2 vereint viele Eigenschaften, die für die Realisierung eines on line arbeitenden Expertensystems unverzichtbar sind. Hierzu gehört regelbasiertes Schlußfolgern (Inferenzmaschine), prozedurale Datenverarbeitung, Netzwerkfähigkeiten, Datenein- und -ausgabemöglichkeiten, objektorientierte Programmierweise, modularer Systemaufbau und graphische Benutzerschnittstelle.

Alle G2-Wissensbasen werden auf rechteckigen Arbeitsbereichen den WORKSPACES entwickelt. Auf dem Arbeitsbereich müssen Elemente der Wissenbasis plaziert und verwaltet werden. Dynamisch erzeugte – transiente – Elemente müssen nicht auf einem Arbeitsbereich abgelegt sein, sie werden aber beim Zurückstellen der Anwendung gelöscht. Elemente in G2 teilen sich in zahlreiche Unterklassen. Zu den Wichtigsten zählen die Regeln, die Prozeduren, die Anzeige-Werkzeuge und die Objekte. Jedes Element unter G2 hat eine Eigenschaftstabelle, welche die Information in Bezug auf dieses Element enthalten. Die Eigenschaften (Attribute) eines Elements sind abhängig von der Klassenzugehörigkeit.

G2 bietet vielfältige Möglichkeiten Wissen zu implementieren und zu verarbeiten. Die Inferenzmaschine erlaubt in G2 sowohl daten- als auch zielgetriebene Schlußfolgerungen. Methoden Wissen wird in G2 in Form von Regeln, Faktenwissen in Form von Objekten implementiert.

Ein Objekt definiert sich aus seiner Klassenzugehörigkeit, seinen spezifischen und vererbten Eigenschaften (Attribute) und den Werten, die den Attributen zugeordnet sind. Die Attribute können selbst wieder Objekte umfassen. Eine herausragende Eigenschaft von G2 ist die Möglichkeit, sowohl in Regeln als auch in Prozeduren Operationen auf ganze Objektklassen zu beziehen. Jede Regel und Prozedur innerhalb von G2 ist selbst ein eigenständiges Objekt, dessen Eigenschaften wie z. B. die Zeit und Art der Überprüfung durch die Inferenzmaschine (Regeln), Kategorien (Regeln) und Prioritäten (Prozeduren und Regeln) individuell eingestellt werden können. Um die Echtzeitfähigkeit des Systems zu gewährleisten, ist G2 mit einem eigenen Time-scheduler ausgestattet. Über eine integrierte Schnittstelle ist es möglich, externe Prozeduren auszuführen und Daten ein- und auszulesen. In Tabelle 3-2 sind die wichtigsten Eigenschaften der verwendeten Expertensystemshell zusammengefaßt.

Tabelle 3-2 Eigenschaften der Expertensystemshell G2

Wissensbasiert Leistungsfähige Inferenzmaschine Zahlreiche Regeltypen	Objektorientiert Alle Elemente des Systems als Objekte implementiert Objekt Hierarchien mit Vererbung
Echtzeitfähig Jedes Objekt ist mit Prioritäten versehen Integrierter Echtzeit-Daten-Server	Graphisch Alle Objekte graphisch dargestellt Vielfältige Objekte zur Visualisierung und Animation
Simulation Integrierter Simulator mit Daten-Server	Kommunikation Schnittstelle implementiert Anbindung externer Routinen möglich

4 Ein Expertensystem zur Überwachung von Bioprozessen

In diesem Kapitel wird das Expertensystem zur Überwachung von Bioprozessen vorgestellt, daß im Rahmen dieser Arbeiten, entwickelt wurde. Die zentrale Aufgabe der Expertensystementwicklung ist die Wissensakquisition, auf deren Methoden und Ergebnisse zunächst näher eingegangen wird. Da die Meßdaten des Prozeßleitsystems die wesentliche Informationsquelle für das Expertensystem bei der On-line-Ermittlung des Prozeßzustand sind, bilden die vorgestellten Methoden zur Datenanalyse eine wichtige Voraussetzung für die Unterstützung des Bedieners. Die in der Wissensakquisition gewonnenen Erkenntnisse werden bei der anschließenden Konzeption des Expertensystems berücksichtigt. Anschließend erfolgt die Beschreibung des Expertensystems in Hinsicht auf die Konzepte, die zur Problemlösung mit den Mitteln der Entwicklungsumgebung praktisch umgesetzt wurden. Zum Abschluß wird an Hand von Anwendungsbeispielen die Leistungsfähigkeit des entwickelten Systems vorgestellt.

4.1 Wissensakquisition

Die Wissensakquisition wird die Wissenserhebung, die Analyse und Abbildung des Wissens durch verschiedene Techniken der Wissensrepräsentation in einer Wissensbasis bezeichnet [Curth et al. 1990]. Das Ergebnis der Wissensakquisition stellt eine Konzeptionalisierung der Expertise innerhalb einer beschränkten Domain dar [Stender 1989]. Dieses Modell gibt sprachliche und formale Elemente wieder, die allgemein nicht ausführbar sind. Es gibt erst wenige ausgereifte Methoden zur Wissensakquisition, weshalb man auch von dem Flaschenhals der Expertensystementwicklung spricht [Sundermeyer 1991]. Mit Ansätzen wie dem maschinellen Lernen wird beispielsweise versucht, durch eine Automatisierung der Wissensakquisition, diesen Hals zu erweitern. Für die Wissensakquisition sind verschiedene Techniken und Methoden entwickelt worden, die von Karbach und Linster ausführlich diskutiert werden [1990]. Diese Techniken eignen sich auch für die Entwicklung von Echtzeit-Expertensysteme. Allerdings sind in diesem Anwendungsbereich verschiedene Besonderheiten zu beachten, welche in erster Linie

auf den unterschiedlichen Informationsquellen beruhen, die einem menschlichen Experten bei der Überwachung und Fehlerdiagnose zur Verfügung stehen. Die Verfügbarkeit unterschiedlicher Informationsquellen für die Diagnose hat einen großen Einfluß auf die anzuwendende Strategie bei der Problemlösung.

Im wesentlichen gibt es vier Hauptmethoden, um Wissen zu erheben:

- Literaturstudium und die Analyse bestehender Programme
- Interviewtechniken
- Wissenserwerbssysteme
- Induktion

Bei der Interviewtechnik steht ein spezieller Wissensingenieur zur Erhebung des Wissens den Experten zur Seite und versucht, das für die Problemlösung notwendige Wissen zu erfragen. Die Antworten der Experten stellen dabei einen formalen Zwischenschritt der Wissenserhebung dar, die durch den Wissensingenieur noch weiter analysiert werden müssen. Bei der Interviewtechnik treten folgende Probleme auf:

- Wissensingenieure sind bei der objektiven Wissenserhebung mit Beschränkungen und Vorurteilen behaftet.
- Menschliche Experten begründen häufig nicht in einer musterhaften Weise.
- Expertenwissen ist vielfach unbewußt gespeichert und kann deshalb nur mit Schwierigkeiten erhalten werden.
- Experten haben nicht selten Probleme ihr Handeln zu begründen.

Eine Abhilfe für diese Beschränkungen bieten Wissenserwerbssysteme, die den Weg über den Interviewer umgehen. Durch sie erhält der Experte eigene Möglichkeiten und Werkzeuge, um sein Wissen direkt zu formulieren und im Expertensystem zu speichern. Je besser ein Werkzeug an das spezielle Problemfeld angepaßt ist, desto besser ist auch die Unterstützung des Experten. Häufig bietet es sich an, daß ein Wissensingenieur in einem frühen Projektstadium ein spezifisches Werkzeug zur Wissensakquisition erstellt.

Die Induktion verwendet maschinelle Lerntechniken, die eine vollständige Automatisierung der Wissensakquisition ermöglichen sollen. Nur einfache induktive Lerntechniken bieten zur Zeit die Möglichkeit die Wissensakquisition partiell zu automatisieren. Manche Systeme sind z. B. in der Lage aus Fallbeispielen Regeln abzuleiten. Diese Technik kann nur in Fällen eingesetzt werden, in denen eine vollständige, repräsentative Sammlung von Fallbeispielen – die von einem Experten klassifiziert wurden – vorliegt.

Die in dieser Arbeit erstellten Wissenserwerbssysteme beinhalten dialogorientierte Benutzerführungen, die speziell zur Akquisition von allgemeinem Wissen des Problemfelds und methodischem Wissen zur Problemlösung konzipiert sind. Diese sind in dem Kapitel 4.4.5.4 beschrieben. Für das spezielle Gebiet der Echtzeit-Expertensysteme, in denen die On-line-Daten eine wichtige Informationsquelle des Problemfelds darstellen, bieten sich induktive Ansätze besonders an, um partiell die Wissenserhebung zu automatisieren.

Die Wissensakquisition verfolgt in dieser Arbeit zwei Ziele. Zum einen soll mit dem Expertensystem Wissen zur Verfügung gestellt werden, welches allgemein eine Überwachung von Bioprozessen zuläßt. In diesen Bereich gehört z. B. das gesamte Wissen der Ablaufsteuerung und Benutzerführung. Dieses allgemeine Wissen kann in dem System statisch implementiert sein.

Im Gegensatz zu dem allgemeinen Wissen steht das spezielle Prozeßwissen, das getrennt gespeichert und die prozeßspezifische Wissensbasis darstellt. Bei einer Anpassung des Expertensystems an einen neuen Prozeß erfordert dies die neue Akquisition dieses prozeßspezifischen Wissens.

Eine andere Möglichkeit das Wissen, das zur Überwachung von Bioprozessen benötigt wird, zu unterteilen ist in die beiden Teilbereiche des analytischen Wissens und empirischen Wissens. Hierbei umfaßt das analytische Wissen funktionelle Zusammenhänge und Meßprinzipien, die bei der Durchführung von Bioprozessen zur Anwendung kommen. Es handelt sich um allgemein anerkanntes Wissen, das in der Regel gut strukturiert ist und nur in geringem Maß Unsicherheiten enthält. Das empirische Wissen enthält Wissen über den biologischen Prozeß sowie Zusammenhänge zwischen Prozeßvorgängen und deren Auswirkungen auf die Meßsignale. Als empirisches Wissen (Erfahrungswissen) wird Wissen aus unterschiedlichen Bereichen bezeichnet, das während des langjährigen praktischen Einsatzes eines Prozesses gewonnen wurde. Es umfaßt Faktenwissen, Heuristiken, Methoden und Strategien für die Überwachung, die Fehlerdiagnose und die Fehlerbehandlung. Empirisches Wissen ist häufig vage formuliert und kann widersprüchlich oder mit Unsicherheiten behaftet sein. Trotz der Schwierigkeiten bei der Akquisition und Formalisierung von empirischen Wissen ist die Integration von Erfahrungswissen von großer Bedeutung für Expertensysteme. Das empirische Wissen gliedert sich im wesentlichen in die Prozeßmodelle und die Fehlermodelle. Die Prozeßmodelle enthalten das Wissen über den durchgeführten Prozeß. Die Fehlermodelle enthalten Wissen über die Auswirkung von Störungen auf den Verlauf der Kultivierung und dadurch auch auf die Meßsignale. Abbildung 4-1 zeigt diese Zusammenarbeit des Wissensingenieurs und des Experten bei der

Wissensakquisition. Der Wissensingenieur erlangt durch Interviews Wissen der betrachteten Domäne. Mit diesem Wissen ist der Wissensingenieur in der Lage sowohl die allgemeine Wissensbasis als auch die prozeßspezifische Wissensbasis zu füllen.

Damit die Experten selbständig die Wissensakquisition durchführen können, schafft der Wissensingenieur Werkzeuge, mit denen Experten direkt die prozeßspezifische Wissensbasis füllen können.

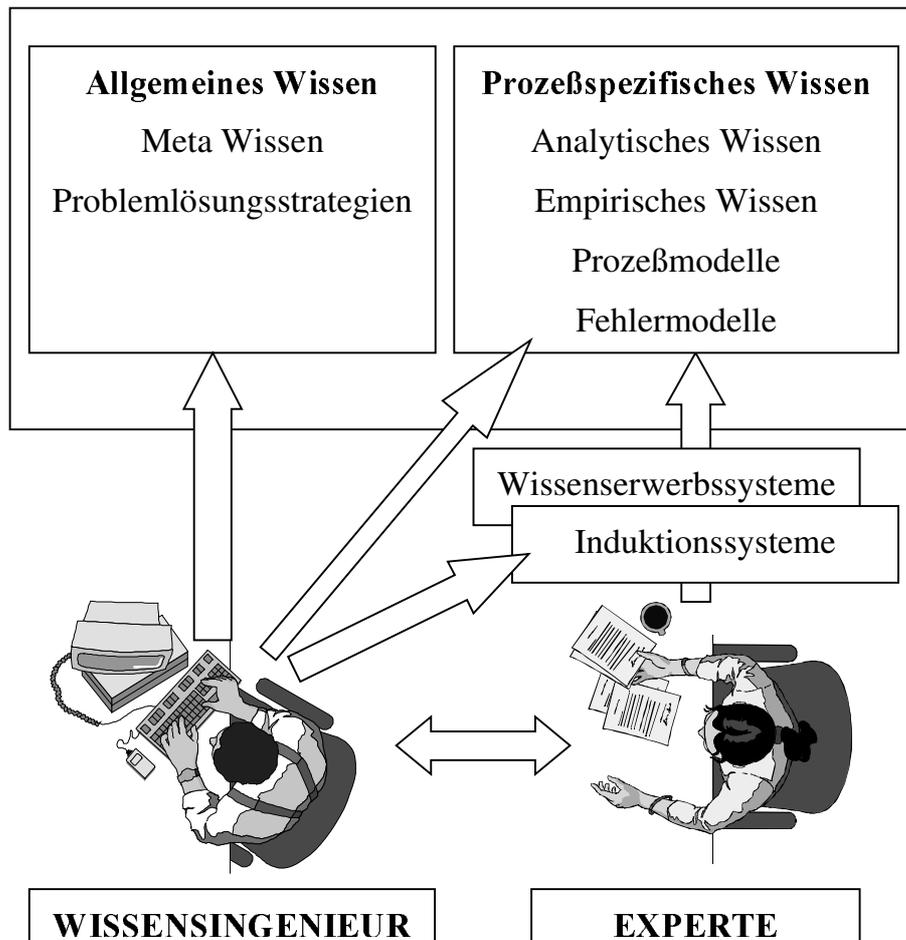


Abbildung 4-1 Zusammenarbeit eines Wissensingenieurs und eines Experten bei der Wissensakquisition

Wegen ihrer Bedeutung für die Überwachung von Bioprocessen soll im folgenden auf das in den Prozeßmodellen enthaltene Wissen näher eingegangen werden.

4.1.1 Heuristische Prozeßmodelle

Die Erstellung von Prozeßmodellen ist nicht nur für die computergestützte Überwachung von Bioprozessen wichtig. Der erfahrene Bediener kann zwar direkt die Prozeßapparatur beobachten; die in dem Reaktor stattfindenden Vorgänge können jedoch nur indirekt über die Signale der angeschlossenen Sensoren verfolgt werden. Anhand der Signalveränderungen ist er in der Lage zu beurteilen, ob der Prozeß den gewünschten Verlauf nimmt. Hierbei nutzt er die Zusammenhänge in der Wirkungskette zwischen dem Zustand des Mikroorganismus, dem Medium, der Meßsensoren und den daraus resultierenden Meßsignalen. Für jeden Übergang gibt es eigene – zum Teil analytische – Zusammenhänge. Der Bediener stellt aber häufig direkt qualitative Zusammenhänge zwischen einem beobachteten Merkmal des Signals und dem Prozeßzustand auf. Im Fall einer Störung müssen zur Fehlerlokalisierung Zusammenhänge der einzelnen Systemübergänge bekannt sein. Abbildung 4-2 verdeutlicht die Wirkungskette zwischen Mikroorganismus und Meßsignal.

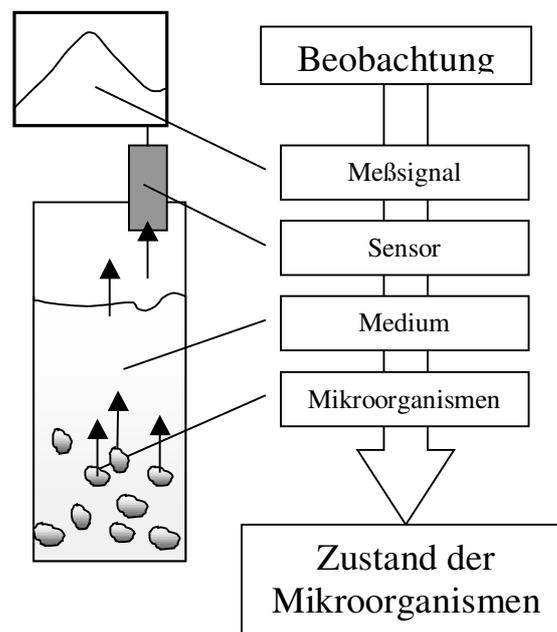


Abbildung 4-2 Der Zustand des Mikroorganismus wirkt sich nach mehreren Übergängen auf das Meßsignal aus; aus den Beobachtungen kann deshalb auf den Zustand zurück geschlossen werden

Bevor die Zusammenhänge zwischen Meßsignalen und Prozeßzustand interpretiert werden können, muß zunächst ein Verständnis über Vorgänge, die bei einer Kultivierung stattfinden, gewonnen werden.

4.1.1.1 Ablauf einer Kultivierung und Phasenunterteilung

Durch die **Mediumsvorbereitung** werden für das Wachstum der Organismen optimale Bedingungen geschaffen. Die wesentlichen Komponenten des Mediums sind die Kohlenstoffquellen (Substrate), Stickstoffquellen, Nährsalze und Wasser. Darüber hinaus muß ein optimaler pH-Wert eingestellt und das gesamte Medium sterilisiert werden.

Der Beginn der Kultivierung zeichnet sich durch das **Animpfen** des Mediums mit der Vorkultur aus. In der Vorkultur sind die Mikroorganismen in einem kleineren Maßstab zu einer hohen Konzentration herangezogen worden, die in dem Medium rasch anwachsen kann. Der Wachstumsprozeß durchläuft in der Regel bestimmte Phasen. In der **Adaptionsphase** –auch Lag-Phase genannt– muß der Mikroorganismus sich zunächst an die neuen Mediumsbedingungen gewöhnen. Hierbei stellt sich der Stoffwechsel auf das günstigste Substrat ein. Dies äußert sich in der Regel durch ein geringes Wachstum. Ist der Stoffwechsel umgestellt, so durchläuft der Prozeß eine **exponentielle Wachstumsphase**, die durch eine konstante Verdopplungsphase (μ -max) gekennzeichnet ist. Dieses Wachstum dauert solange an, wie die Substratkonzentrationen nicht limitierend wirken oder toxische Ausscheidungsprodukte der Mikroorganismen eine optimale Vermehrung verhindern. Darauf folgt die **stationäre Phase**, in der die Vermehrungs- sowie die Absterberaten gleich sind. Übersteigt die Absterberate die Vermehrungsrate, so spricht man von der **Absterbephase**. Sind in dem Medium mehrere Substrate vorhanden, so wird nach der 1. exponentiellen Phase der Stoffwechsel auf ein weiteres Substrat umgestellt. Dies äußert sich in dem Auftreten einer 2. Exponentiellen Phase nach der stationären Phase. Man spricht dann von einer **Diauxie**.

Gegenüber der Vielzahl von Einflußgrößen, die das Stoffwechselverhalten der Organismen während der gesamten Kultivierung bestimmen, ist für den Metabolismus innerhalb einer Prozeßphase nur eine Untermenge von Größen entscheidend. Damit eignet sich die phasenartige Einteilung von Kultivierungen vorzüglich zur Dekomposition des komplexen Gesamtprozesses in eine Sequenz von einfacheren Teilprozessen. Jeder Phase können eigene Modelle, Regelstrategien und andere qualitative Phänomene zugeordnet werden. Zur Beschreibung jedes Abschnitts genügen häufig eine geringe Anzahl der verfügbaren Meßinformation.

In der **Aufarbeitung** müssen die gewünschten Produkte aus dem Medium isoliert werden.

4.1.1.2 Klassifikation des Prozeßzustandes

Bei Batch- und Fed-Batch-Fermentationen resultiert der phasenartige Ablauf im wesentlichen aus dem Verbrauch der Nährstoffe, mit dem die Adaption des Zellmetabolismus korreliert ist. Andere mögliche Ursachen für strukturelle Veränderungen im biologischen System sind z. B. die Anhäufung von toxischen Produkten, Änderungen im Membran-Transport-Mechanismus, morphologische Veränderungen, durch chemische Faktoren oder durch Temperaturgradienten hervorgerufene Genexpression / -repression in rekombinanten Organismen [Gehlen 1992].

Die Qualität der automatischen Erkennung von Stoffwechseleränderungen hängt entscheidend von den verfügbaren Kenngrößen ab. Erfahrenes Bedienpersonal kann bei den meisten Kultivierungen die Prozeßphasen anhand weniger Merkmale sicher erkennen [Halme 1991]. Die für die Klassifikation verwendeten Kenngrößen müssen trotzdem einigen Kriterien genügen [Konstantinov et al. 1990b]:

- Ein klarer, qualitativer Zusammenhang zwischen Kenngröße und Zellzustand muß gegeben sein.
- Die Größen müssen automatisiert berechnet werden können.
- Die Bestimmung muß zuverlässig möglich sein (geringe Störungen).
- Die Kenngrößen müssen innerhalb eines vertretbaren Zeitverzugs vorliegen.

Einige Kenngrößen haben sich in dieser Hinsicht zur Klassifikation des Zellzustandes bewährt [Konstantinov et al. 1990b]:

- Spezifische, auf die Biomasse bezogene Aufnahme-, Transfer- bzw. Produktionsraten (z. B. Sauerstoffaufnahme rate OUR, Kohlendioxidtransferrate CTR, Kohlendioxidproduktionsrate CPR),
- Quotienten von spezifischen Aufnahme-, Transfer- oder Produktionsraten und
- Verhältnisse von spezifischen Aufnahme-, Transfer- oder Produktionsraten, bezogen auf maximale Umsatzraten.

Die Erkennung struktureller Veränderungen in biologischen Systemen auf der Basis nur einer Kenngröße ist wegen im allgemeinen gestörter Meßwerterfassung nicht möglich [Konstantinov et al. 1990b]. Robuste Erkennung erfolgt durch die Einführung alternativer Übergangszustände. Die Nutzung mehrerer paralleler Wechselzustände ist bei zusätzlichen Meßgrößen möglich.

4.2 Datenanalyse

Da einem Expertensystem die audiovisuelle Sensorik eines Menschen fehlt, stehen ihm als einzige Quelle an Information über den aktuellen Betriebszustand nur die On-line-Meßsignale der gemessenen Prozeßgrößen zur Verfügung. Diese Signale liefern wesentliche Information über den Prozeßzustand. Die Erkennung und Charakterisierung von Veränderungen der Meßsignale ist von entscheidender Bedeutung für die Durchführung einer Diagnose. Eine geeignete Beschreibung der Signalverläufe ist sowohl für die Wissensakquisition als auch für die Durchführung der Diagnose und die Kommunikation mit dem Benutzer erforderlich. Innerhalb eines Signalverlaufs erkennt der Bediener Merkmale und Muster, die er bestimmten Ereignissen zuordnet. Aus dem Auftreten und Beurteilen der Merkmale vergangener Kultivierungen ist der Bediener in der Lage, diese Merkmale bestimmten Ereignissen zuzuordnen. Als Kriterium für die qualitative Einschätzung der Merkmale dient neben dem Wert des Merkmals auch das zeitliche Auftreten bestimmter Merkmale. Ein erfahrener Anwender benutzt zur Charakterisierung von Abweichungen der Signale von ihren üblichen Verläufen Begriffe wie z. B. „pH-Signal ist zu hoch“, „pO₂-Signal sinkt nicht“, „pO₂-Signal verrauscht“, „O₂-Signal weist Spikes auf“ oder „Signal steigt zu früh“, die er mit bestimmten Ursachen assoziiert. Für eine dem menschlichen Problemlösungsverhalten analoge Vorgehensweise des Expertensystems müssen Merkmale aus den Signalen gewonnen werden, die den Verlauf der Signale zweckmäßig beschreiben (Merkmalsextraktion). Diese Merkmale sind im allgemeinen numerische Kennzahlen, die mit Methoden der digitalen Signalverarbeitung und der Zeitreihenanalyse aus den Meßdaten gewonnen werden. Für die Entwicklung und Akzeptanz eines Expertensystems ist es erforderlich, die spezielle Fachterminologie des Anwendungsbereichs zu übernehmen [Puppe 1991]. Es ist deshalb notwendig eine weitgehende Abstraktion der numerischen Merkmale zu verbal formulierten, symbolischen Merkmalen anzustreben.

Aufgabe des Überwachungssystems ist es, Merkmale der Signale on line zu berechnen und zu beurteilen. Die Beurteilung der Merkmale durch das System ist aber nur möglich, wenn – analog den Erfahrungen des Bedieners – Wissen zur Beurteilung der Merkmale vorliegen. Dies geschieht über Methoden der Datenanalyse, die prozeßspezifisches Faktenwissen über den Verlauf der Merkmale erzeugen und dem Expertensystem zur Verfügung stellen. Das Kapitel gibt zunächst einen Überblick des Stands der Technik in der Datenanalyse. Danach wird die Charakterisierung der Meßsignale in dem Expertensystem beschrieben. Anschließend werden

die Methoden erklärt, mit denen innerhalb des Expertensystems Datenanalysen durchgeführt werden und welches Wissen über den Signalverlauf daraus generiert wird.

4.2.1 Charakterisierung der Meßsignale

Die Signale, die dem Expertensystem zur Beurteilung des Prozesses zur Verfügung stehen, stellen diskrete Funktionen $S(n\Delta t)$ der Zeit dar, wobei $1/\Delta t$ die von dem Prozeßleitsystem vorgegebene Meßfrequenz und n die Anzahl der zum aktuellen Zeitpunkt t gemessenen Werte darstellt. Die Merkmale, die zur Beurteilung des Signals herangezogen werden, können in zwei Gruppen gegliedert werden.

1. Merkmale zur Charakterisierung eines Einzelwertes
2. Merkmale zur Charakterisierung des zeitlichen Verlaufs eines Meßsignals

Die Merkmale, welche für die Charakterisierung des zeitlichen Verlaufes des Meßsignals verwendet werden, sind der Rohwert, die Zeit des Werts, der Meßwertabstand und die Distanz zur Meßbereichsgrenze. Abbildung 4-3 stellt die Merkmale zur Charakterisierung eines einzelnen Meßwerts dar. Die Merkmale erlauben eine schnelle Überprüfung des Meßsignals bezüglich seiner Validität.

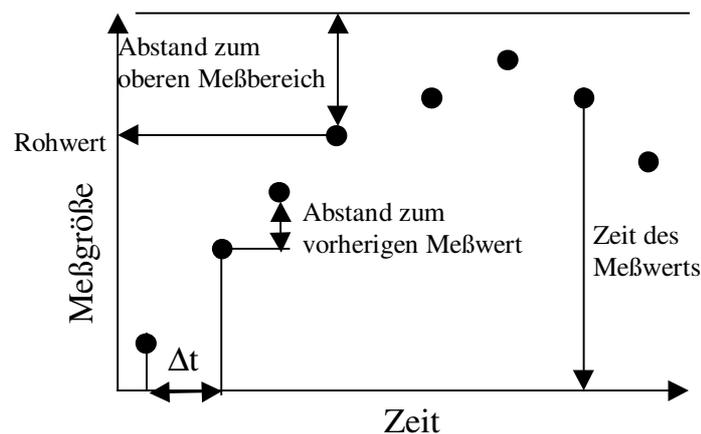


Abbildung 4-3 Merkmale zur Charakterisierung eines einzelnen Meßwerts

Die Merkmale, welche für die Charakterisierung des zeitlichen Verlaufs des Meßsignals berechnet werden, sind der Mittelwert einer Stunde, Median, Rauschfaktor, Signal-Rauschverhältnis, Varianz, Standardabweichung, lineare Steigungen unterschiedlicher Zeitintervalle und

die Anzahl der Wertwiederholungen. In Abbildung 4-4 sind die Merkmale zur Charakterisierung des zeitlichen Verlaufs eines Meßsignals dargestellt. Diese Merkmale besitzen z. T. einen höheren Informationsgehalt als die Merkmale des Einzelwerts. Sie werden dafür aber nur mit einer größeren zeitlichen Verzögerung erhalten.

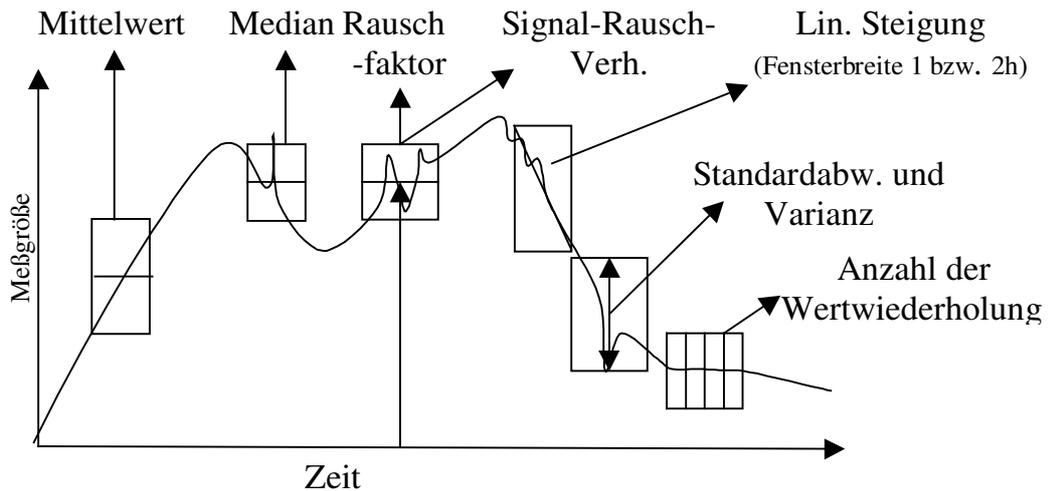


Abbildung 4-4 Merkmale zur Charakterisierung des zeitlichen Verlaufs des Meßsignals

Die berechneten numerischen Merkmale stellen unterschiedliche Informationen über das Meßsignal zur Verfügung. So sind der Mittelwert und Median des Meßsignals Kennzahlen, die den Einfluß hochfrequenter Störgrößen eliminieren und somit eine zuverlässige Aussage über die Meßgröße zulassen. Im Gegensatz dazu gibt der Rauschfaktor einen Hinweis auf die hochfrequenten Störsignale. Zu bedenken ist aber, daß die Berechnung der Kenndaten immer mit einem Informationsverlust behaftet ist. Die Eliminierung von Rauschen durch Filter ist z. B. mit einer Verzerrung des ursprünglichen Signals behaftet, die sehr stark von dem verwendeten Filter und den Filterparametern abhängt. Die Ermittlung der linearen Steigung über eine Stunde – welche Informationen über die Änderung des Meßsignals gibt – spiegelt nur die mittlere Änderungsrate der Meßdaten einer Stunde wieder. Hier werden beispielsweise kurzfristige mittlere Änderungen nicht erfaßt, die besser durch den Meßwertabstand erfaßt werden. Abbildung 4-5 zeigt den Einfluß eines Mittelwert- und Median-Filters auf den Verlauf von O_2 -Daten. Man sieht, daß der Median-Filter gut geeignet ist, typische Impulstörungen des Signals – wie sie in der Biotechnik häufig anzutreffen sind – zu eliminieren und darüber hinaus, eine geringere zeit-

liche Verzögerung aufweist als der Mittelwert-Filter [Gollmer 1996]. In der Tabelle 4-1 und Tabelle 4-2 sind die Formeln zur Berechnung der Merkmale zusammengestellt.

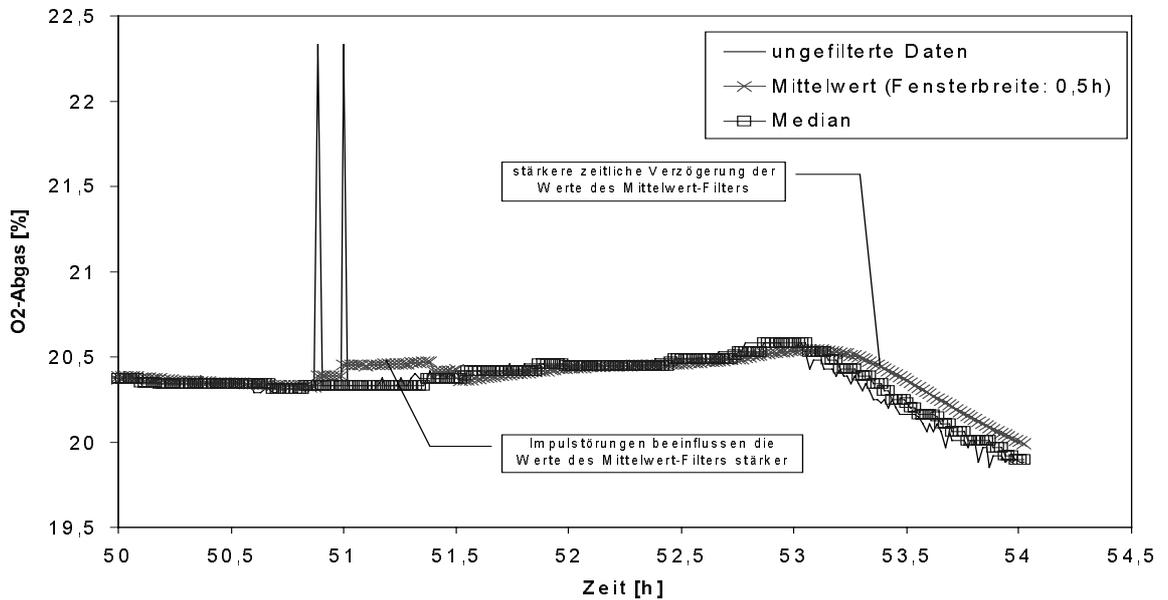


Abbildung 4-5 Vergleich zwischen Mittelwert- und Median-Filter

Tabelle 4-1 Formeln der numerischen Merkmale, die den Einzelwert eines Meßsignals charakterisieren

Merkmal	Berechnung
Meßfrequenz	$1/\Delta t$
Rohwert	$S(n\Delta t)$
Zeit	$n\Delta t$
Abstand zum oberen Meßbereich	$Abs_{Me\beta b.}(n\Delta t) = Max - S(n\Delta t)$
Abstand zum vorherigen Wert	$Abs_{Me\beta b.}(n\Delta t) = S((n-1)\Delta t) - S(n\Delta t)$

Tabelle 4-2 Formeln der numerischen Merkmale, die den zeitlichen Verlauf eines Meßsignals charakterisieren

Merkmal	Berechnung
Mittelwert	$av(n\Delta t) = \frac{\sum_{i=n-N}^{i=n} S(i\Delta t)}{N+1}$
Standardabweichung	$STD(n\Delta t) = \sqrt{\frac{(N+1) \sum_{i=n-N}^{i=n} S(i\Delta t)^2 - \left(\sum_{i=n-N}^{i=n} S(i\Delta t) \right)^2}{N(N+1)}}$
Varianz	$VAR(n\Delta t) = \frac{(N+1) \sum_{i=n-N}^{i=n} S(i\Delta t)^2 - \left(\sum_{i=n-N}^{i=n} S(i\Delta t) \right)^2}{N(N+1)}$
Rauschfaktor	$R(n\Delta t) = \frac{\sum_{i=n-N}^{i=n} ABS(S(i\Delta t) - S((i-1)\Delta t))}{N+1}$
Signal-Rauschverhältnis	$SRV(n\Delta t) = \frac{\sum_{i=n-N}^{i=n} ABS(S(i\Delta t) - S((i-1)\Delta t))}{S(n\Delta t)N+1}$
Lineare Steigung	$m(n\Delta t) = \frac{(N+1) \left(\sum_{i=n-N}^{i=n} (i\Delta t) S(i\Delta t) \right) - \left(\sum_{i=n-N}^{i=n} (i\Delta t) \right) \left(\sum_{i=n-N}^{i=n} S(i\Delta t) \right)}{N+1 \left(\sum_{i=n-N}^{i=n} (i\Delta t)^2 \right) - \left(\sum_{i=n-N}^{i=n} (i\Delta t) \right)^2}$

4.2.2 Berechnung von Richtwerten und Intervallen zur qualitativen Beurteilung und Klassifizierung von numerischen Merkmalen

Die Beurteilung und Klassifizierung der berechneten numerischen Merkmale erfolgt durch Vergleiche mit Richtwerten. So stellt z. B. das Merkmal ‚Abstand zur Meßbereichsgrenze‘ voraus, daß die Meßbereichsgrenze bekannt ist. Hierbei teilt der Richtwert ‚oberer Meßbereich‘ die möglichen Merkmalswerte in zwei Bereiche ein, für die jeweils ein eigener Zustand definiert ist. So liegt der ‚normale‘ Zustand vor, wenn der Meßwert unterhalb des oberen Meßbereichs liegt und damit die Kennzahl größer Null ist. Der ‚unnormale‘-Zustand liegt vor,

wenn der Wert den oberen Meßbereich überschreitet – d. h. die Kennzahl kleiner als Null ist. Diese Klassifizierung von Merkmalen entspricht dem Vorgehen des menschlichen Experten, der Merkmale in ähnlicher Weise qualitativ beurteilt. In dem Expertensystem wird die Zugehörigkeit zu einer bestimmten Klasse durch ein Symbol repräsentiert, das den Zustand näher charakterisiert.

Für die meisten Merkmale bietet es sich an, einen Konfidenzbereich zu definieren, der durch ein Intervall festgelegt wird. Die Merkmalswerte werden durch den Konfidenzbereich in drei Bereiche unterteilt und somit können auch drei Zustände eingenommen werden. In Abbildung 4-6 ist diese Unterteilung für die Werte eines Merkmal dargestellt. Zum einen kann der Wert des Merkmals innerhalb des Konfidenzbereichs liegen. Dieser Zustand wird als ‚normal‘ bezeichnet. Zum anderen liegt der Zustand ‚hoch‘ bzw. ‚niedrig‘ vor, wenn sich die Werte des Merkmals ober- bzw. unterhalb des Bereichs befindet.

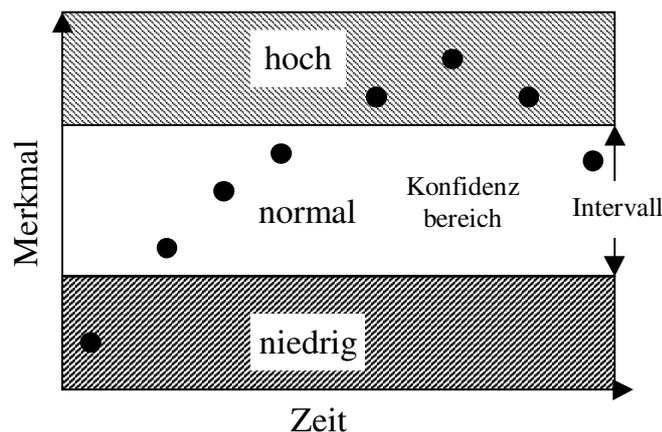


Abbildung 4-6 Der Konfidenzbereich wird durch ein Intervall definiert; dadurch können die Werte des Merkmals drei Zustände einnehmen, denen die Symbole ‚hoch‘, ‚normal‘ und ‚niedrig‘ zugeordnet werden

Da der Verlauf der Kultivierung starken zeitlichen Schwankungen unterworfen ist, ergibt sich daraus, daß die Richtwerte zur Klassifizierung der numerischen Merkmale auch zeitlichen Änderungen unterworfen sein können. Es hat sich als sinnvoll herausgestellt, die Richtwerte nach ihrer zeitlichen Abhängigkeit in drei Klassen zu unterteilen die entweder:

- über den gesamten Verlauf der Kultivierung konstant,
- über einen zeitlichen Signalabschnitt konstant oder
- zeitlich nicht konstant sind.

In der Datenanalyse wird versucht, anhand schon durchgeführter Kultivierungen die Richtwerte und Intervallgrenzen der Konfidenzbereiche der numerischen Merkmale zu berechnen. Grundvoraussetzung ist, daß die Kultivierungen schon möglichst häufig und reproduzierbar durchgeführt wurden. Für die Berechnung der zeitunabhängigen Richtwerte wird der gesamte Signalverlauf in Betracht gezogen. Diese Richtwerte erlauben eine grobe Einschätzung des Signals. Die berechneten Richtwerte sind:

Absolutes Maximum, absolutes Minimum, größte Steigung, kleinste Steigung, maximaler Meßwertabstand, maximale Rauschen, maximale Anzahl an auftretenden Wertwiederholungen, maximale Varianz.

Die Richtwerte, die zur näheren Charakterisierung einer Signalphase berechnet werden, setzen eine zuverlässige Ermittlung der Signalphasen voraus. Der in dieser Arbeit verfolgte Ansatz zur Unterteilung der Signale in charakteristische Abschnitte wird im Kapitel 4.2.3 beschrieben.

Die Richtwerte, die für jede Signalphase ermittelt werden, sind:

Maximaler und minimaler Wert des Signalabschnitts, Maximale und minimale Werte am Anfang bzw. am Ende des Signalabschnitts, maximale und minimale Steigung des Signalabschnitts, maximale und minimale Steigung am Anfang bzw. am Ende des Signalabschnitts.

Konfidenzbereiche

Um den zeitlichen Verlauf sowie die große Varianz von Kultivierungen bei der Klassifizierung der Meßsignale zu berücksichtigen, werden Konfidenzbereiche aus den Daten berechnet. Die Konfidenzbereiche sind durch eine obere und untere Intervallgrenze (Int_{max} und Int_{min}) definiert, die Funktionen der Zeit darstellen. Abbildung 4-7 zeigt den Vorgang der Berechnung der Intervalle aus den Signalverläufen mehrerer Kultivierungsläufe. Zunächst wird mit den in Kapitel 4.2.1 beschriebenen Formeln der zeitliche Verlauf der Merkmale berechnet. Im folgenden Schritt wird aus Gründen der Datenreduktion für den Zeitraum von jeweils einer Stunde der Mittelwert, die Standardabweichung und der minimale und maximale Wert des Merkmals berechnet. Im letzten Schritt werden die Merkmale mehrerer Kultivierungen zu Intervallen zusammengefaßt, indem für jeweilige Zeitabschnitte der Gesamtmittelwert, die Gesamtstandardabweichung sowie das absolute Maximum und absolute Minimum des Merkmals aller Kultivierungen in diesem Zeitabschnitt berechnet wird. Die obere Intervallgrenzen werden durch den Gesamtdurchschnitt zuzüglich der dreifachen Standardabweichung definiert. Analog ist die

untere Intervallgrenze des Konfidenzbereichs durch den Gesamtdurchschnitt minus der dreifachen Standardabweichung festgelegt.

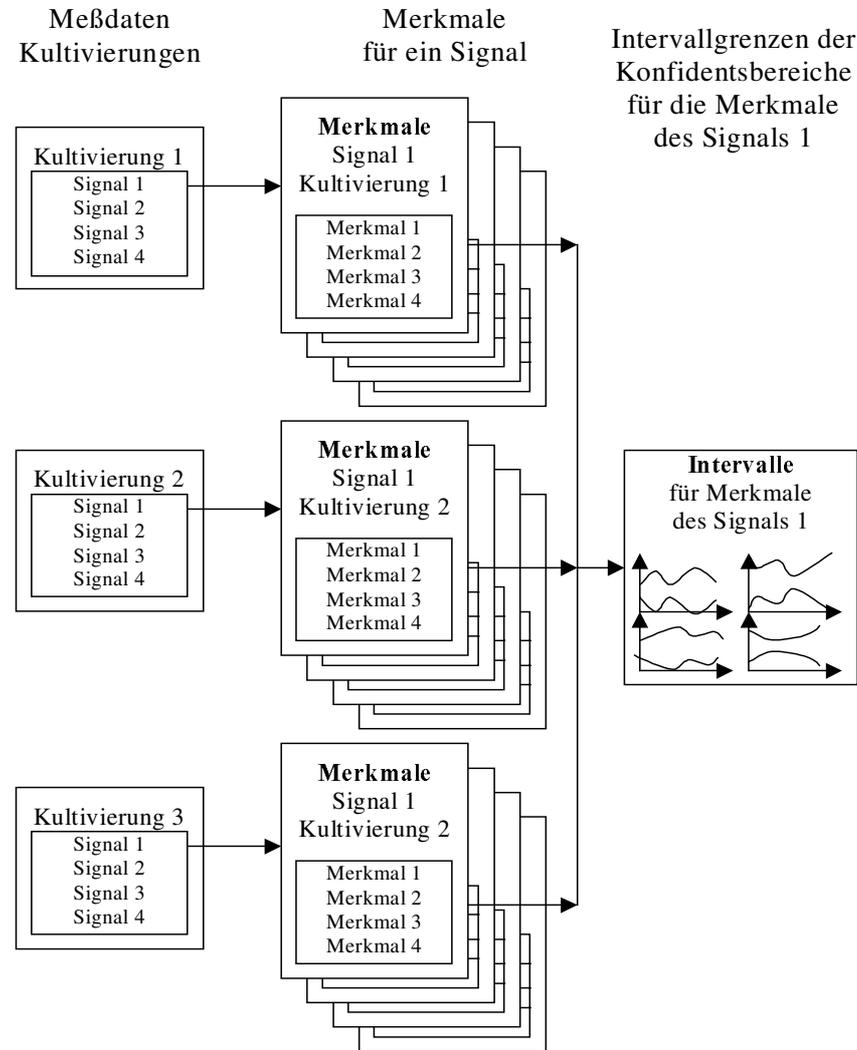


Abbildung 4-7 Berechnung von Intervallgrenzen der Konfidenzbereiche für die Merkmale eines Signals aus den Signalverläufen mehrerer Kultivierungen

In Abbildung 4-8 sind die Intervallgrenzen, die für den CO₂-Wert zweier Kultivierung ermittelt wurden, dargestellt. Es ist deutlich zu sehen, daß die Intervalle für Zeitbereiche, in denen sich das Signal stark verändert, auch die Intervallgrenzen stark aufgeweitet sind. Trotzdem kann mit Hilfe der Intervalle der Wert eines Merkmals qualitativ gut eingeschätzt werden. Die Berechnung der Intervalle stellt sich in der vorgestellten Form als sehr empfindlich gegenüber zeitlichen Verschiebungen des Kultivierungsverlaufs dar. Diese zeitlichen Verschiebungen sind häufig in zeitlich unterschiedlichen lang verlaufenden Lag-Phasen begründet, die z. B. auf eine längere Anpassungszeit der Organismen an das Medium zurückzuführen sein können.

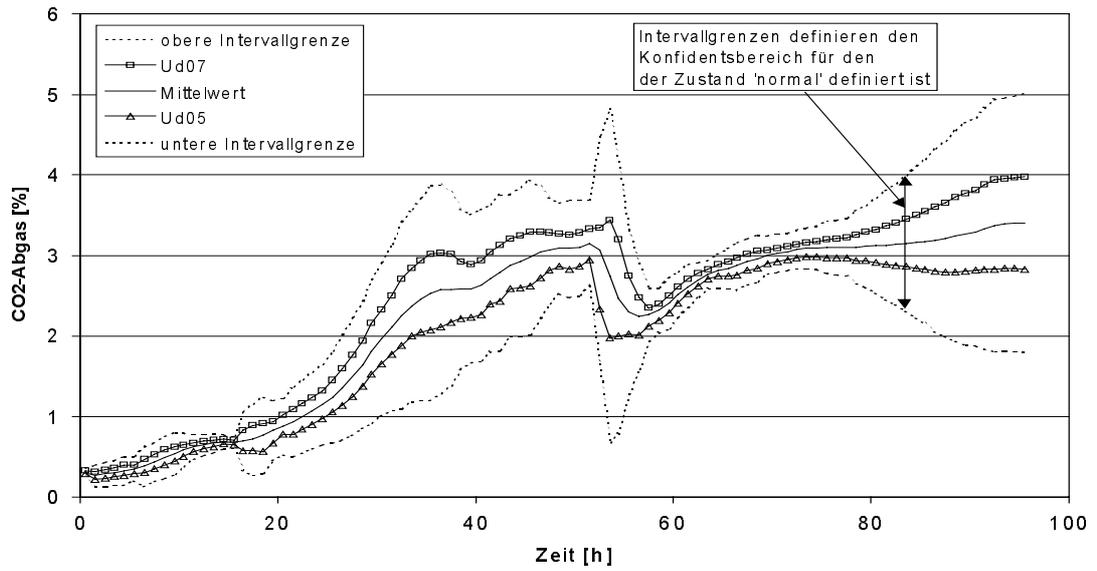


Abbildung 4-8 Intervallgrenzen für den Konfidenzbereich der CO₂-Werte der zwei Kultivierungen UD05 und UD07

Abbildung 4-9 zeigt den Verlauf der CO₂-Werte zweier Hefe-Kultivierungen [Kirchner, 1997], die unter reproduzierbaren Bedingungen durchgeführt wurden.

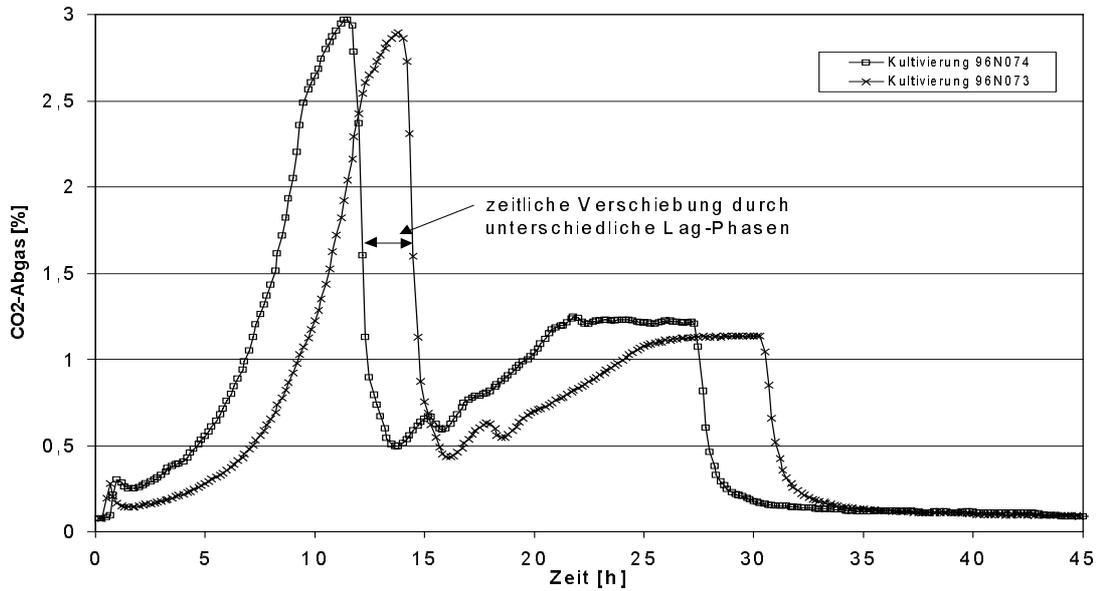


Abbildung 4-9 Verlauf der CO₂-Werte zweier Hefe-Kultivierungen, die durch eine zeitlich unterschiedlich lange Lag-Phase gekennzeichnet sind

Man kann erkennen, daß sich beide Verläufe sehr stark bis auf eine zeitliche Verzögerung von ca. 2 Stunden ähneln. Wird diese Verzögerung in der Berechnung der Intervalle nicht berücksichtigt, so erhält man dementsprechend breite Intervalle. Eine Möglichkeit zur Berücksichtigung zeitlicher Verschiebungen der Kultivierungsverläufe besteht in der Berechnung der maximalen Überlappung der Signale mit Hilfe der Kreuzkorrelation.

Nach Formel 4-1 werden zwei momentane Amplituden der Funktion $y_1(t)$ bei t' und $y_2(t)$ bei $t' + t''$ multipliziert und das Produkt im ganzen zur Verfügung stehenden Bereich von t' , von $-\tau$ bis τ integriert.

$$\phi_{y_1, y_2}(t'') = \lim_{\tau \rightarrow \infty} \left(\frac{1}{2\tau} \int_{-\tau}^{\tau} y_1(t') y_2(t' + t'') dt' \right) \quad \text{Formel 4-1}$$

Die Kreuzkorrelationsfunktion stellt also als Funktion der Zeitdifferenz t'' den zeitlichen Mittelwert des Produktes der beiden Funktionen dar. Die Daten werden vor der Berechnung der Kreuzkorrelation normiert. In Abbildung 4-10 ist die Kreuzkorrelationsfunktion für den CO_2 -Wert der beiden Kultivierungen dargestellt. Man sieht, daß die maximale Überlappung der Signale bei einer zeitlichen Verschiebung von 2,2 h erreicht ist.

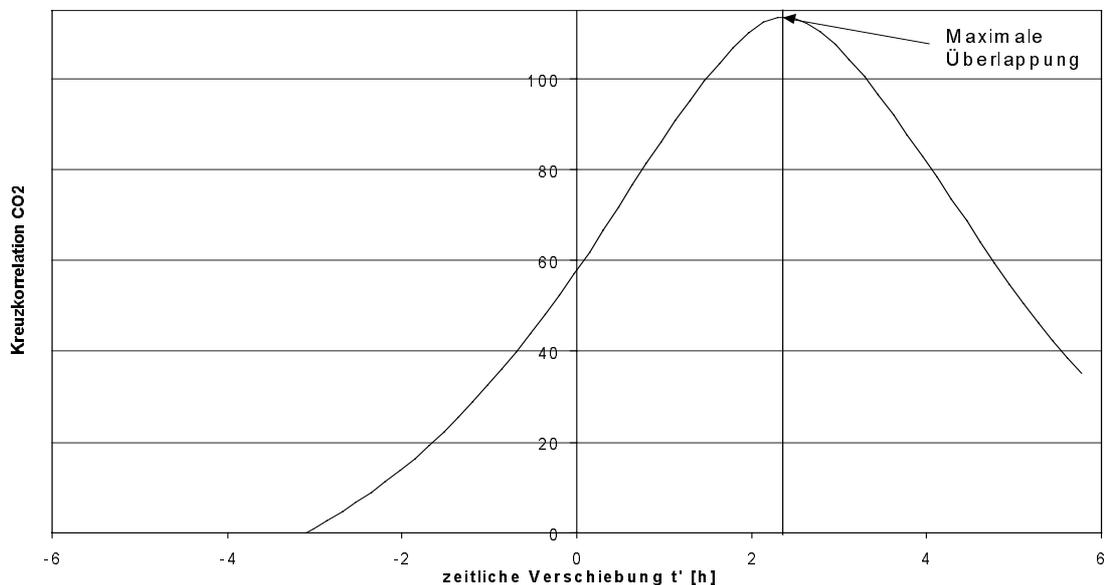


Abbildung 4-10 Kreuzkorrelation für den CO_2 -Wert der beiden Hefe-Kultivierungen 96N074 und 96N073

Abbildung 4-11 zeigt die Intervallgrenzen, die erhalten werden, wenn die zeitliche Verschiebung nicht berücksichtigt wurde. An dem Verlauf des Mittelwerts ist erkennbar, daß der charakteristische Verlauf des Signalverlaufs nur grob wiedergegeben wird.

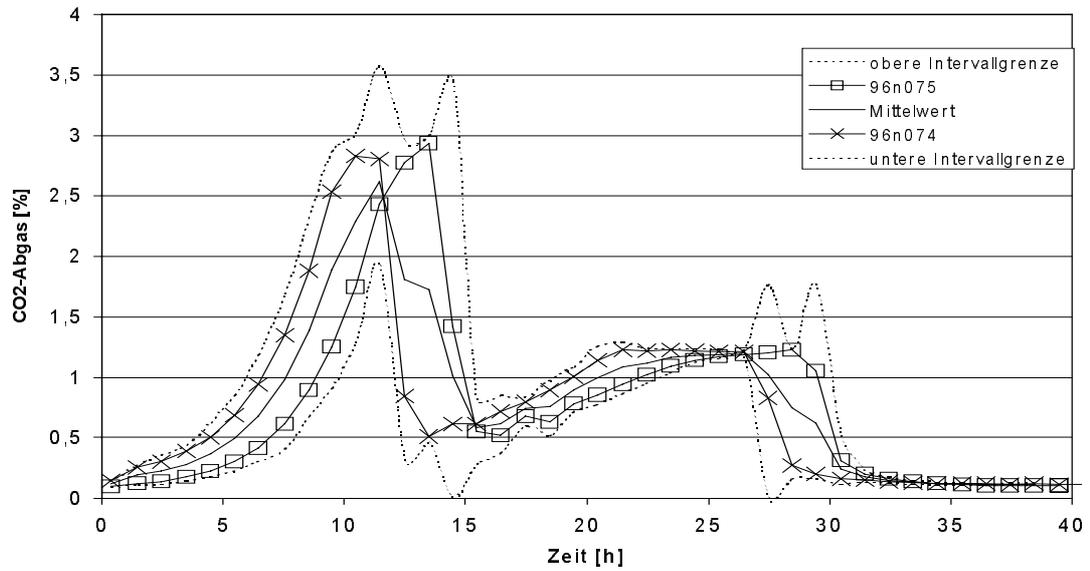


Abbildung 4-11 Intervalle berechnet für zwei Kultivierungen ohne Berücksichtigung der Lag-Phase

In Abbildung 4-12 hingegen sind die Intervallgrenzen unter Berücksichtigung der zeitlichen Verschiebung dargestellt. Der Mittelwert gibt jetzt den wesentlichen Charakter des Verlaufs wieder und die Intervallgrenzen liegen viel enger im Bereich des Signals, wodurch eine genauere Beurteilung des Signals ermöglicht wird.

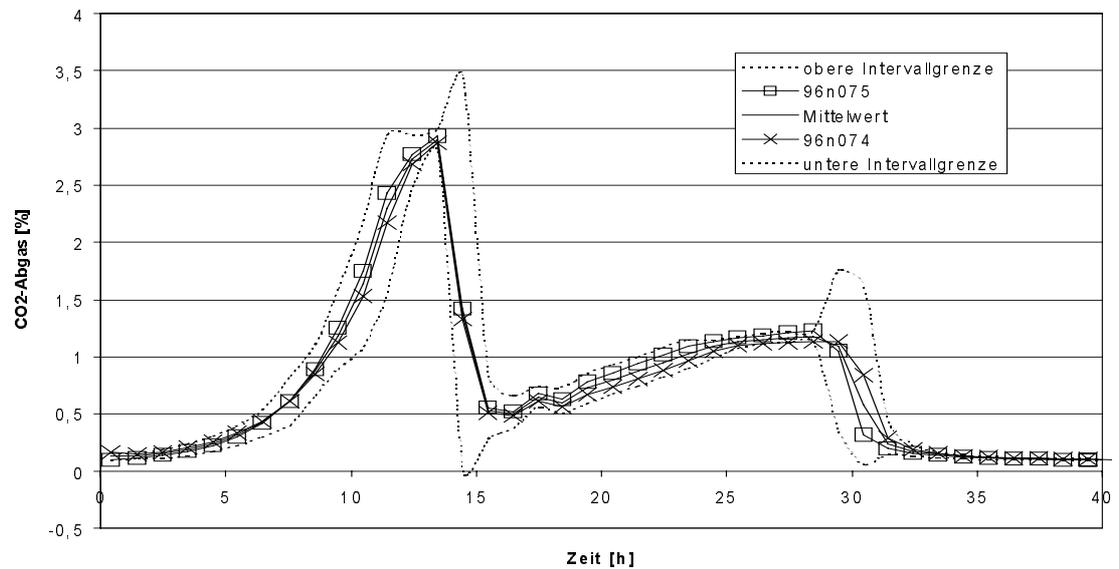


Abbildung 4-12 Intervalle unter Berücksichtigung der Lag-Phase

Um bei der Berechnung der Intervalle eine zeitliche Verschiebung zu berücksichtigen, wird von allen zu analysierenden Kultivierungen zunächst ein Datensatz als zeitliche Referenz ausge-

wählt. Von dieser wird zu allen anderen Signalverläufen das Maximum der Kreuzkorrelation ermittelt. Die Daten des Signals werden vor weiteren Analyse zeitlich verschoben.

Zu Berücksichtigen ist aber, daß nur ein ausgeprägtes Maximum der Kreuzkorrelation ein eindeutiger Hinweis auf eine zeitliche Verschiebung des Signalverlaufs ist. Bei Größen, deren zeitlicher Verlauf nicht charakteristisch ist, ergibt die Kreuzkorrelation keinen Aufschluß über eine zeitliche Verschiebung der Kultivierungsdaten.

4.2.3 Mustererkennung zur Signalunterteilung

Bei einer Mustererkennung wird aus dem Rohdatenvektor der Prozeßvariablen ein Merkmalsvektor geringerer Dimension erzeugt, der sowohl Ähnlichkeiten als auch Unterschiede des Prozeßgeschehens verstärkt widerspiegelt. Ein Signal wird anhand des Merkmalsvektors in eine Abfolge von repräsentativen Mustern unterteilt. Dabei können zeitliche Abschnitte, in denen das gleiche Muster vorliegt, zu einer Signalphase zusammengefaßt werden. Die Mustererkennung, die in dem Expertensystem zur Anwendung kommt, sollte in der Off-line-Phase den Benutzer in der Analyse seiner Daten unterstützen und Hinweise auf den charakteristischen Prozeßverlauf geben. Durch den Benutzer erfolgt während der Wissensakquisition eine heuristische Klassifikation, in der das Auftreten charakteristischer Muster spezifischen Prozeßphänomenen zugeordnet wird. Das Ergebnis dieser Klassifikation sind Regeln, die in der Wissensbasis von dem Benutzer abgelegt werden und die in der On-line-Überwachung zur Erkennung von Prozeßzuständen und Diagnosesituationen genutzt werden können. Diese Vorgehensweise funktioniert nur zuverlässig, wenn sowohl on line als auch off line die gleichen Methoden zur Mustererkennung Verwendung finden. Die Mustererkennung sollte folgenden Ansprüchen genügen:

- Es sollten wenige Parameter zur Einstellung der Mustererkennung notwendig sein, die bestenfalls automatisch festgelegt werden können.
- Die Ergebnisse der Mustererkennung sollten in einer für den Benutzer verständlichen Form aufbereitet werden können.
- Es sollte kein Unterschied zwischen den Ergebnissen der On-line- und Off-line-Mustererkennung geben.

- Die Algorithmen sollten einfach und robust sein, damit im On-line-Betrieb immer eine genügend schnelle Antwortzeit gewährleistet bleibt.

In dem entwickelten Expertensystem kam eine Variante des Matched-Filter Ansatzes zum Einsatz. Diese Methode setzt kein Prozeßwissen voraus, ist einfach zu handhaben und eignet sich gut für den On-line-Einsatz, da nur wenige Parameter für die Einstellung der Mustererkennung vorhanden sind. Im folgenden soll die Methode vorgestellt werden.

Matched-Filter

Der Matched-Filter ist eine sehr einfache Methode zur Mustererkennung. In ihm wird aus einem Satz von Referenzmustern ein Muster bestimmt, das den Verlauf des Signals über ein festgelegtes Zeitintervall am besten wiedergibt. Hierfür wird der Abstand $A_i(t)$ zwischen jedem Referenzmuster R_i und dem Signal S berechnet. Das Referenzmuster mit dem kleinsten Abstand wird als zutreffendes Muster definiert. Wird in aufeinander folgenden Zeitabschnitten das gleiche Muster bestimmt, so werden diese zu einem Signalabschnitt zusammengefaßt. Abbildung 4-13 zeigt ein Signal sowie die für diskrete Zeitabschnitte gefundenen Muster. Hierbei symbolisieren die Buchstaben A-D verschiedene Muster. In dem Beispiel wird das Signal in fünf Signalabschnitte aufgeteilt, wobei der Anfangs- und Endzeitpunkt des Abschnitts sowie das Muster die jeweiligen Signalabschnitte kennzeichnen.

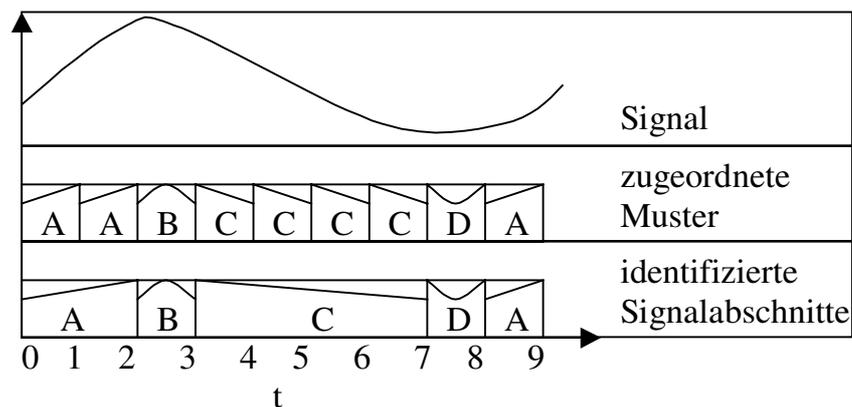


Abbildung 4-13 Mustererkennung und Unterteilung eines Signals in charakteristische Abschnitte

Soll die Mustererkennung on line eingesetzt werden, so muß berücksichtigt werden, daß zu einem gegebenen Zeitpunkt nur die aktuellen und alle vergangenen Daten vorhanden sind. Des Weiteren soll aber möglichst zu jedem Zeitpunkt eine Aussage über ein aktuelles Muster gemacht werden. Dies wird durch die gleitende Fenster-Methode erreicht, in der zu einem gege-

benen Zeitpunkt t_1 und einem Signal $S(k\Delta t)$, welches eine diskrete Funktion der Zeit darstellt, das Muster für das Zeitintervall $S(t_1 - t_{\text{Fenster}})$ bis $S(t_1)$ bestimmt und als zutreffendes Muster definiert wird. Hierbei gibt t_{Fenster} die Zeitdauer des betrachteten Datenbereichs an. t_{Fenster} stellt den wesentlichen Parameter der Mustererkennung dar und muß zu Beginn der Mustererkennung festgelegt werden. Abbildung 4-14 zeigt das Prinzip des gleitenden Fensters zur Festlegung des Datenbereichs. Dieses Fenster wird nach Ermittlung des Musters jeweils um einen Datenpunkt verschoben und der Vorgang der Mustererkennung wiederholt sich.

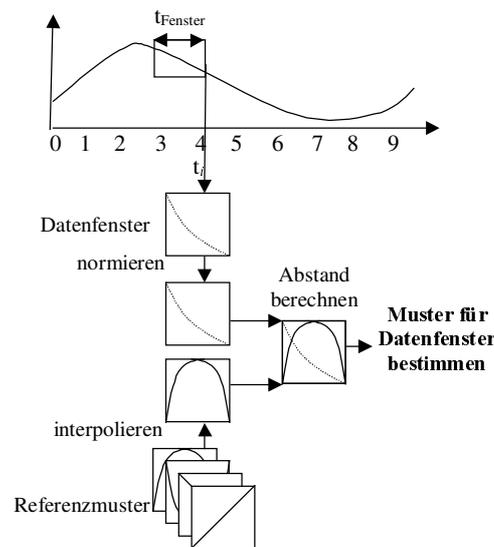


Abbildung 4-14 Prinzip der Mustererkennung mit gleitendem Datenfenster

Man erhält also für N Daten zunächst $N - N_{\text{Fenster}}$ Muster wobei N_{Fenster} die Anzahl der gemessenen Daten in dem Zeitintervall t_{Fenster} darstellt. Diese Anzahl hängt direkt mit der Meßfrequenz $1/\Delta t$ des Signals ab. Es gilt $N_{\text{Fenster}} = t_{\text{Fenster}} / \Delta t$. Eine Datenreduktion kommt erst bei der Zusammenfassung gleicher Muster zu Signalabschnitten zustande. Um für eine gegebenes Fenster das zugehörige Muster zu ermitteln, werden die Zeit und die Amplitude der Daten zunächst normiert, damit die so gewonnenen Daten invariant gegenüber Änderungen in Offset und Verstärkung sind. Das normierte Signal S_n erhält man nach Formel 4-2:

$$S(t)_n = \frac{S(t) - S_{\min}}{S_{\max} - S_{\min}} \quad \text{mit} \quad S_{\max} : \text{Maximum} \quad S_{\min} : \text{Minimum} \quad \text{Formel 4-2}$$

Die zeitliche Normierung nach Formel 4-3:

$$t_n = \frac{t - t_{\min}}{t_{\max} - t_{\min}} \quad \text{mit} \quad t_{\max} : \text{Maximum} \quad t_{\min} : \text{Minimum} \quad \text{Formel 4-3}$$

Man erhält eine normierte Meßfrequenz $1/\Delta t_n$ bei der gilt, $\Delta t_n = 1/N_{\text{Fenster}}$. Die Referenzmuster liegen in dem System als diskrete normierte Werte R_i vor. Folgende Kriterien wurden bei der Auswahl der Referenzmuster berücksichtigt:

- Die Mustererkennung sollte ohne Veränderung der Referenzmuster bei verschiedenen Prozeßgrößen einsetzbar sein.
- Des Weiteren sollten die Mustererkennung bei unterschiedlichen Kultivierungen vergleichbare Ergebnisse liefern.
- Unterschiede in den Daten, wie z. B. abweichende Meßfrequenzen sollten die Mustererkennung nicht beeinflussen.

Abbildung 4-15 zeigt die in dem System untersuchten und die für den praktischen Einsatz ausgewählten Referenzmuster. Es hat sich herausgestellt, daß es nur sinnvoll ist, wenige, stark unterschiedliche Referenzmuster zu verwenden. Werden zu viele oder sich stark ähnelnde Referenzmuster verwendet, so wird das Signal in sehr viele kleine Abschnitte unterteilt. Dies entspricht aber nicht dem Vorgehen des menschlichen Experten, der ein Signal in wenige Abschnitte unterteilt und nur die wesentlichen Charakteristika berücksichtigt.

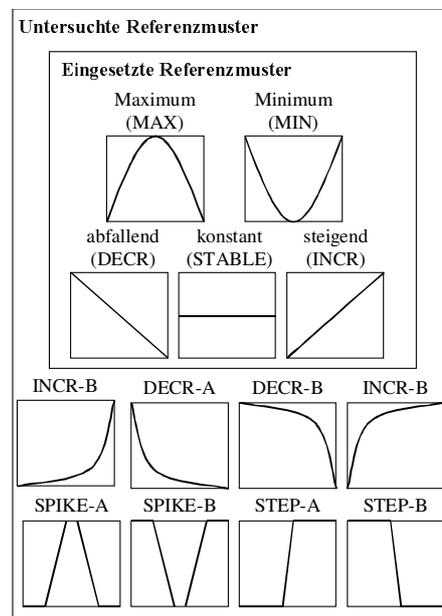


Abbildung 4-15 Darstellung aller Referenzmuster die in der Mustererkennung untersucht wurden, sowie die Auswahl der Muster, die in dem Expertensystem eingesetzt werden; über jedem Referenzmuster ist das Symbol gezeigt, mit dem sie im Expertensystem repräsentiert wird

Um das zutreffende Muster des Datenfensters zu ermitteln, wird der Abstand A_i zwischen den Referenzmustern und dem Datenfenster ermittelt. Die Berechnung des Abstands A_i erfolgt

nach Formel 4-4. Sie ergibt sich aus der Summe der quadratischen Abweichungen zwischen den Daten des Signals und einem Referenzmuster R_i . Damit der Abstand A_i berechnet werden kann, muß zu jedem Zeitpunkt des Signals $S_n(j\Delta t_n)$ auch das Referenzmuster R_i einen Funktionswert besitzen.

$$A_i = \sum_{j=1}^{j=N_{\text{Fenster}}} (R_i(j\Delta t_n) - S_n(j\Delta t_n))^2 \tag{Formel 4-4}$$

Da aber die Anzahl der Werte der Referenzmuster kleiner ist als die Anzahl der Daten in dem betrachteten Fenster, werden vor der Berechnung des Abstands zunächst fehlende Werte durch lineare Interpolation nach Formel 4-5 berechnet:

$$\left[\begin{array}{l} R_i(t_j) = \frac{(m+1)\Delta t_r - t_j}{\Delta t_r} R_i(m\Delta t_r) + \frac{t_j - m\Delta t_r}{\Delta t_r} R_i((m+1)\Delta t_r) \\ \text{mit } t_j = j\Delta t_n \\ \text{und } m < M-1 \\ \text{und } m\Delta t_r < t_j < (m+1)\Delta t_r \\ \text{und } \Delta t_r = 1/M \end{array} \right] \tag{Formel 4-5}$$

In der Abbildung 4-16 ist das Ergebnis der Mustererkennung für den CO₂-Wert einer Kultivierungen von *Acremonium chrysogenum* dargestellt. Für die Mustererkennung hat sich eine Breite des Datenfensters von 1-2 Stunden bewährt. In diesem konkreten Fall wurde ein Datenfenster von zwei Stunden festgesetzt. Es fand keine weitere Datenvorverarbeitung statt.

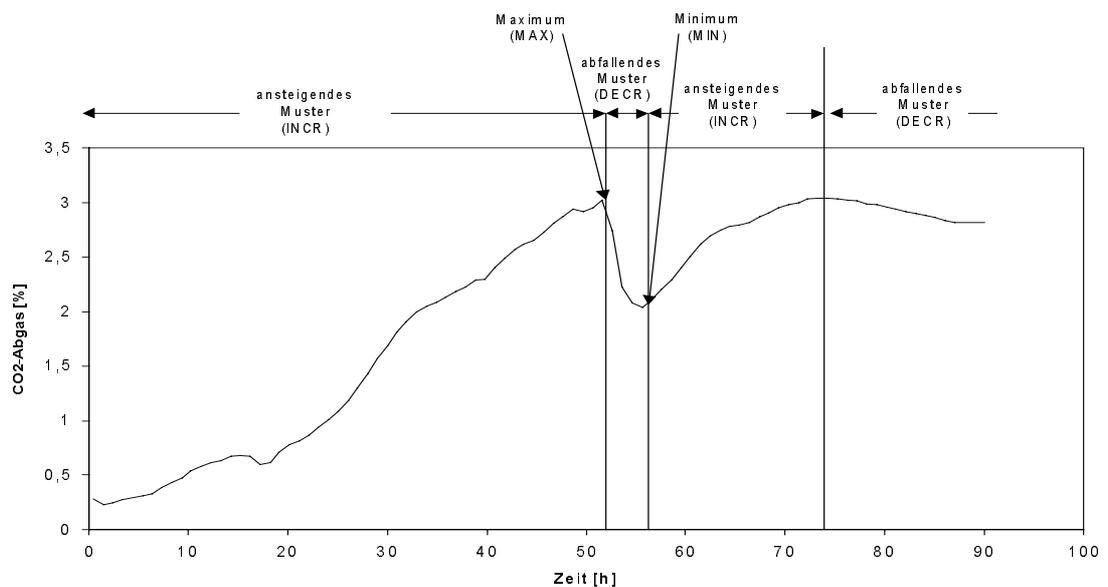


Abbildung 4-16 Ergebnisse der Mustererkennung für den CO₂-Wert einer Kultivierung von *Acremonium chrysogenum*

Das Signal wird in vier Abschnitte unterteilt, wobei die Erkennung von Minima und Maxima nicht als Abschnitte betrachtet werden, sondern als zusätzliche Muster, die Abschnittswechsel charakterisieren. Sichtbar ist die zeitlich verspätete Erkennung einzelner Muster, da zu einem gegebenen Zeitpunkt ein Muster passend zu den Daten der letzten zwei Stunden ermittelt wird. Durch die Verkleinerung des Datenfensters können Muster mit kleinerer zeitlicher Verzögerung erkannt werden. Hierbei treten aber Einflüsse des Signalrauschens stärker in den Vordergrund, so daß die spezifischen Eigenschaften des Signalverlaufs nicht mehr hervorgehoben werden.

4.3 Konzeption des Expertensystems

Durch die Entwicklung eines Expertensystems zur On-line-Überwachung von Bioprocessen soll der unerfahrene Anwender bei der Durchführung von Kultivierungen unterstützt werden. Ein Schwerpunkt liegt in der Programmierung eines graphisch orientierten Interfaces, da die Akzeptanz eines Expertensystems entscheidend von der Bedienerfreundlichkeit abhängig ist.

Die in Kapitel 2.7 beschriebenen Expertensysteme demonstrieren das Potential unterschiedlicher Methoden, zur Lösung spezieller Probleme der Prozeßführung und -überwachung. Die Übertragung der Methodik dieser Expertensysteme auf einen neuen, leicht veränderten Prozeß wirft aber immer zahlreiche Fragen auf, die häufig schwer von einem Experten des Fachbereichs zu beantworten sind. Neben dem Wissen aus dem Fachbereich des Experten wird noch sehr viel zusätzliches Wissen vorausgesetzt. So ist z. B. die Auswahl, Konfiguration und das Trainieren von neuronalen Netzen kein triviales Problem, deren Lösung das Wissen eines eigenständigen Experten erfordert. Je komplexer die Methodik wird, desto größer sind auch die Anforderungen an die Wissenserhebungssysteme. Dies gilt insbesondere dann, wenn das Expertensystem flexibel und schnell auf andere Prozesse übertragbar sein soll.

In dieser Arbeit soll zunächst ein hybrides System zur Überwachung der Bioprocessen eingesetzt werden. In diesem System werden numerische Methoden verwendet, um aus den on line gemessenen Daten zunächst numerische Merkmale zu berechnen. Anhand dieser Merkmale soll in einem nachgeschalteten wissensbasiertem Modul mit Regeln der Prozeßzustand überwacht und Störfälle diagnostiziert werden. Die Regeln sollten möglichst allgemein formuliert werden,

damit eine Allgemeingültigkeit der Wissensbasis gewährleistet bleibt. Dies kann z. B. durch eine strikte Trennung des Methodenwissens von dem prozeßspezifischen Faktenwissen erreicht werden.

Bei der Durchführung der Diagnose sollte eine Fokussierung der Inferenz auf relevante Regeln stattfinden, damit auch bei einer umfangreichen Wissensbasis ein Echtzeitverhalten gewährleistet bleibt.

Für die verwendeten Methoden sollen dem Bediener umfassende Werkzeuge zur Verfügung gestellt werden, damit der Fachexperte direkt die Wissensbasis manipulieren und das Expertensystem für die Überwachung unterschiedlicher Prozesse eingesetzt werden kann.

4.4 Systembeschreibung

In diesem Kapitel wird die Struktur und Funktionalität des Expertensystems beschrieben. Zunächst wird dazu die Gesamtarchitektur des Systems vorgestellt und darauf folgend, die wichtigsten Module des Systems näher beschrieben.

4.4.1 Architektur des Gesamtsystems

Die grundlegende Architektur des Systems mit der Anbindung an den Prozeßleitsystem über das Lokale Netzwerk (LAN) ist in Abbildung 4-17 dargestellt.

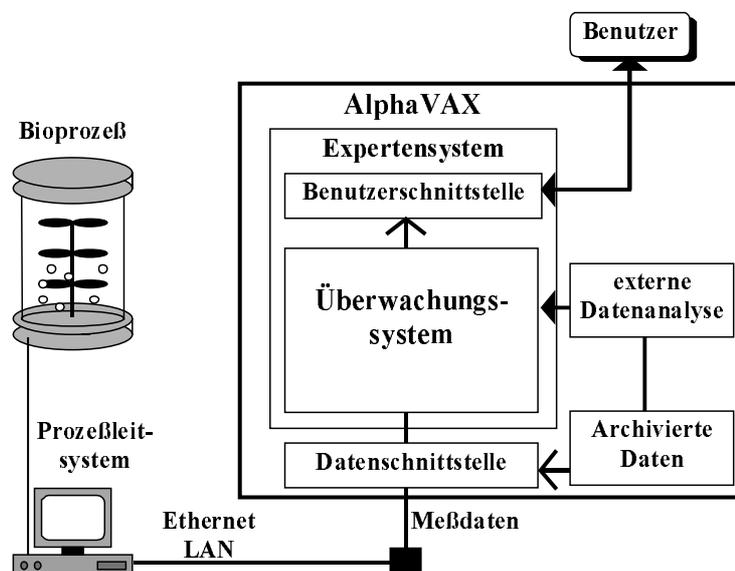


Abbildung 4-17 Architektur des Gesamtsystems und Anbindung des Prozeßleitsystems des Bioprocesses über das lokale Netzwerk (LAN)

Das System ist modular aufgebaut und enthält neben dem eigentlichen Überwachungssystem zwei Module, die Schnittstellenfunktionen zum Prozeßleitsystem und zum Anwender beinhalten. Ein weiteres Modul zur externen Datenanalyse führt aufwendige Off-line-Analysen mit archivierten Daten durch. Die verschiedenen Systemkomponenten werden im folgenden beschrieben.

4.4.2 Datenschnittstelle

Die Datenschnittstelle hat die Aufgabe, die vom Prozeßleitsystem aufgenommenen Meßdaten für das Expertensystem verfügbar zu machen. Hierzu werden die Daten über das Netzwerk auf den Rechner, auf dem das Expertensystem installiert ist, übertragen. Die Schnittstelle sollte schnelle Antwortzeiten bieten, damit die Echtzeitfähigkeit des Systems gewährleistet bleibt. Des Weiteren soll sie robust sein, damit im Falle von Netzwerkproblemen oder Systemausfällen eine Wiederaufnahme der Überwachung möglich ist. Diese Anforderungen werden durch eine Schnittstelle erreicht, dessen Aufbau in Abbildung 4-18 schematisch dargestellt ist.

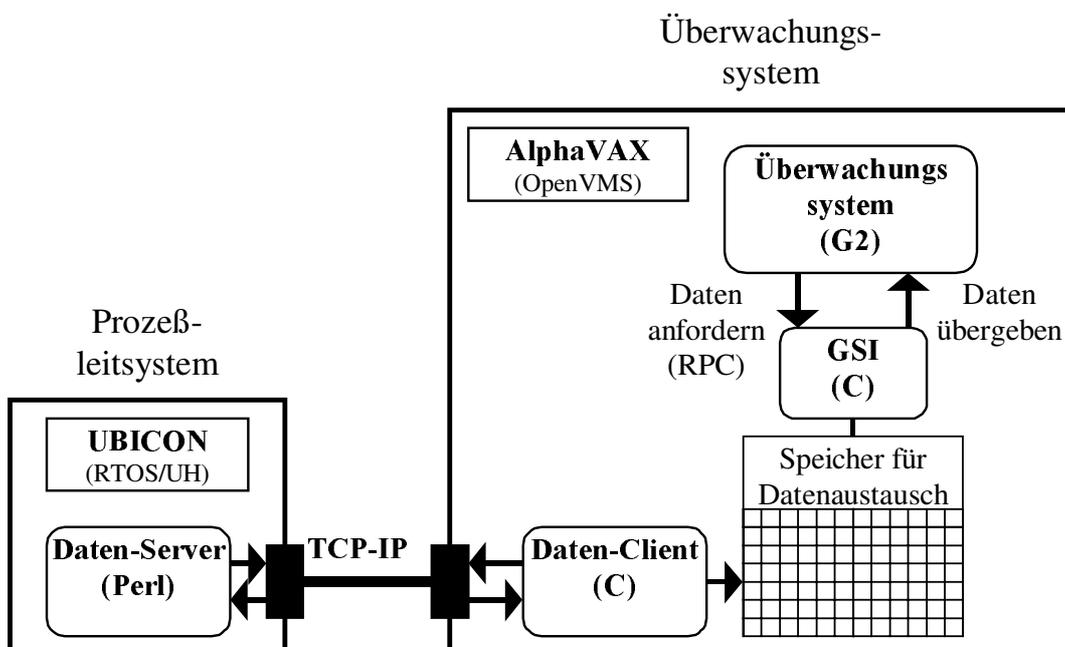


Abbildung 4-18 Schnittstelle zwischen dem Prozeßleitsystem und dem Expertensystem

Die Kommunikation zwischen dem Prozeßleitsystem UBICON und der Überwachungskomponente wurde mit Hilfe des TCP-IP-Socket Protokolls und eines Client-Server Konzepts realisiert [Seeger 1997]. Auf dem Prozeßleitsystem wurde ein Datenserver installiert, der auf Anfrage eines Clients Daten aus dem Prozeßleitdatenpool liest. Die Weiterleitung erfolgt asynchron, d.h. die Meßdaten werden in einem Speicherbereich (shared common) zwischengepuffert, falls das Überwachungssystem nicht sofort zur Aufnahme bereit ist. Das Eintreffen eines neuen Meßwerts wird dem Überwachungssystem durch das Setzen eines sogenannten Flags mitgeteilt.

Zum Einlesen der während der Kultivierung gemessenen Daten in das Expertensystem werden vom Überwachungssystem Schnittstellenfunktionen aufgerufen (remote procedure calls, RPC), welche die angeforderten Daten aus dem ‚shared common‘ in das Expertensystem einlesen. Die Schnittstellenfunktionen sind unter Verwendung des Gensym Standard Interface (GSI), welche die Standard Schnittstellenfunktionen der Expertensystemshell darstellen, implementiert.

Die Zwischenpufferung der Daten in einem shared common hat entscheidende Vorteile. Der auf dem Überwachungsrechner installierte Datenclient ist robust programmiert, so daß eine hohe Verfügbarkeit garantiert ist. Im Falle eines Systemfehlers des Expertensystems, werden weiterhin die Daten regelmäßig automatisch auf den Überwachungsrechner übertragen. Das Expertensystem kann nachdem es die Überwachung wieder aufgenommen hat, innerhalb kürzester Zeit den aktuellen Überwachungszustand wiedererlangen. Darüber hinaus ermöglicht die Datenschnittstelle auch den Off-line-Betrieb des Expertensystems mit archivierten Daten. Dazu werden die Dateien mit den Meßdaten von der Datenschnittstelle in den Speicher gelesen.

Ein weiterer Vorteil dieses Speichermodells ist die Zugriffsmöglichkeit anderer Programme auf die Daten, wie z. B. ein eigenständiges Visualisierungsprogramm. Somit dient der Überwachungsrechner als zentraler Datenserver.

4.4.3 Externe Datenanalyse

Der Ablauf der externen Datenanalyse ist in Abbildung 4-19 dargestellt. Die Datenanalyse historischer Daten dient zur Ermittlung von prozeßspezifischem Faktenwissen eines neuen Prozesses. Die Datenanalyse in dem Expertensystem gliedert sich in vier Schritte, die in Abbildung 4-19 dargestellt sind. Im ersten Schritt werden Informationen über durchgeführte Kultivierungen und den dazu vorliegenden Dateien vom Benutzer erfragt und in Objekten der Klasse

'Datensatz' abgelegt. Im zweiten Schritt werden gezielt Analysen der Daten vom Benutzer angefordert. Die Datenhaltung sowie die Durchführung und Speicherung der Analysen werden außerhalb des Expertensystems durchgeführt, was wesentliche Vorteile bietet. Zum einen müssen die Dateien nicht in das Expertensystem eingelesen werden, zum anderen bieten Programme, die in einer herkömmlichen Programmiersprache (wie C) geschrieben sind, eine wesentlich höhere Berechnungsgeschwindigkeit als die Prozeduren innerhalb von G2. Die Ergebnisse der Analyse werden in Dateien gespeichert, die in das Expertensystem eingelesen und zur Erstellung von Objekten der Klasse 'Intervalle', 'Signalverlauf' und 'Phasenverlauf' verwendet werden.

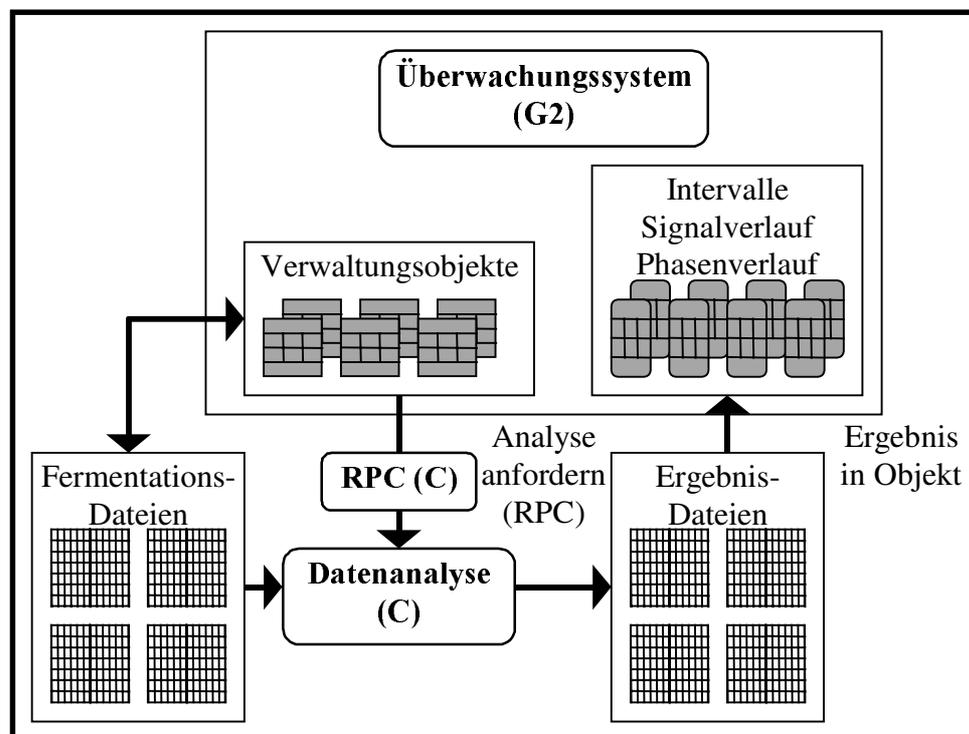


Abbildung 4-19 Ablauf der externen Datenanalyse

4.4.4 Überwachungssystem

Das Überwachungssystem vereint herkömmliche Techniken der numerischen Meßdatenverarbeitung mit regelorientierter (symbolischer) Wissensverarbeitung. Die effiziente Integration von numerischer und symbolischer Datenverarbeitung ist eine Grundvoraussetzung für die Entwicklung eines wissensbasierten Echtzeitsystems [Laffey 1988]. G2 bietet die Möglichkeit herkömmliche numerische Datenverarbeitung innerhalb von Prozeduren mit symboli-

scher Wissensverarbeitung zu kombinieren. Im folgenden wird dargestellt wie das Wissen, welches für die Überwachung benötigt wird, in dem Expertensystem implementiert ist.

4.4.4.1 Wissensrepräsentation

Tabelle 4-3 zeigt unterschiedliche Typen von Wissen und wie sie in dem System implementiert sind.

Tabelle 4-3 Wissensrepräsentation in dem Expertensystem

Wissensrepräsentation	Wissensart	Erläuterung
Objekte	Faktenwissen	Wissen über konkrete und abstrakte Objekte
Relationen	Faktenwissen	Beziehungen zwischen Objekten
Regeln	Methodenwissen	Wissen über kausale Zusammenhänge
Prozeduren	Explizites Wissen	Berechnung expliziter Algorithmen

4.4.4.1.1 Objekte

Ein Objekt stellt eine Instanz einer Objektklasse dar, wobei die Eigenschaften (Attribute) der Klasse in der Objektdefinition festgelegt werden. Um Zusammenhänge und Ähnlichkeiten zwischen Objektklassen zum Ausdruck zu bringen, können Objekthierarchien durch Vererbung aufgebaut werden. Bei der Festlegung von Objektklassen beinhaltet eine Klasse, die höher in der Klassenhierarchie steht, alle gemeinsamen Eigenschaften der zugehörigen Unterklassen. Je höher eine Klasse innerhalb der Klassenhierarchie angeordnet ist, desto allgemeiner sind demnach ihre Eigenschaften.

Vornehmliche Aufgabe des Expertensystems ist die Überwachung von Bioprozessen. Um eine, dem Menschen analoge Vorgehensweise bei der Überwachung von Bioprozessen mit dem Computer nachzuahmen, müssen Objekte und somit auch das Wissen zahlreicher Wissensdomäne in dem System abgebildet werden. Die für die Überwachung benötigten Objekte gliedern sich in die vier Oberklassen Prozeß, Meßwerte, Überwachung und Datenanalyse. Im folgenden soll auf diese Objektklassen und die daraus abgeleiteten Unterklassen näher eingegangen werden.

Objektklasse Prozeß

Unter der Klasse Prozeß sind alle Objekte definiert, welche konkrete Bestandteile des Bioprozesses repräsentieren. Die in Abbildung 4-20 aufgeführten Objektklassen stellen nur eine Auswahl, der in der Wissensbasis verfügbaren Objektklassen zur Repräsentation von Wissen über den Bioprozeß dar. Die Akquisition dieses Wissens und die Erstellung dieser Objekte findet in einem zentralen Wissensakquisitionsmoduls des Gesamtexpertensystems statt. Es sei an dieser Stelle auf die Arbeiten von Ludewig [1998] hingewiesen, welche die Funktionen und den Inhalt der zentralen Wissensakquisition beschreibt.

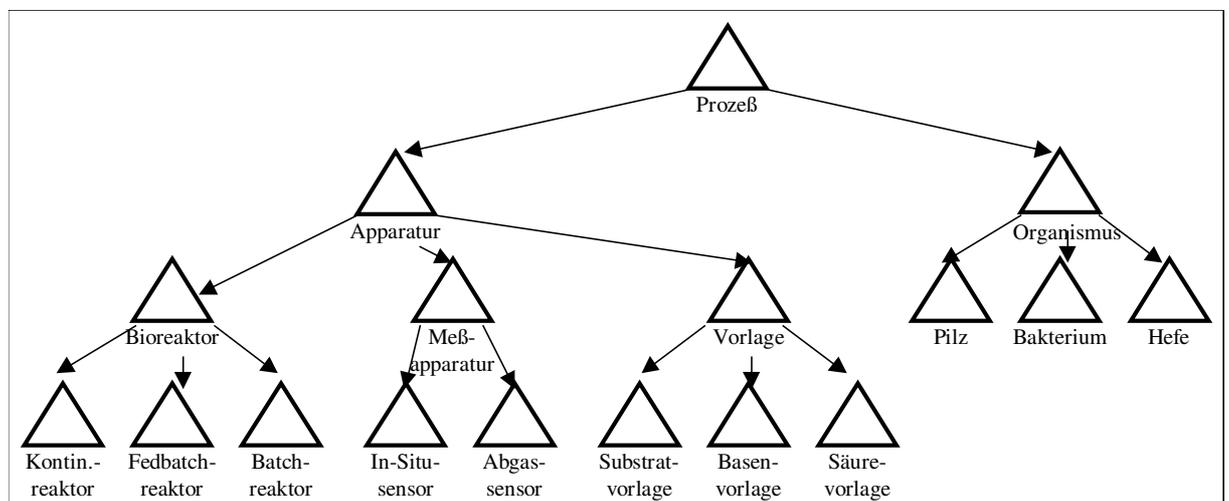


Abbildung 4-20 Die Objektklasse 'Prozeß'

Objektklasse Meßwerte

Damit das Expertensystem Aussagen über den Zustand des Bioprozesses und seiner konkreten Objekte machen kann, müssen dem System Informationen zur Beurteilung zur Verfügung gestellt werden. Die von dem Prozeßleitsystem on line gemessenen Prozeßvariablen wie z. B. der pH-Wert, die Temperatur oder der Kohlendioxidgehalt im Abgas sind für das Expertensystem die einzige Quelle an Information über den aktuellen Zustand des Prozesses. Alle Objekte, die diese Meßwerte und alle daraus resultierenden Eigenschaften beinhalten sind unter der Objektklasse 'Meßwerte' zusammengefaßt. In der Abbildung 4-21 sind die Objektklasse 'Meßwert' und alle daraus abgeleiteten Unterklassen dargestellt. Um spezielle Eigenschaften der unterschiedlichen Meßwerte zu erfassen, sind die Unterklassen 'On-line-', 'Off-line-' und 'Berechnete-Größen' daraus abgeleitet.

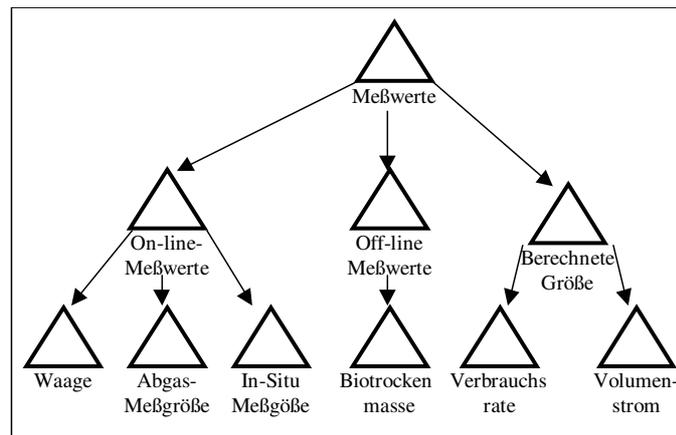


Abbildung 4-21 Die Objektklasse ‘Meßwerte’ und die zugehörigen Unterklassen

Die Eigenschaften der Objekte ‘Meßwerte’ gliedern sich in die vier logischen Eigenschaftsklassen Rohdaten, numerische Merkmale, symbolische Merkmale und Eigenschaften für die Verwaltung des Objekts, die in Tabelle 4-4 dargestellt ist. Unter Rohdaten sind alle zeitlichen Verläufe des Meßwerts zusammengefaßt. Hierzu gehören die Rohdaten, die über die Schnittstelle eingelesen wurden, sowie zahlreiche gefilterte Datenverläufe wie z. B. die Stundenmittelwerte oder der Medians.

Tabelle 4-4 Gliederung der Eigenschaften der Objekte ‘Meßwerte’

Attribute	Beschreibung
Daten	Über die Schnittstelle eingelesenen Daten
Numerische Merkmale	Mit Prozeduren berechnete Merkmale
Symbolische Merkmale	Symbole der qualitativen Einschätzung der numerischen Merkmale und des Zustands
Verwaltung	Von dem Steuersystem verwendeten Attribute

Unter numerische Merkmale sind die zeitlichen Verläufe der Merkmale der Meßwerte zusammengefaßt, die mittels numerischer Algorithmen berechnet wurden. Hierzu gehören z. B. die erste und zweite Ableitung des Signals, die Steigung des Signals über eine bzw. zwei Stunden oder das mittlere Rauschen des Signals. Den symbolischen Merkmalen des Meßwerts werden alle Eigenschaften zugeordnet, die eine qualitative Beurteilung des Meßwerts beinhalten. So existiert für jedes numerische Merkmal ein symbolisches Merkmal, das die Einschätzung des numerischen Merkmals charakterisiert. Unter Verwaltung werden Attribute zusammengefaßt,

die von dem Steuersystem des Expertensystems während des Ablaufs der Überwachung benötigt werden.

Objektklasse Überwachung

Die Objekte, die Ergebnisse und Zwischenergebnisse der Überwachung bereitstellen, sind unter der Klasse ‘Überwachung’ zusammengefaßt, die in der Abbildung 4-22 dargestellt ist. Von dieser Klasse abgeleitet sind z. B. die Klassen ‘Symptom’, ‘Hypothese’, ‘Störfall’ und ‘Prozeßzustand’. Die Klasse ‘Störfall’ beinhaltet zusätzlich die Unterklassen ‘Biologischer-Störfall’, ‘Regelungsstörfall’ und ‘Sensorstörung’.

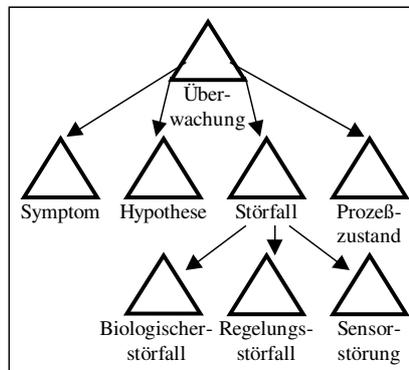


Abbildung 4-22 Die Objektklasse ‘Überwachung’

In der Tabelle 4-5 sind die Eigenschaften der Objekte ‘Störfall’ dargestellt. Neben einem Bezeichner für die Art des Störfalls, besitzt ein Objekt der Klasse Störfall eine detaillierte Beschreibung, sowie eine Liste von Symptomen, die den jeweiligen Störfall näher charakterisieren. Die Eigenschaft ‘Maßnahmen’ bietet Maßnahmen an, die zur Behebung des Störfalls durchgeführt werden können. Die letzten beiden Eigenschaften werden zur Speicherung des Zeitpunkts verwendet an dem ein Störfall während einer Kultivierung auftrat.

Tabelle 4-5 Eigenschaftstabelle des Objekts ‘Störfall’

Attribute	Beschreibung
Störfall-Art	Art der Störung
Beschreibung	Beschreibung des Störfalls
Symptome	Symptome, die den Störfall charakterisieren
Maßnahmen	Maßnahmen, um Störfall zu beseitigen
Frühste-Erkennung	Frühster Zeitpunkt, an dem der Störfall auftrat
Späteste-Erkennung	Spätester Zeitpunkt, an dem der Störfall auftrat

Objektklasse Datenanalyse

Das für die Überwachung von Bioprocessen notwendige Expertenwissen erlangt der Bediener in der Regel durch die Auswertung von Erfahrungen, die mit früheren Kultivierungen gesammelt wurden. Dadurch wird es ihm möglich, bestimmte Werte und Verläufe in den Meßsignalen qualitativ einzuschätzen und ihren Ursachen zuzuordnen. Damit das Expertensystem eine qualitative Einschätzung der Meßwerte und ihrer Merkmale vornehmen kann, müssen vor der Überwachung historische Daten ausgewertet und prozeßspezifisches Faktenwissen für die Einschätzung generiert werden. Die Klasse Datenanalyse beinhaltet alle Objekte, die für die Datenanalyse historischer Daten und die Verfügbarmachung des Faktenwissens während der Überwachung benötigt werden. In der Abbildung 4-23 sind die Objektklassen, die sich von der Klasse ‚Datenanalyse‘ ableiten, dargestellt.

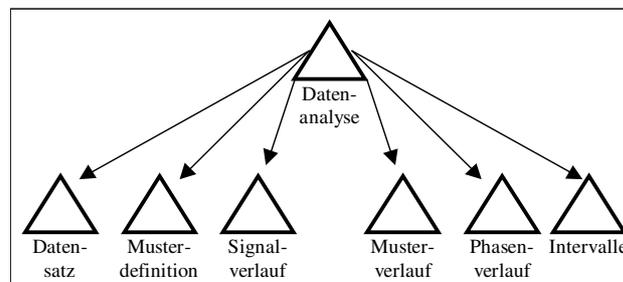


Abbildung 4-23 Die Objektklasse 'Datenanalyse'

Hierzu gehören die Objektklassen 'Datensatz', 'Musterdefinition', 'Signalverlauf', 'Musterverlauf', 'Phasenverlauf' und 'Intervalle'. Die Klasse 'Datensatz' findet für die Verwaltung und Verarbeitung von Off-line-Daten Verwendung. In der Tabelle 4-6 sind die wesentlichen Eigenschaften des Objekts 'Datensatz' dargestellt. Die Ergebnisse einer Datenanalyse werden in Objekten der Klasse 'Signalverlauf', 'Musterverlauf', 'Phasenverlauf' und 'Intervalle' gespeichert.

Tabelle 4-7 zeigt die Eigenschaftstabelle des Objekts 'Signalverlauf', in einem solchen Objekt wird die Information über den Signalverlauf eines Meßwerts gespeichert. Hierzu gehören z. B. in der Datenanalyse berechneten Angaben über die Anzahl der charakteristischen Signalabschnitte oder die maximal zu erwartende Lag-Phase. In einem Objekt der Klasse 'Musterverlauf' wird das Ergebnis der Off-line-Mustererkennung gespeichert. Tabelle 4-8 stellt die Eigenschaftstabelle des Objekts 'Musterverlauf' dar. Die Analyse des Phasenverlaufs der Kultivierung wird in einem Objekt der Klasse 'Phasenverlauf' gespeichert, deren Eigenschaftstabelle in

der Tabelle 4-9 aufgelistet ist. Um on line den Verlauf der numerischen Merkmale qualitativ einschätzen zu können, werden in der Datenanalyse Intervalle berechnet, die in dem Expertensystem in Objekten der Klasse 'Intervalle' abgelegt sind. Tabelle 4-10 zeigt die Eigenschaftstabelle dieser Objektklasse.

Tabelle 4-6 Eigenschaftstabelle des Objekts ‚Datensatz‘

Attribute	Typ	Beschreibung
Pfad	Text	Pfad des Datensatzes
DateiName	Text	Dateiname
Kultivierungsdatum	Text	Zeitpunkt der Kultivierung
Notizen	Text	Notizen über den Ziel der Kultivierung, Medium etc.
Störungen	Text	Aufgetretene Störungen
AnzahlDaten	Integer	Anzahl der aufgenommenen Daten (Zeilen)
AnzahlMeßwerte	Integer	Anzahl der gemessenen Werte (Spalten)
MeßwertName	Text-Feld	Bezeichnung der einzelnen Meßwerte
MeßwertNameUBICON	Text-Feld	Bezeichnung der Meßwerte im Prozeßleitsystem
Typ	Text-Feld	Typ des Meßwerts (Off-line, Waage, Feed-Rate)
AnzahlAnalysen	Integer	Anzahl der durchgeführten Datenanalysen
AnalyseObjekte	Text-Feld	Namen der erstellten Ergebnisobjekte

Tabelle 4-7 Eigenschaftstabelle des Objekts ‚Signalverlauf‘

Attribute	Typ	Beschreibung
DateiNamen	Text-Feld	Dateinamen der analysierten Dateien
MeßwertName	Text	Name des analysierten Meßwerts
AnzahlAbschnitte	Integer	Anzahl der charakteristischen Signalabschnitte
MaxLagPhase	Float	Maximale ermittelte Lag-Phase
AnzahlMax	Integer	Anzahl der lokalen Maxima im Signalverlauf
MaxWerte	Float-Feld	Absoluter Wert der lokalen Maxima
ZeitMaxWerte	Float-Feld	Zeitpunkt des Auftretens lokaler Maxima
AnzahlMin	Integer	Anzahl der lokalen Minima im Signalverlauf
MinWerte	Float-Feld	Absoluter Wert der lokalen Minima
ZeitMinWerte	Float-Feld	Zeitpunkt des Auftretens lokaler Minima
MaxTrend	Float	Maximale Steigung
ZeitMaxTrend	Float	Zeitpunkt der maximalen Steigung
MinTrend	Float	Mainimale Steigung
ZeitMinTrend	Float	Zeitpunkt der minimalen Steigung
Wertwiederholungen	Integer	Maximale Anzahl der Wertwiederholungen
RegelungsBeginMin	Float	Frühster Zeitpunkt des Einsetzen einer Regelungen
RegelungsBeginMax	Float	Spätester Zeitpunkt des Einsetzen einer Regelungen
SignalRauschen	Float	Mittlerer Rauschfaktor
SignalRauschenVarianz	Float	Varianz des Rauschenfaktors

Tabelle 4-8 Eigenschaftstabelle des Objekts ‚Musterverlauf‘

Attribute	Typ	Beschreibung
AnzahlMuster	Integer	Anzahl der gefundenen Muster
ZeitAnfang	Float-Feld	Anfang des Musters
ZeitEnde	Float-Feld	Ende des Musters
Muster	Symbol-Feld	Für den Zeitabschnitt ermittelte Muster
WertAnfang	Float-Feld	Wert am Anfang der Signalabschnitts
WertEnde	Float-Feld	Wert am Ende der Signalabschnitts

Tabelle 4-9 Eigenschaftstabelle des Objekts ‚Phasenverlauf‘

Attribute	Typ	Beschreibung
AnzahlPhasen	Integer	Anzahl der ermittelten Phasen
ZeitAnfang	Float-Feld	Anfang der Phase
ZeitEnde	Float-Feld	Ende der Phase
Anfangsmerkmal	Symbol-Feld	Merkmal, das die Phase einleitet
Endmerkmal	Symbol-Feld	Merkmal, das die Phase beendet

Tabelle 4-10 Eigenschaftstabelle des Objekts ‚Intervalle‘

Attribute	Typ	Beschreibung
Zeit	Float-Feld	Zeit des Intervallabschnitts
DurchschnittWert	Float-Feld	Durchschnittswert des Intervallabschnitts
StandardabweichungWert	Float-Feld	Standardabweichung des Intervallabschnitts
MaxWert	Float-Feld	Maximaler Wert des Intervallabschnitts
MinWert	Float-Feld	Minimaler Wert des Intervallabschnitts
MaxTrend1	Float-Feld	Maximale Steigung des Intervallabschnitts 1
MinTrend1	Float-Feld	Minimale Steigung des Intervallabschnitts 1
MaxTrend2	Float-Feld	Maximale Steigung des Intervallabschnitts 2
MinTrend2	Float-Feld	Minimale Steigung des Intervallabschnitts 2
MaxTrend3	Float-Feld	Maximale Steigung des Intervallabschnitts 3
MinTrend3	Float-Feld	Minimale Steigung des Intervallabschnitts 3

4.4.4.1.2 Relationen

Mit der Hilfe von Relationen werden innerhalb von dem Expertensystem Beziehungen zwischen unterschiedlichen Objekten zum Ausdruck gebracht. So bringt die Relation ‘ist-Meßwert-von’ die Beziehung, daß ein Objekt den Meßwert einer bestimmten Meßapparatur darstellt. In der Tabelle 4-11 sind die wichtigsten Relationen für die Überwachung sowie der korrespondierenden Klassen aufgelistet.

Tabelle 4-11 Relationen des Expertensystems und ihre Bedeutung

Name der Relation	Klasse 1	Klasse 2
ist-Meßwert-von	Meßwert	Meßapparatur
ist-Symptom-von	Symptom	Meßwert
ist-Hypothese-von	Hypothese	Meßwert
ist-Störfall-von	Störfall	Objekt

4.4.4.1.3 Regeln

Mit Regeln wird Methodenwissen repräsentiert und Wissen über kausale Zusammenhänge dargestellt. Eine Regel besteht aus einer Prämisse und einer Konklusion. Sind die in der Prämisse gestellten Anforderungen erfüllt, wird der Aktionsteil der Regel ausgeführt. Die Schlußfolgerung kann eine Veränderung der Wissensbasis bewirken, so daß weitere Regeln auf die neue Situation reagieren die wiederum weitere Schlußfolgerungen nach sich ziehen. In der Expertensystemshell G2 besitzen Regeln zahlreiche Eigenschaften, mit deren Hilfe die Verarbeitung durch die Inferenzmaschine beeinflusst werden kann. In dem Expertensystem sind Regeln implementiert, um auf bestimmte Änderungen der Wissensbasis zu reagieren und auf Anforderung des Benutzers bestimmte Aktionen zu triggern. Diese Meta-Regeln repräsentieren im wesentlichen das Steuerungswissen und das Wissen zur Verwaltung des Expertensystems.

Das spezifische Wissen zur Überwachung und Diagnose des Bioprocesses wird mit zwei unterschiedlichen Regel-Typen implementiert: Die Low-level-Regeln und die Diagnose-Regeln. Low-level-Regeln stellen ständig aktive 'Whenever-Regeln' dar, die auf Änderungen bestimmter Eigenschaften der Überwachungsobjekte reagieren und nötigenfalls Diagnose-Regeln aktivieren, die für genau die Diagnose des unnormalen Zustands verantwortlich sind. Die Aktivierung der Diagnose-Regeln findet über die Attribute 'Categories' und 'Focal-Objects' – dies entspricht der Fokussierung auf ein bestimmtes Objekt – statt. Abbildung 4-24 zeigt die Eigenschaftstabelle einer Low-level-Regel. Sie stellt eine generische Regel dar, die immer von der Inferenzmaschine überprüft wird, wenn ein Objekt der Klasse 'Meßwertanalyse' einen neuen Wert erhält. Die Bedingungsteil der Regel ist erfüllt, wenn das Symbol der qualitativen Einschätzung des Meßwerts ‚HOCH‘ ist. Das Auftreten dieses Merkmals stellt ein Symptom dar, welches in einer Diagnose auf mögliche Ursachen weiter untersucht werden kann. Hierfür wird in der Konklusion der Regel über den Befehl 'invoke diag-wert-hoch rules for M1' alle Diagnose-Regeln aktiviert, deren Eigenschaftstabelle die Kategorie 'diag-wert-hoch' beinhaltet und deren Fokales-Objekt M1 entspricht.

RULE-269, a rule	
Options	not invocable via backward chaining, not invocable via forward chaining, may not cause data seeking, may not cause forward chaining
Notes	RULE-XXX-46: OK
User restrictions	none
Names	RULE-269
Tracing and breakpoints	default
whenever the wert-symbol of any messwert-analyse M1 receives a value and when the wert-symbol of M1 = the symbol HOCH then invoke diag-wert-hoch for M1	
Scan interval	none
Focal classes	none
Focal objects	none
Categories	none
Rule priority	6
Depth first backward chaining precedence	1
Timeout for rule completion	use default

Abbildung 4-24 Eigenschaftstabelle einer Low-level-Regel

Abbildung 4-25 zeigt die Eigenschaftstabelle einer Diagnose-Regel. In der Eigenschaftstabelle der Regel ist die Kategorie 'Diagnose-Max' und das Fokale-Objekt 'Meßwert-CO₂' eingetragen.

RULE-161, a rule	
Options	invocable via backward chaining, not invocable via forward chaining, may cause data seeking, may cause forward chaining
Notes	OK
User restrictions	none
Names	RULE-161
Tracing and breakpoints	default
if the sub-ws-active of the diagnose-regel-ws superior to the workspace of this rule = the symbol true then start message-auf-scroll(this rule,"CO2 Maximum entdeckt!","einfache Regel") and start deactivate-rule (the workspace of this rule ,2.0)	
Scan interval	none
Focal classes	none
Focal objects	messwert-co2
Categories	diagnose-max
Rule priority	6
Depth first backward chaining precedence	1
Timeout for rule completion	use default

Abbildung 4-25 Eigenschaftstabelle einer Diagnose-Regel

Eine Diagnose-Regel kann als Schlußfolgerung die Inferenz weiterer Diagnose-Regeln bewirken. Die gegenseitigen Abhängigkeiten der Regeln nach Kategorien, fokalen Objekten und deren Aktivierung stellen eine Baumstruktur dar. Eine Diagnose stellt dabei das Durchsuchen eines bestimmten Astes innerhalb dieses Diagnose-Baums dar und wird durch eine Low-level-Regel initiiert. In der Abbildung 4-26 reagiert die Low-level-Regel 1 auf das Auftreten eines

Symptoms des Objekts 1 in dem sie alle Regeln aktiviert, welche die Kategorie A und den Fokus auf dem Objekt 1 aufweisen. Die Diagnose-Regel 1 aktiviert im Verlauf der Inferenz weitere Diagnose-Regeln, welche die Kategorie E und den Fokus auf dem Objekt 1 besitzen. Dieser Vorgang wird solange weitergeführt bis Regeln ansprechen, die keine Aktivierungen durchführen oder die Prämisse keiner Regeln mehr erfüllt sind.

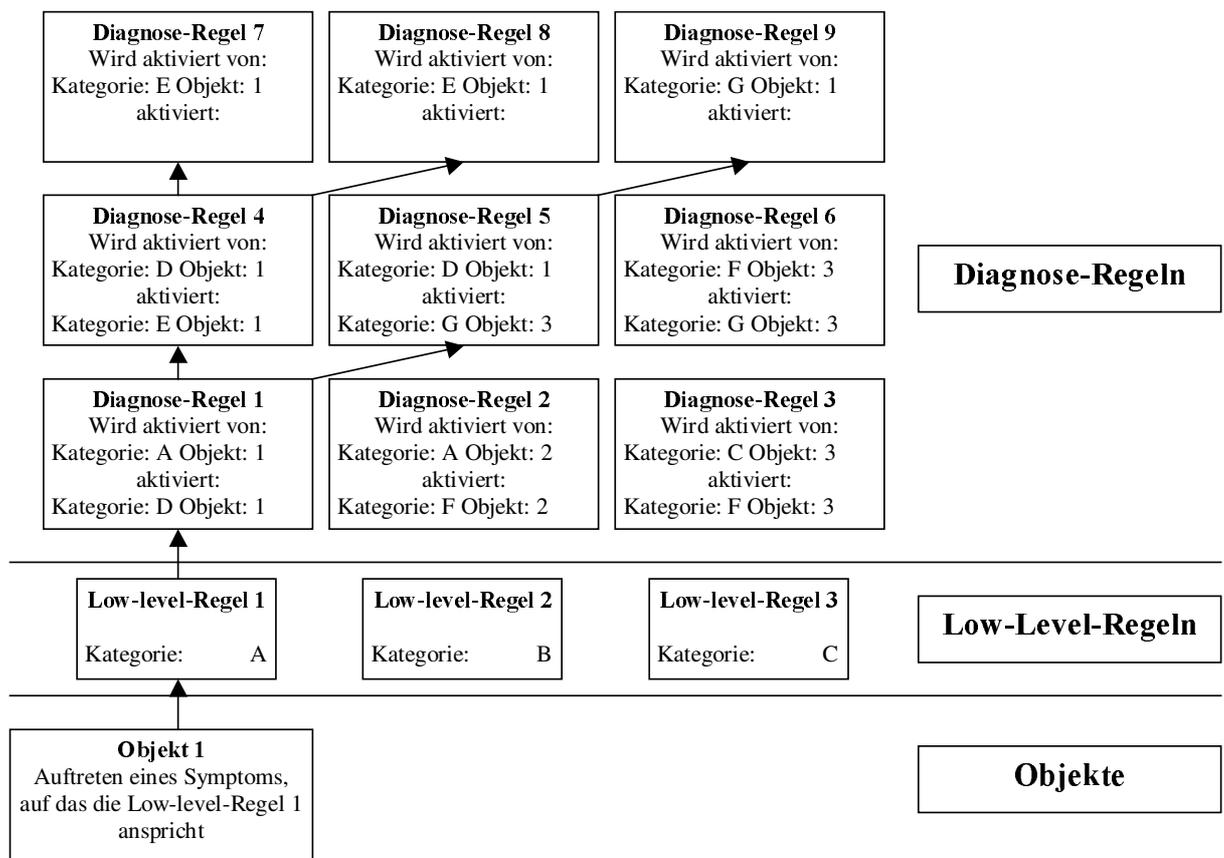


Abbildung 4-26 Schematische Darstellung des Inferenzvorgangs; Low-level-Regel 1 reagiert auf das Symptom des Objekts 1 indem es alle Diagnose-Regeln aktiviert, welche die Kategorie A und den Fokus auf dem Objekt 1 aufweisen.

In dem Expertensystem sind zahlreiche Low-level-Regeln implementiert. Tabelle 4-12 stellt die unterschiedlichen Kategorien dar, die von ihnen aktiviert werden.

Tabelle 4-12 Auflistung der Kategorien, die – von dem im System implementierten Low-level-Regeln – aktiviert werden

Aktivierte Kategorie	Attribute (Typ)	Focal-Objekt
Diagnose-Hypothese-bestaetigt	Bestaetigte-Hypothese	Aktueller-Prozeß
Diagnose-Hypothese-verworfen	Verworfen-Hypothese	Aktueller-Prozeß
Diag-Wert-Hoch	Wert (Symbol)	Alle Meßgrößen
Diag-Wert-niedrig	Wert (Symbol)	Alle Meßgrößen
Diag-Trend1-hoch	Meßwertabstand (Symbol)	Alle Meßgrößen
Diag-Trend1-niedrig	Meßwertabstand (Symbol)	Alle Meßgrößen
Diag-Trend2-hoch	Trend 1 Stunde (Symbol)	Alle Meßgrößen
Diag-Trend2-niedrig	Trend 1 Stunde (Symbol)	Alle Meßgrößen
Diag-Trend3-hoch	Trend 2 Stunden (Symbol)	Alle Meßgrößen
Diag-Trend3-niedrig	Trend 2 Stunden (Symbol)	Alle Meßgrößen
Diag-Wert-wiederholung	Meßwertabstand (Symbol)	Alle Meßgrößen
Diag-Spike	Meßwertabstand (Symbol)	Alle Meßgrößen
Diag-sr-hoch	Signalrauschen (Symbol)	Alle Meßgrößen
Diag-max	Muster (Symbol)	Alle Meßgrößen
Diag-min	Muster (Symbol)	Alle Meßgrößen
Diag-decrease	Muster (Symbol)	Alle Meßgrößen
Diag-increase	Muster (Symbol)	Alle Meßgrößen
Diag-stable	Muster (Symbol)	Alle Meßgrößen

4.4.4.1.4 Prozeduren

Algorithmisierbares Wissen, das z. B. in Form von mathematischen Gleichungen vorliegt, läßt sich innerhalb von G2 in Prozeduren ablegen. Abbildung 4-27 zeigt die Eigenschaftstabelle einer Prozedur in G2. Zu diesem Wissensbereich gehören beispielsweise Algorithmen zur Berechnung von Merkmalen, Formeln für statistische Berechnungen und Gleichungen, die mathematische Modelle repräsentieren.

NAME-EXIST, a procedure	
Notes	OK
User restrictions	none
Tracing and breakpoints	default
Class of procedure invocation	none
Default procedure priority	6
Uninterrupted procedure execution limit	use default
<pre> name-exist(name-to-check:text,logical:truth-value) M:class float-array ; begin logical = true; if there exists an float-array M such that (the text of the names of M = name-to-check) then logical = false; end </pre>	

Abbildung 4-27 Eigenschaftstabelle einer Prozedur in G2

Eine weitere wichtige Wissensart, die in Prozeduren abgelegt ist, stellt das Wissen des Steuersystems dar. Dieses Wissen faßt alle Schritte zusammen, die für den Programmablauf des Expertensystems verantwortlich sind, nicht aber explizit mit dem speziellen Wissen des Bioprozesses zu tun haben. Innerhalb der Prozeduren von G2 können die Attribute einzelner Objekte für Berechnung referenziert werden.

4.4.4.2 Arbeitsablauf der Überwachung

Die On-line-Daten werden einmal pro Minute eingelesen. Aus den eingelesenen Daten werden zunächst numerische Merkmale wie z. B. Trends und Varianzen berechnet. Das unscharfe Wissen, das in den Regeln enthalten ist, kann jedoch nicht direkt auf die exakten numerischen Merkmale zur Anwendung kommen. Die Repräsentation unscharfer Formulierungen findet mit Hilfe von Intervallen statt. So wird einem sprachlich formuliertem Wert (z. B. *hoch*) ein Zahlenintervall zugeordnet. Liegt der Meßwert eines Merkmals innerhalb dieses Intervalls, so ist eine Aussage, die sich auf den entsprechenden formulierten Wert bezieht, wahr. Jedem sprachlich formulierten Wert wird in der Datenbasis ein Symbol (Zeichen bzw. Zeichenfolge) zugeordnet. Nach der Transformation der numerischen Merkmale in unscharfe Formulierungen – repräsentiert durch Symbole – können die Regeln angewendet werden.

Die Vorgehensweise der Erklärungskomponente ist in Abbildung 4-28 dargestellt. Um einen kontinuierlichen Echtzeit-Betrieb zu gewährleisten, muß die Antwortzeit weniger als eine Minute (bei einer Meßfrequenz von 1 min^{-1}) in Anspruch nehmen. Wenn in jedem Zyklus alle Regeln überprüft werden müssen, so wird – bei wachsender Wissensbasis – diese Antwortzeit schnell überschritten. Es findet deshalb eine Fokussierung – der Aktivität – auf relevante Regeln statt. Dies geschieht durch die Unterteilung der Regeln in ständig aktive Low-level-Regeln und nicht aktive Diagnose-Regeln. Die Low-level-Regeln reagieren auf Änderungen symbolischer Merkmale der Signale und generieren Symptome. Um das Symptom einer Ursache zuzuordnen, d.h. eine Diagnose durchzuführen, werden weitere Regeln aktiviert. Diese Regeln erzeugen Hypothesen über mögliche Ursachen. Im weiteren Verlauf der Schlußfolgerungen können die Hypothesen bestätigt oder verworfen werden.

Die Erstellung von Hypothesen ist besonders wichtig, wenn in einem Inferenzzyklus keine eindeutige Diagnose durchgeführt werden kann. Sie stellen damit Zwischenergebnisse der Dia-

gnose dar, auf die in späteren Inferenzzyklen zurückgegriffen werden kann. Der Einsatz von Hypothesen bietet unter anderem sich unter folgenden Voraussetzungen an:

- Wenn die Symptome zur Erkennung des Zustands nur kurzfristig – im schlimmsten Fall nur während eines Inferenzzyklusses – in Erscheinung treten. Damit kann nicht gewährleistet werden, daß alle Symptome, die zur Durchführung einer Diagnose notwendig sind, zur gleichen Zeit auftreten. In diesem Fall wird die Hypothese als Zwischenergebnis verwendet, auf das in einem späteren Inferenzzyklus Bezug genommen werden kann.
- Über den Zeitpunkt der Erstellung der Hypothese auf das zeitliche Auftreten von Symptomen geschlossen werden.

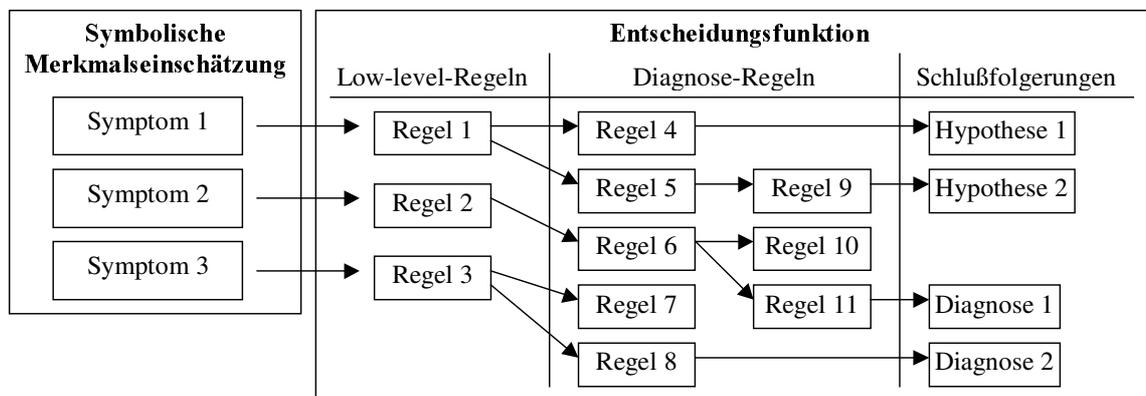


Abbildung 4-28 Schematische Darstellung der Entscheidungsfunktion; Low-level-Regeln reagieren auf neue Symptome und aktivieren Diagnose-Regeln, die Diagnosen durchführen. Dabei dienen Hypothesen als Zwischenergebnisse der Diagnose.

Kommt es innerhalb eines Inferenzzyklus zu einer Diagnose, so werden dem Bediener durch eine Protokollfunktion die Schlußfolgerungen und Zwischenergebnisse des Expertensystems mitgeteilt. Je nach Priorität werden die Mitteilungen auf einer Bildlaufleiste und im HTML-Format über einen WWW-Server verfügbar gemacht oder als elektronische Nachricht (e-mail) an vorher festgelegte Personen gesendet. Abbildung 4-29 zeigt den generellen Ablauf der Überwachung.

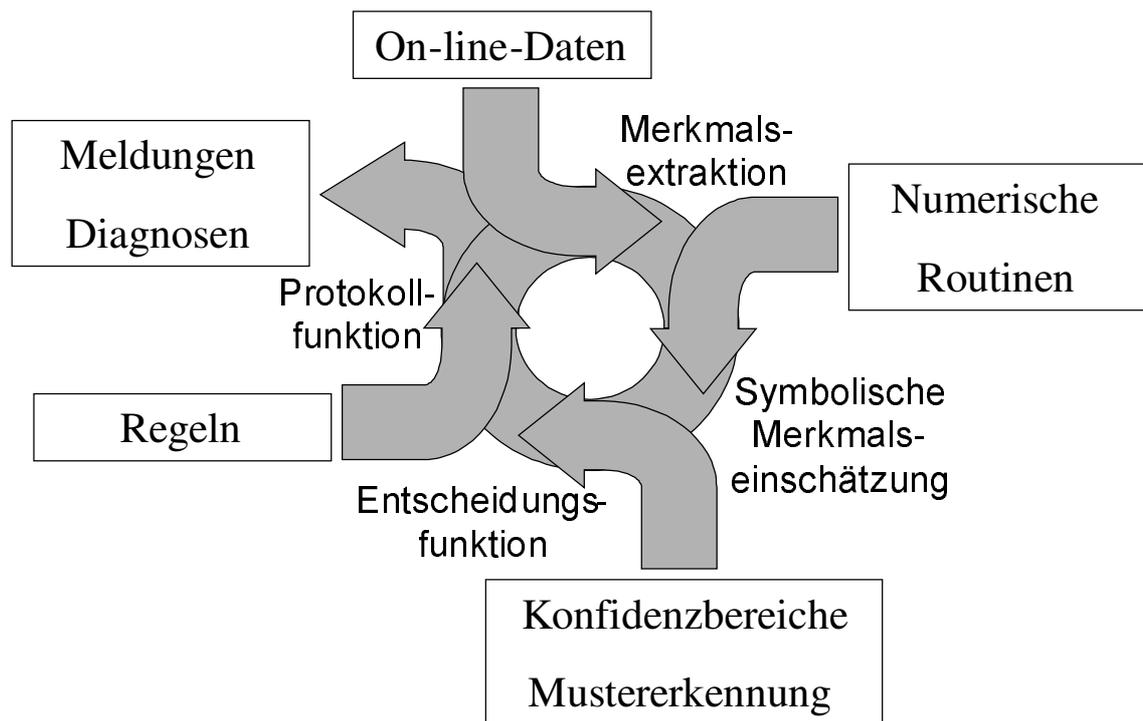


Abbildung 4-29 Arbeitsablauf des Überwachungssystems

4.4.5 Benutzerschnittstelle

Für die Kommunikation mit dem Bedienpersonal wurde eine graphische Benutzerschnittstelle entwickelt. Funktionell gliedert sich diese in die Bereiche Überwachung und Wissensakquisition. Der Bereich Überwachung beinhaltet alle Funktionen zum Protokollieren und Visualisieren des überwachten Prozesses. Im Bereich Wissensakquisition sind alle Funktionen enthalten, die benötigt werden, um das für die Überwachung notwendige Wissen zu akquirieren und in das System zu implementieren.

4.4.5.1 Protokollfunktion

Mit Hilfe der Protokollfunktionen werden Informationen über den Zustand des Bioprozesses und über den Ablauf der Diagnose ausgegeben. Die Information wird in verschiedenen Text- und Grafiken dem Bediener mitgeteilt. Abbildung 4-30 zeigt den Bildschirmaufbau des Expertensystems, wenn ein Fehler aufgetreten ist. Wesentliche Bestandteile sind die Bildlaufleiste, auf der alle Textmeldungen des Systems ausgegeben werden. Jede Meldung beinhaltet

die Namen der Regel bzw. Prozedur, welche die Meldung erzeugt hat und einem Meldungstext.

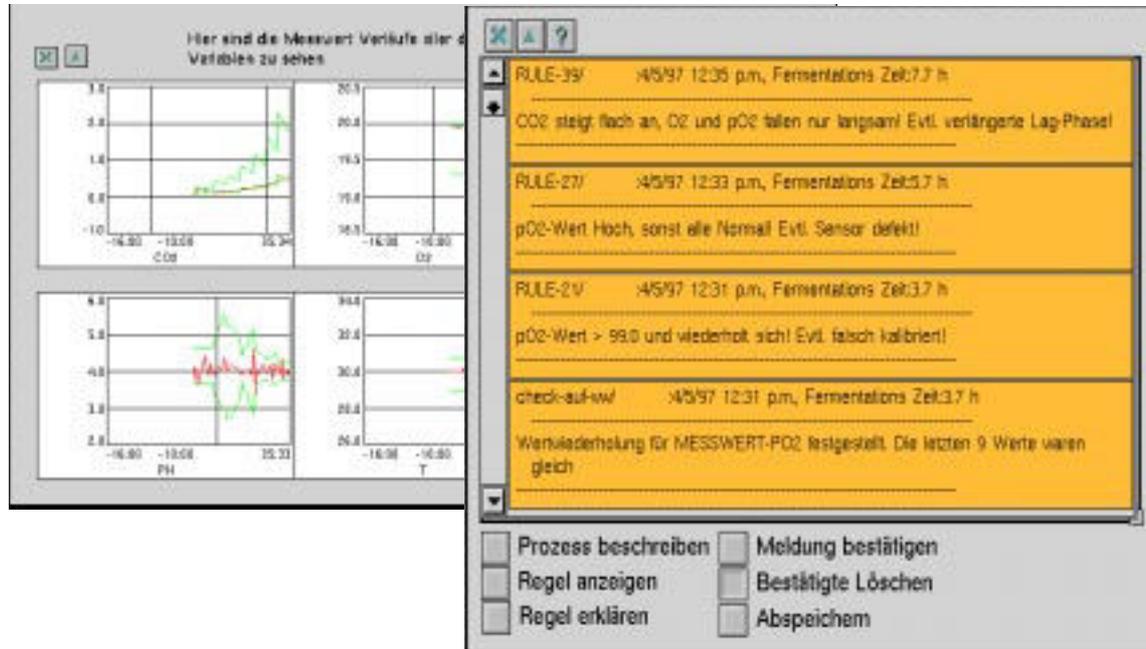


Abbildung 4-30 Bildlaufleiste mit Meldungen (rechtes Fenster) sowie Datenvisualisierung (linkes Fenster) während einer Überwachung

Von der Arbeitsfläche der Bildlaufleiste können weitere Menüpunkte aktiviert werden. Tabelle 4-13 enthält eine Liste der Menüpunkte mit einer kurzen Beschreibung.

Tabelle 4-13 Menüpunkte der Bildlaufleiste

Funktion	Beschreibung
Prozeß beschreiben	Gibt eine zusammenfassende Beschreibung des Prozeßzustands.
Regel anzeigen	Zeigt den Text der Regel an, welche die Meldung erzeugt hat.
Regel erklären	Ruft die Erklärungskomponente für eine ausgewählte Regel auf.
Meldung bestätigen	Ausgewählte Meldungen bestätigen.
Bestätigte Löschen	Bestätigte Meldungen löschen.
Abspeichern	Alle Meldungen in einer Datei abspeichern.

Wesentliche Aufgabe des Expertensystems ist die Überwachung eines Bioprozesses, um so den Bediener bei Routineaufgaben zu unterstützen. Stellt das Überwachungssystem einen kritischen Zustand beim Prozeß fest, so wird unmittelbar auf die Diagnose des Expertensystems hingewiesen. Zur weiteren Erläuterung der Schlußfolgerungen, werden zusätzlich Graphiken der relevanten Signalverläufe ausgegeben.

4.4.5.2 Erklärungskomponente

Die Erklärungskomponente beinhaltet Funktionen, um die Regeln der Wissensbasis durchsuchen und Schlußfolgerungen des Expertensystems transparent darstellen zu können. Der Diagnoseablauf des Systems kann mit dem kompletten Durchsuchen eines komplexen Regelnetzwerks verglichen werden, wobei die einzelnen Regeln in dem Netzwerk über "Kategorien" und "Focal-Objekte" miteinander in Verbindung stehen. Die Wissensbasis kann mit der Erklärungskomponente sowohl vorwärts- als auch rückwärtsgerichtet durchsucht werden. Bei der vorwärtsgerichteten Suche werden von der gegebenen „Kategorie“ und dem gegebenen „Focal-Objekt“ ausgehend alle Regeln gesucht, die der gesuchten Kategorie angehören. Werden von den so ermittelten Regeln weitere Regel-Kategorien aktiviert, so erfolgt für diese Kategorien auch eine Auflistung aller Regeln. Dieser Vorgang wird solange wiederholt, bis keine weiteren Regeln gefunden werden.

Mit einer rückwärtsgerichteten Suche werden für eine zu untersuchende Regel alle Regeln gesucht, die diese Regel aktivieren. Besitzen die so gefundenen Regeln wiederum Kategorien, werden für diese Kategorien die Regeln gesucht, die sie aktivieren. Dieser Vorgang wird solange wiederholt, bis keine weiteren Regeln gefunden werden. Bei der rückwärtsgerichteten Suche sollte für eine gegebene Diagnose-Regel das Regelnetzwerk bis zu den Low-level-Regeln führen. In der Abbildung 4-31 ist die Arbeitsfläche der Erklärungskomponente dargestellt, in der eine rückwärtsgerichtete Suche für eine Regel durchgeführt wurde. Die beiden angegebenen Low-level-Regeln sind hierbei direkt verantwortlich für die Aktivierung der untersuchten Regel. Das bedeutet daß diese Low-level-Regeln die Regel direkt aktivieren.

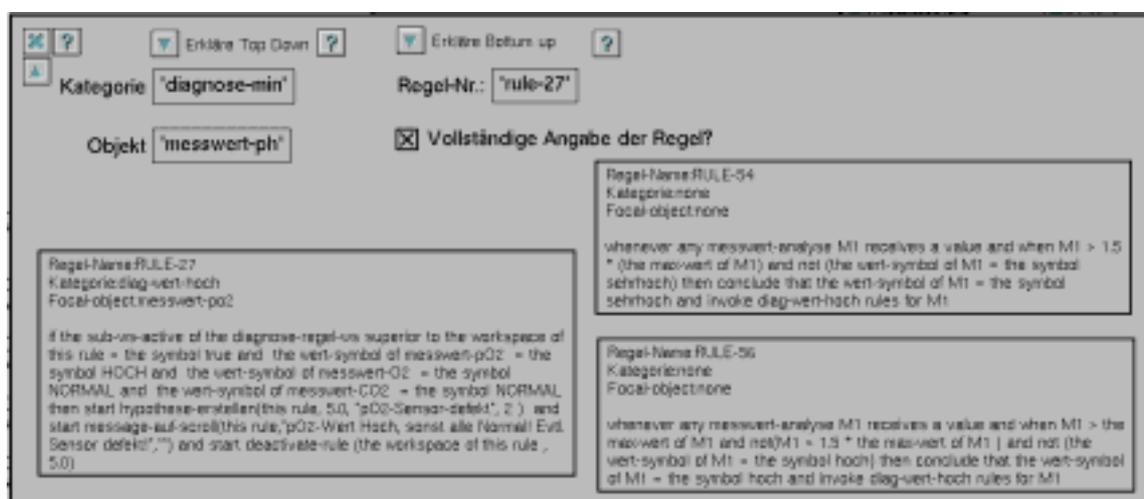


Abbildung 4-31 Arbeitsbereich der Erklärungskomponente

Neben einer übersichtlichen Darstellung der Zusammenhänge zwischen den einzelnen Regeln innerhalb des Diagnose-Netzwerks können durch die Erklärungskomponente zwei Fehler erkannt werden, die zu einer inkonsistenten Wissensbasis führen:

1. Aktivierungsschleifen
2. Nicht-aktivierbare Regeln

Aktivierungsschleifen kommen dann vor, wenn Regeln eine zyklische Aktivierungsreihenfolge besitzen. Im einfachsten Fall aktiviert eine erste Regel die Kategorie einer zweiten Regel, die selbst wiederum die Kategorie der ersten aktiviert. Wird solch eine Aktivierungsschleife angestoßen, so kann es zu einer Endlosschleife kommen, die schließlich zu einem Systemfehler führt.

Eine nicht-aktivierbare Regel wird bei der rückwärtsgerichteten Suche dann festgestellt, wenn die Suche keine Low-level-Regel liefert, da nur durch diese innerhalb des Diagnose-Netzwerks eine Inferenz angestoßen wird.

Die Abbildung 4-32 zeigt die vorwärts- und rückwärtsgerichtete Suche sowie den Fall einer nicht-aktivierbaren Regel und einer Aktivierungsschleife.

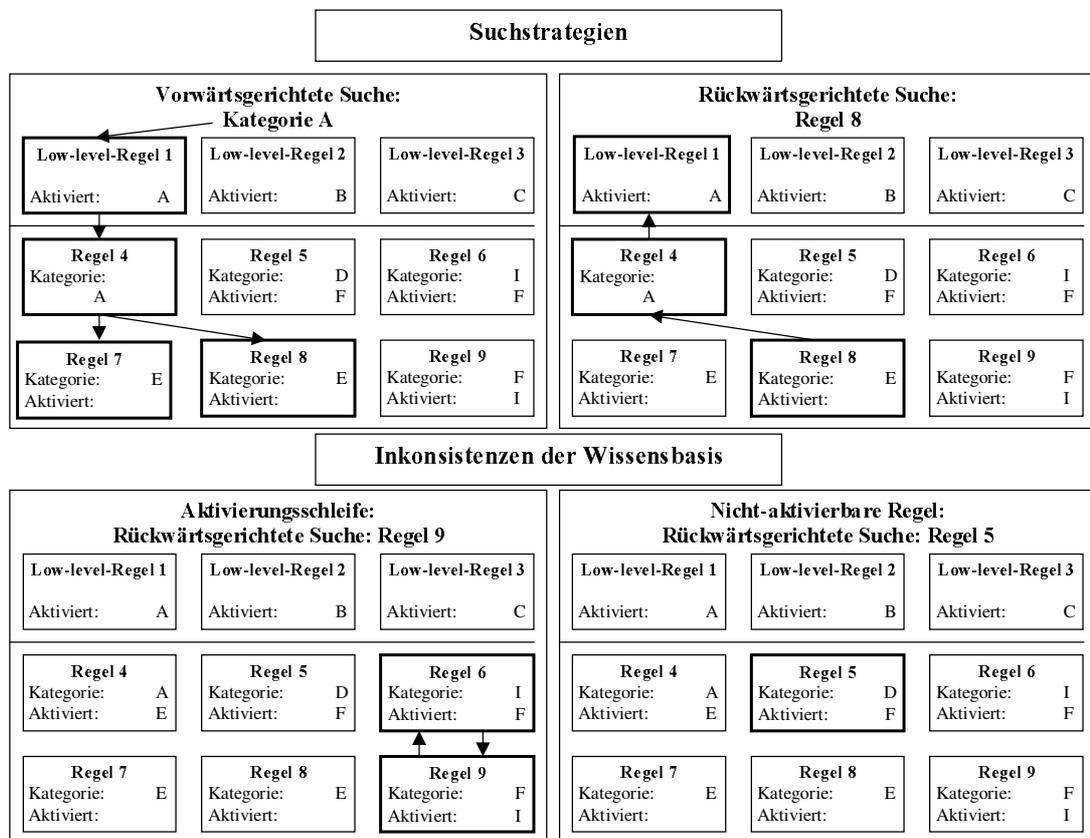


Abbildung 4-32 Schematische Darstellung der vorwärts- und rückwärtsgerichteten Suche sowie Beispiele für eine Aktivierungsschleife und eine nicht-aktivierbare Regel

Die Notwendigkeit einer Erklärungskomponente wächst mit dem Umfang der Wissensbasis. Insbesondere wenn mehrere Regelebenen innerhalb des Regel-Netzwerks zur Diagnose eines Störfalls vorhanden sind, lassen sich Fehler in der Wissensbasis nur schlecht bestimmen.

4.4.5.3 Zusätzliche Visualisierung über das Internet

Das World-Wide-Web, kurz als WWW oder W3 bezeichnet, ist ein Internet-Dienst für die Übertragung von Hyper-Text-Markup-Language-(HTML)-Dokumenten, auch HTML-Seiten genannt [Kirste et.al 1995]. Diese Dokumente können multimediale Erweiterungen enthalten. Das WWW wurde ursprünglich von Tim Berners-Lee am Cern in Genf entwickelt. Dieses System besteht aus Anbieterprogrammen (Servern), Dokumenten und Abfrageprogrammen (Clients). Die Besonderheit des WWW liegt darin, daß Dokumente weltweit und für jeden zugänglich sind, der an das Internet angeschlossen ist. Abfrage- und Visualisierungsprogramme (sog. Browser) sind für alle gängigen Computertypen und Betriebssysteme frei erhältlich. Vorteile des Internets liegen in der ständigen Verfügbarkeit von Informationen, da ein Zugriff rund um die Uhr möglich ist, und des großen Vernetzungsgrads, da mittlerweile fast jede Hochschule sowie große und mittelgroße Unternehmen mit dem Internet verbunden sind. Dadurch ist es möglich, den Verlauf eines Bioprozesses und die Meldungen des Expertensystems, die sonst nur auf der Konsole des Expertensystems dargestellt werden, über das Internet anderen Anwendern zugänglich zu machen.

Als Mensch-Maschine-Schnittstelle in herkömmlichen Systemen dient eine einfache Konsole, welche die Ergebnisse der Messungen in Form von Graphiken und Protokollen darstellt. Dadurch wird der Bediener aber örtlich und zeitlich stark gebunden. Durch eine global zugängliche Visualisierung über das WWW kann der Prozeßverlauf von jedem am Internet angeschlossenen Rechner aus beobachtet werden, die Meldungen des Expertensystems verfolgt und sogar Eingriffe vorgenommen werden. Darüber hinaus bietet sich die Möglichkeit mehrere Prozesse zu verfolgen.

Um die Techniken des WWW von dem Expertensystem aus zu nutzen, mußten spezielle Schnittstellen an einen WWW-Server implementiert werden. Diese Schnittstellen ermöglicht zum einen die Visualisierung der Meßsignale und zum anderen die Ausgabe der Meldungen des Expertensystems. Somit sind die wesentlichen Funktionen der Benutzerschnittstelle über das Internet zugänglich.

Es folgt eine schematische Beschreibung des Ablaufs, der Graphikerstellung aus den Prozeßdaten. Die Daten des Prozeßrechners UBICON werden über die TCP/IP Schnittstelle auf die Workstation – auf dem auch das Expertensystem implementiert ist – übertragen und zentral abgespeichert. Fordert ein Client die HTML-Seite mit den Graphiken an, so wird ein CGI-Script (Common Gateway Interface, eine quasi standardisierte Schnittstelle zwischen WWW-Server und lokalem Programm) ausgeführt, welches aus den aktuellen Daten Graphiken im GIF-Format (Graphical Interchange Format) erstellt. Dieses Format wurde von CompuServe speziell für die Übertragung von Bildern entwickelt, da hier durch eine Komprimierung die zu übertragende Datenmenge stark reduziert wird. Diese GIF-Graphiken werden vom CGI-Script an den WWW-Server weitergeleitet, wodurch sie auf einer HTML-Seite im WWW weltweit verfügbar sind. Zwischen der Abfrage durch einen Client und der Erstellung sowie der vollständigen Übermittlung der HTML-Seite vergehen je nach Art der Netzwerkverbindung 10 bis 60 Sekunden. Abbildung 4-33 zeigt in einem Flußdiagramm die einzelnen Vorgänge, die notwendig sind, um durch einem WWW-Client eine HTML-Seite mit den aktuellen Prozeßdaten anzufordern.

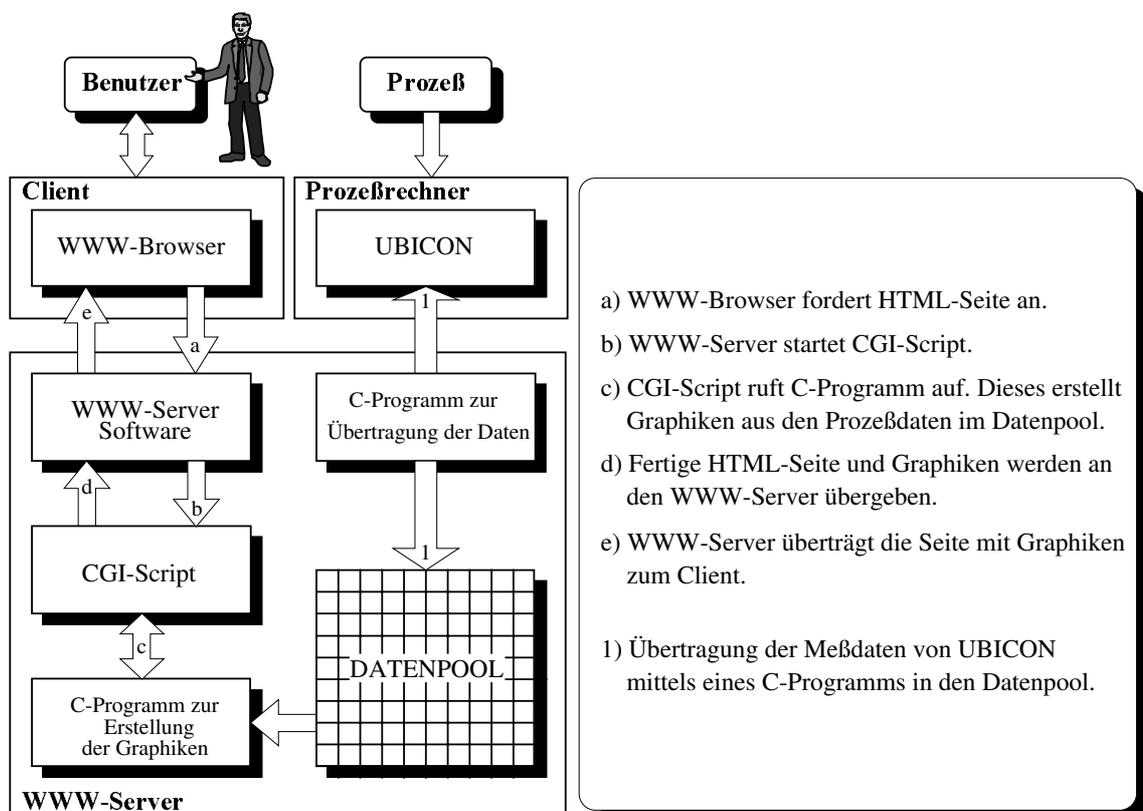


Abbildung 4-33 Flußdiagramm der Vorgänge, wenn ein WWW-Client die Prozeßgraphiken anfordert

Neben der Visualisierung der aktuellen Prozeßdaten ist es für den Bediener wichtig, ebenfalls eine Beurteilung über den Zustand des Prozesses zu erhalten. Zu diesem Zweck werden alle Meldungen des Expertensystems, die auf der Bildlaufleiste angezeigt werden, als HTML-formatierter Text in einem dem WWW-Server und damit weltweit zugänglichen Verzeichnis abgelegt, wodurch dieser Text ebenfalls verfügbar wird. Abbildung 4-34 und Abbildung 4-35 zeigen Bildschirmausgaben eines WWW-Browsers und der HTML-Seiten, die von dem Expertensystem zur Verfügung gestellt werden.

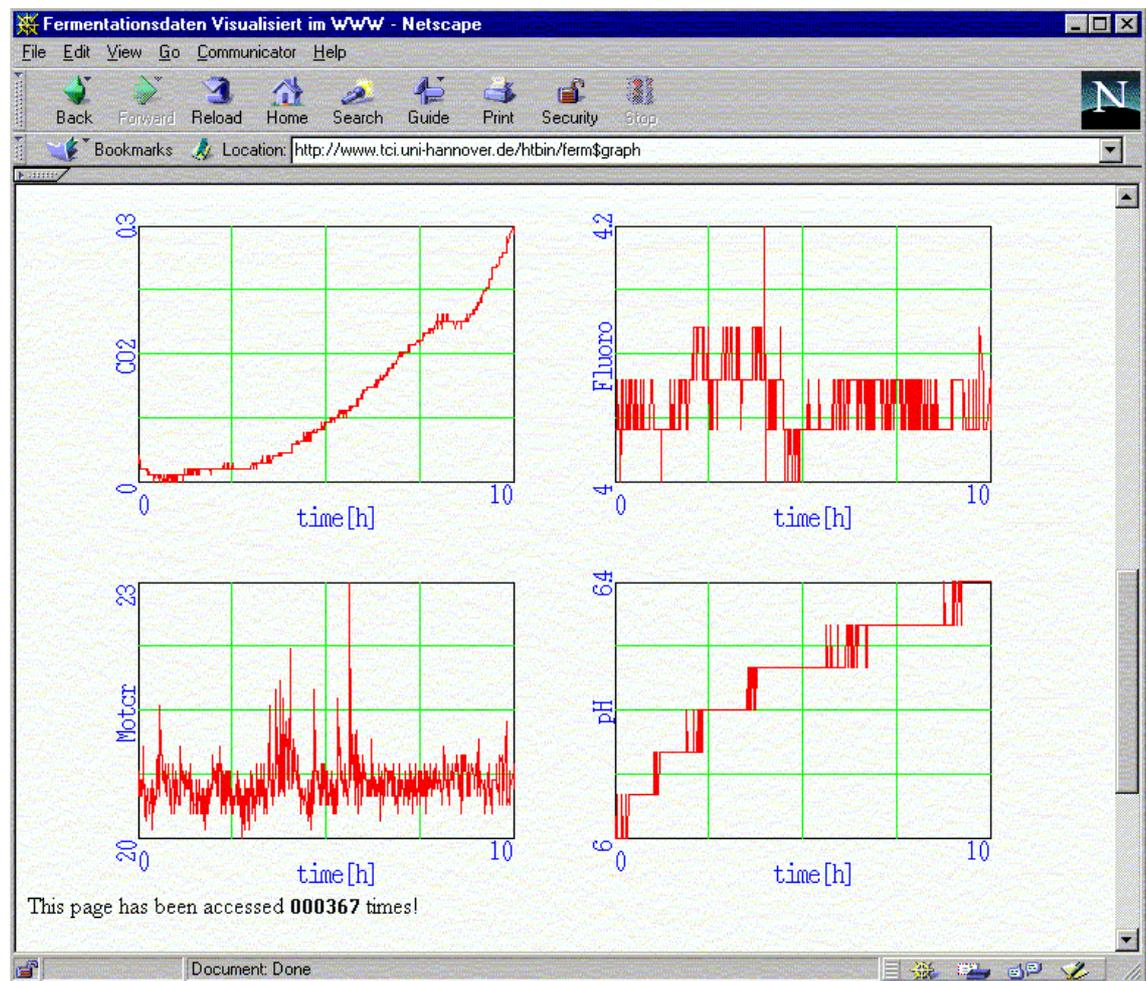


Abbildung 4-34 zeigt einen Bildschirmausdruck des WWW-Browsers mit der HTML-Seite zur Prozeßdatenvisualisierung

Die Vorteile dieses Verfahrens liegen in der hohen Geschwindigkeit, der Datenweiterleitung, sowie der plattformunabhängigen Software, die den Datenzugriff ermöglichen.

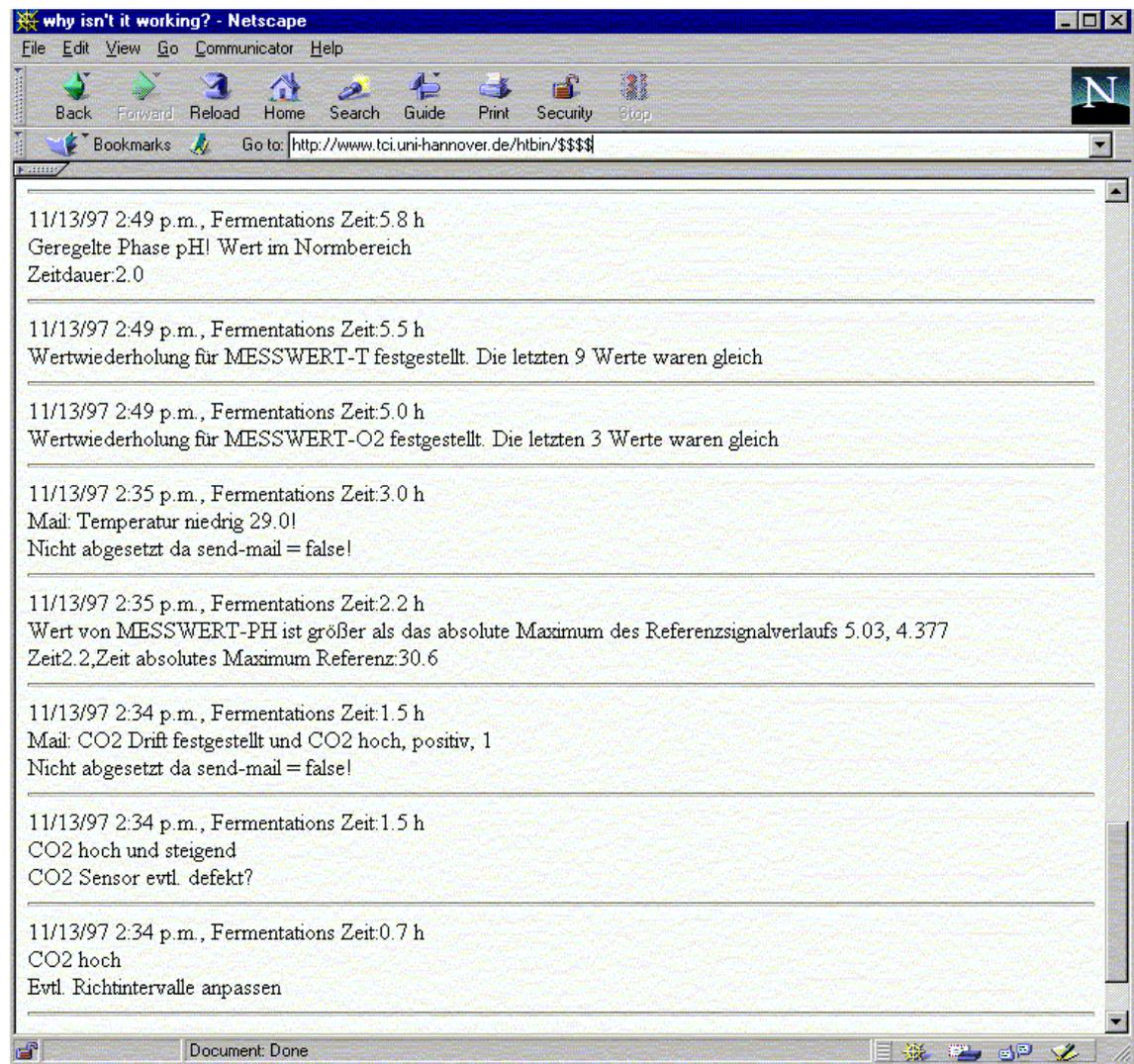


Abbildung 4-35 zeigt einen Bildschirmausdruck des WWW-Browsers mit der HTML-Seite zur Darstellung der Meldungen des Expertensystems

Bei der in Abbildung 4-35 dargestellten Meldungen handelt es sich um Meldungen einer Simulation, weshalb die in den Meldungen angegebene Kultivierungszeit nicht mit dem angegebenen Zeitpunkt der Meldungserstellung übereinstimmt.

4.4.5.4 Wissenserwerbskomponente

Eine Anforderung an das Expertensystem ist die flexible Adaption an neue Prozesse. Dieser Anforderung wird durch eine leistungsfähige Wissenserwerbskomponente Rechnung getragen. Die Anpassung des Expertensystems an einen neuen Prozeß beinhaltet die Wissensakquisition zwei verschiedener Quellen, mit denen eine neue prozeßspezifische Wissensbasis erstellt wird.

Zu dem Wissen, daß akquiriert werden muß, gehört prozeßspezifisches Faktenwissen und methodisches Wissen.

4.4.5.4.1 Prozeßspezifisches Faktenwissen

Die Wissensakquisition von prozeßspezifischen Wissen beinhaltet im wesentlichen die Verwaltung und Analyse von Daten vergangener Kultivierungen. Wobei die Benutzerschnittstelle zahlreiche Dialoge und Funktionen bietet, um den Bediener bei der Durchführung von Datenanalysen zu unterstützen. Abbildung 4-36 zeigt den Ablauf der Datenanalyse und die Aufgabenbereiche, bei denen der Benutzer interaktiv unterstützt wird. Die Ergebnisse der Datenanalyse werden innerhalb des Expertensystems in Objekten der Klasse ‘Datenanalyse‘ gespeichert. Wie in Kapitel 4.4.5.1.4 beschrieben gliedert sich die Klasse ‘Datenanalyse‘ mehrere Unterklassen.

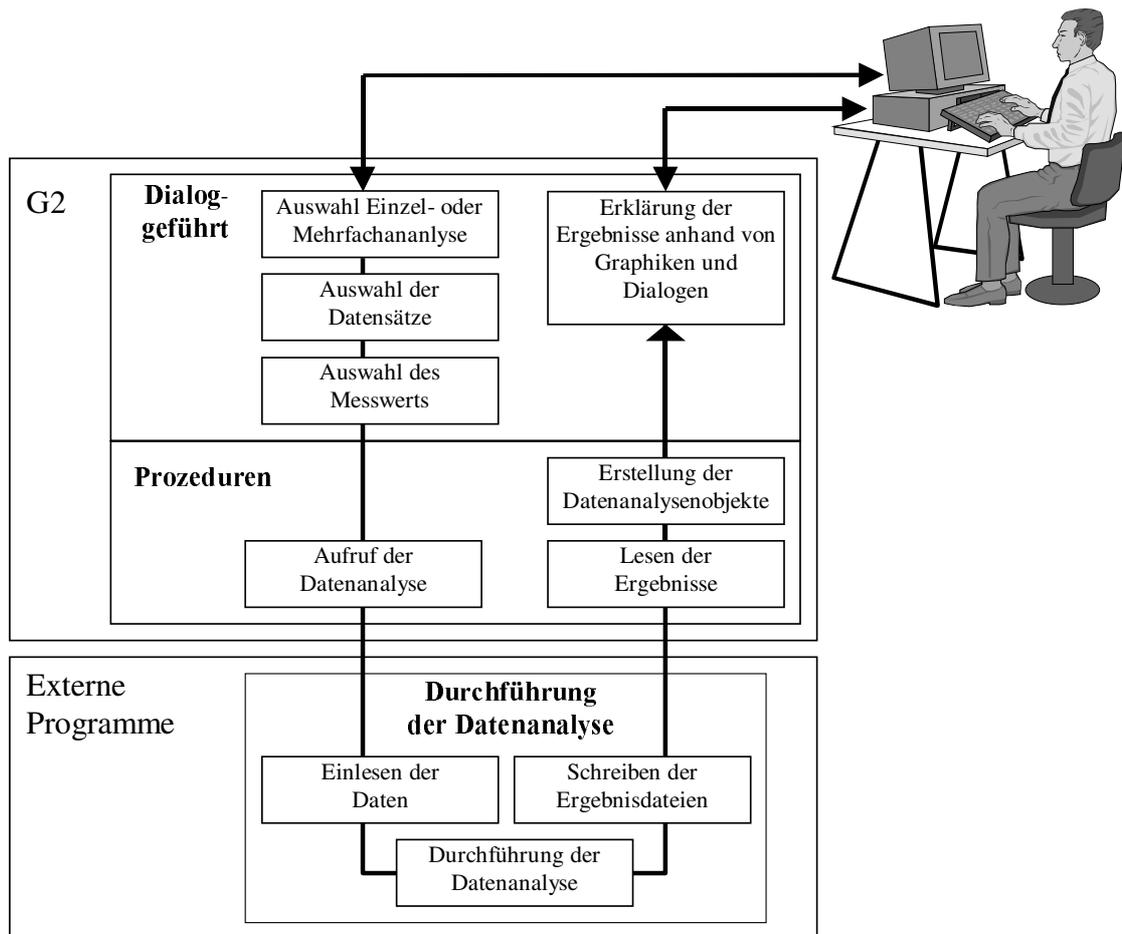


Abbildung 4-36 Durchführung der Datenanalyse

Für jedes Objekt der Klasse ‚Datenanalyse‘ kann eine Beschreibung der Ergebnisse angefordert werden, daß in sprachlich verständlicher Form zusammengefaßt ist. Abbildung 4-37 zeigt den Text der den Inhalt eines bestimmten Objekts beschreibt.

#164 2:56:29 p.m. Dieses Objekt beschreibt den Signalverlauf von MESSWERT-PH. Die Daten wurden mit einer Messfrequenz von 0.017 h gespeichert.

Die Kreuzkorrelation hat ein maximale zeitliche Verschiebung der Daten um -1.3 h festgestellt. Der Verlauf weist insgesamt 5 Maxima auf, das absolute Maximum von 6.889 tritt nach 25.89 h auf.

Der Verlauf weist insgesamt 5 Minima auf, das absolute Minimum von 3.87 tritt um 92.19 h auf.

Der Mittelwert beträgt 5.408 und die Standardabweichung 1.152.

Die 1. Ableitung hat ein Maximum von 20.42 um 133.34 h und ein Minimum von -5.96 um 133.34 h.

Es wurde eine maximale Wiederholung von 7 Werten festgestellt.

Das mittlere Rauschen betrug 0.002 mit einer Standardabweichung von 0.031. Daraus folgt ein mittleres Signal-Rausch-Verhältnis von $3.794e-4$.

Die Analyse hat den Signalverlauf in 10 Trends aufgeteilt.

Abbildung 4-37 Die gespeicherte Information des Objekts kann in sprachlich verständlicher Form ausgegeben werden

Der Verlauf der Intervallgrenzen der Konfidenzbereiche kann auf Anforderung graphisch dargestellt werden. Die von G2 erstellten Graphiken sind in der Abbildung 4-38 dargestellt.

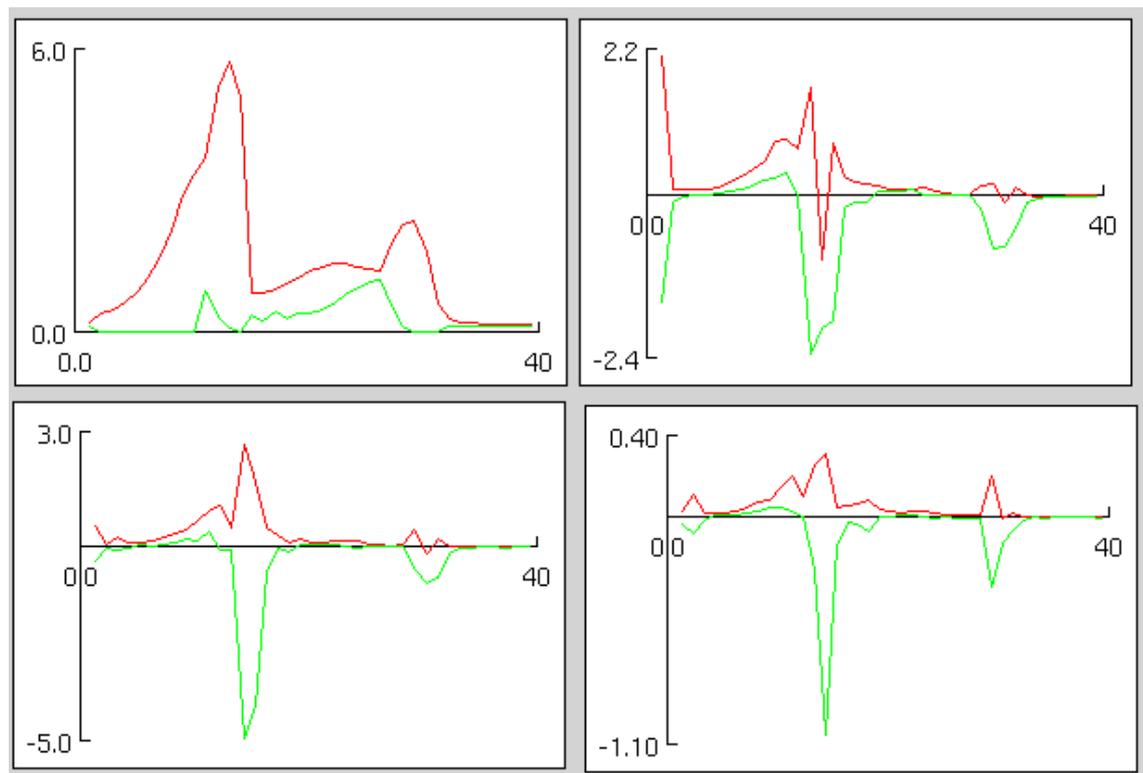


Abbildung 4-38 Die zeitlichen Verläufe der Intervallgrenzen der Konfidenzbereiche können in dem Expertensystem graphisch dargestellt werden

4.4.5.4.2 Methodisches Wissen

Die Akquisition von methodischem Wissen stellt eine Transformation der Heuristiken in eine für den Computer verständliche Form dar. Die Syntax der Regeln zur Repräsentation des methodischen Wissens ist im Kapitel 4.4.4.1.3 dargestellt worden. Anhand der dort aufgeführten Beispiele wird deutlich, daß es eine fast unüberwindliche Aufgabe darstellt – ohne besondere Kenntnisse des Expertensystems – sinnvolle Regeln zu erstellen. Um dieses Hindernis zu umgehen, bietet die Benutzerschnittstelle die Möglichkeit, Regeln für die Überwachung vollständig benutzergeführt zu erstellen. Der Ablauf der Wissensakquisition für die Erstellung einer Diagnose-Regel ist in Abbildung 4-39 dargestellt.

In den ersten Abfragen wird der Bedingungsteil einer Regel erstellt. Hierzu wird aus der Wissensbasis eine Auswahl der in Frage kommenden Objekte zusammengestellt. Für jede Objektklasse besteht eine spezielle Objektdefinition, die Information über die in einer Regel verfügbaren Attribute und ihre zulässigen Werte beinhaltet. Mit dieser Information kann in der Benutzerführung gewährleistet werden, daß der Bedingungsteil der Regel syntaktisch richtig ist.

Abbildung 4-40 zeigt eine Auswahl der Dialoge, die bei der Erstellung des Bedingungsteils der Regel Verwendung finden.

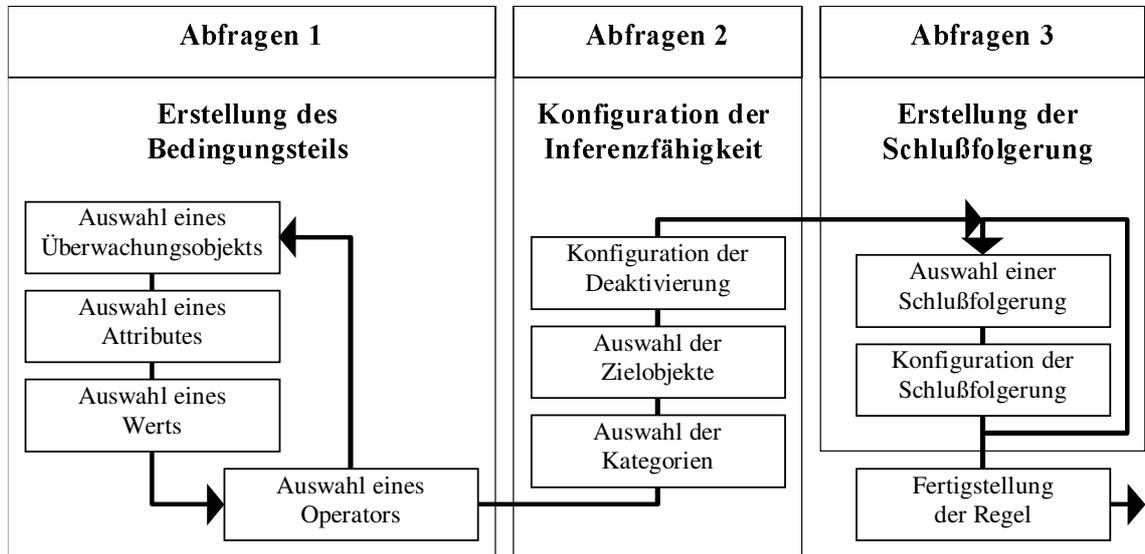


Abbildung 4-39 Schematischer Ablauf der Erstellung einer Diagnose-Regel

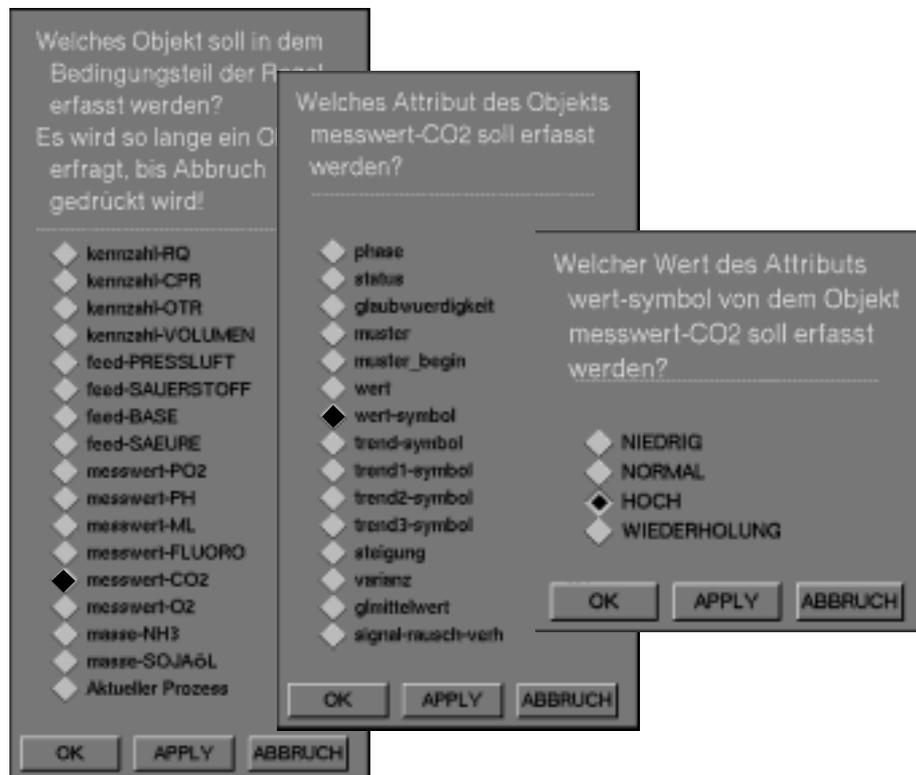


Abbildung 4-40 Dialoge, welche die Prämisse einer Regel erfragen

Die zweite Abfrage legt spezifische Eigenschaften der Regel fest, die für die Inferenz wichtig sind. Hierzu gehören die Kategorien und Focal-Objekte der Regel, die darüber entscheiden, wann die Regel von der Inferenzmaschine aktiviert wird. Darüber hinaus kann eine Regel – nachdem sie ausgeführt worden ist – deaktiviert werden. Dies ist dann sinnvoll, wenn mehrfache Meldungen des Systems unterbunden werden sollen. Abbildung 4-41 zeigt einige der Dialoge, die für die Konfiguration der Regel verwendet werden.

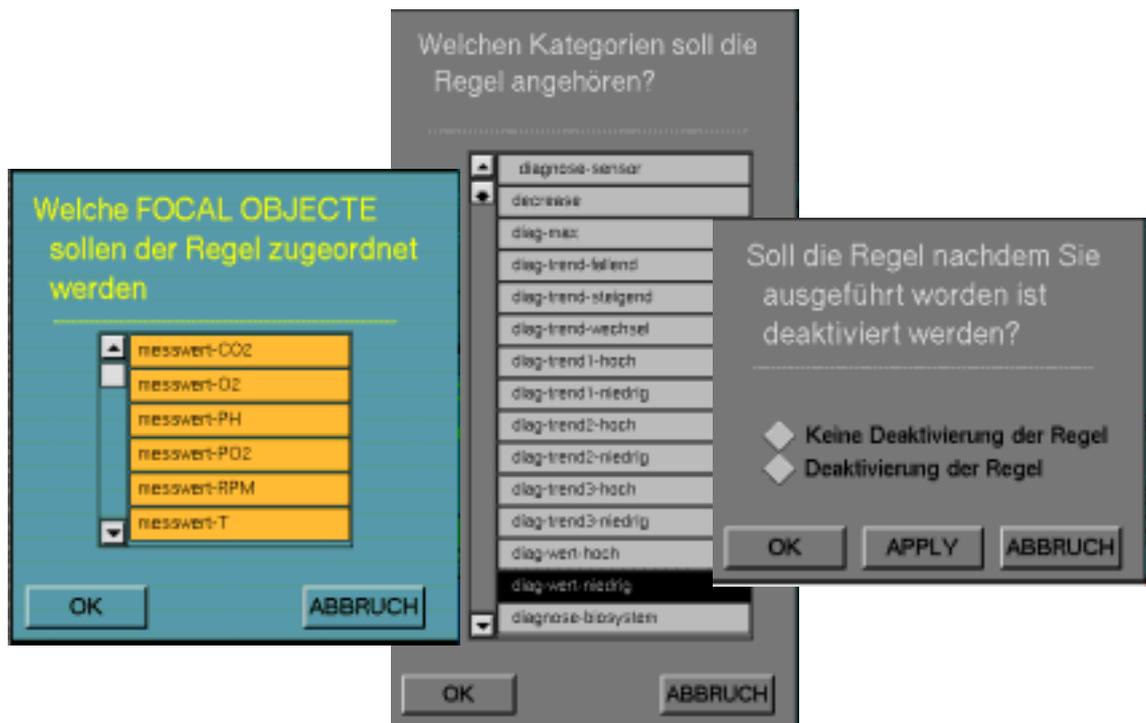


Abbildung 4-41 Dialoge erfragen Eigenschaften der Regel, die für die Inferenzmaschine wichtig sind

In der dritten Abfrage, werden schließlich die Aktionen der Regel festgelegt. Hier stehen drei verschiedene Aktionen zur Auswahl, die Tabelle 4-14 aufgezählt sind.

Tabelle 4-14 Auswahl der Aktionen einer Diagnoseregeln

Aktion	Bedeutung
Benachrichtigungen	Bediener benachrichtigen (Bildlaufleiste oder E-Mail)
Weitere Regeln aktivieren	Aktivierung weiterer Regeln, zur Durchführung weiterer Diagnose
Hypothese erstellen	Erstellung einer Hypothese als Zwischenergebnis der Diagnose
Schlußfolgerung tätigen	Erzeugung von neuem Faktenwissen

Eine Meldung auf der Bildlaufleiste der Bedienoberfläche oder eine Benachrichtigung des Bediener über E-mail wird durch die Aktion ‚Benachrichtigung‘ erzeugt. Hierzu werden der Inhalt der Benachrichtigung sowie die Internet Adressen des Bedieners erfragt, an den die Benachrichtigung verschickt werden soll.

Mit der Aktion ‚Weitere Regeln aktivieren‘ kann die Inferenzmaschine angewiesen werden, weitere Regeln zu aktivieren. Somit kann beispielsweise eine tiefergehende Diagnose ausgelöst werden. Im Dialog werden die für die Aktivierung notwendige Kategorie sowie das ‚Focal-Objekt‘ erfragt.

Bei der Aktion ‚Schlußfolgerung tätigen‘ kann die Wissensbasis gezielt verändert werden. Hierzu können die Werte ausgewählter Objekte verändert werden.

Des Weiteren lassen sich mit der Aktion ‚Hypothese erstellen‘ Objekte der Klasse ‚Hypothese‘ erstellen, die Zwischenergebnisse von Diagnosen darstellen. Abbildung 4-42 zeigt einige Dialoge, die für die Erstellung der Aktion einer Regel verwendet werden.

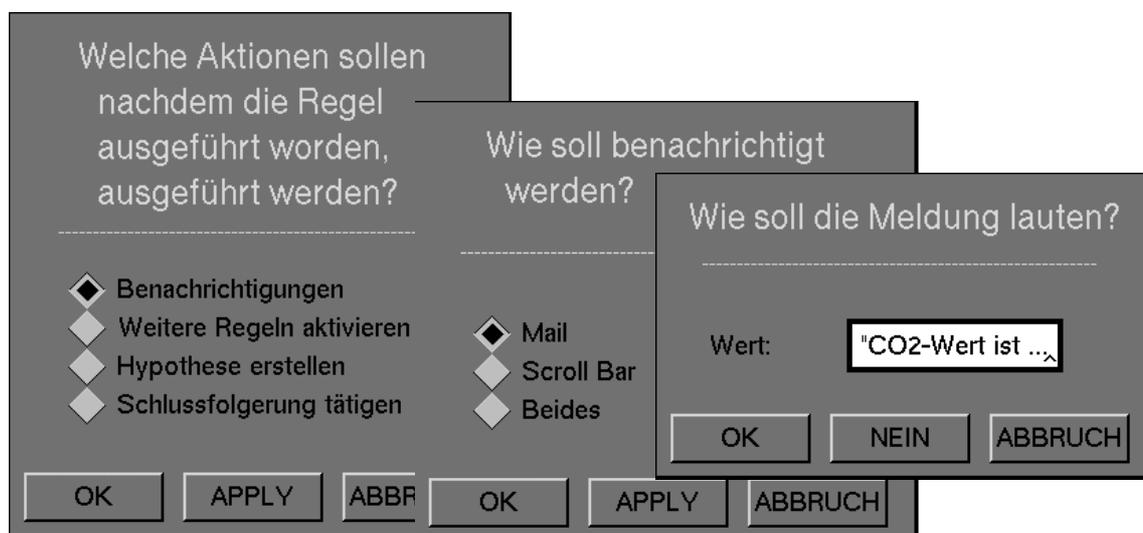


Abbildung 4-42 Dialoge zur benutzergeführten Erstellung der Konklusion einer Regel; dargestellt ist die Konfiguration einer Regel, die eine Meldung an den Benutzer verschickt

4.5 Anwendungsbeispiele

In diesem Kapitel werden ausgewählte Anwendungsbeispiele mit Störfällen vorgestellt, die beim Einsatz des Expertensystems beobachtet wurden. Damit soll die konkrete praktische Bedeutung der Überwachung und Fehlerdiagnose von Bioprocessen herausgestellt und der Ar-

beitsablauf des Expertensystems erläutert werden. Zunächst werden Beispiele beschrieben, die beim On-line-Einsatz des Expertensystems am Prozeß beobachtet wurden, für den das System ursprünglich entwickelt wurde. Um zu zeigen, daß das Expertensystem auch dem Anspruch gerecht wird, flexibel an neue Prozesse angepaßt werden zu können, werden nachfolgend einige Anwendungen mit anderen Prozessen vorgestellt.

4.5.1 Überwachung einer Fed-batch-Kultivierung von *Acremonium chrysogenum*

Heuristisches Prozeßmodell

Um einen unnormalen Prozeßzustand bzw. Signalverlauf von dem 'normalen', erwarteten Verlauf unterscheiden zu können, muß der 'normale' Verlauf der Kultivierung als symbolisches Modell in dem System implementiert werden. Abbildung 4-43 zeigt die Vorgehensweise des Benutzers bei der Implementierung eines heuristischen Prozeßmodells in dem Expertensystem. Ein heuristisches Prozeßmodell beschreibt im wesentlichen spezifische Prozeßvorgänge und ihre Auswirkungen auf den Verlauf einzelner Prozeßgrößen.

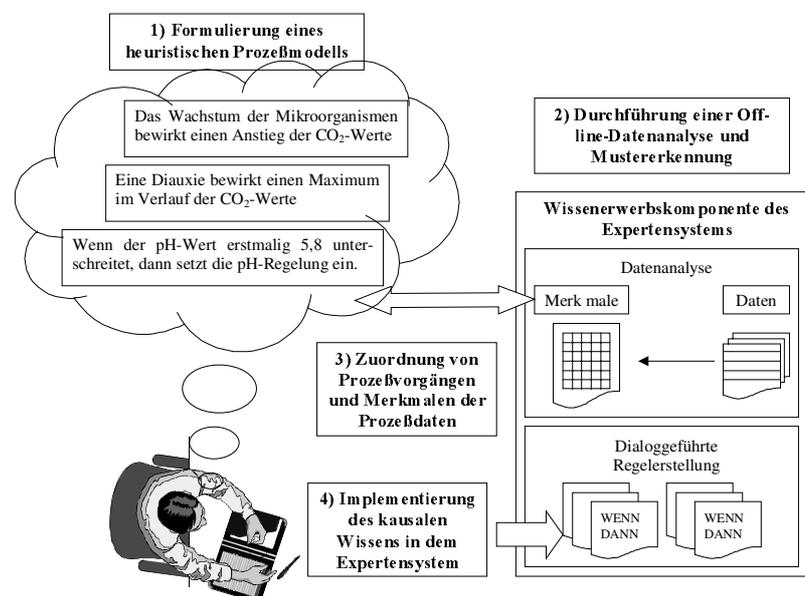


Abbildung 4-43 Vorgehensweise des Benutzers bei der Implementierung eines Prozeßmodells in dem Expertensystem

Dadurch werden spezifische Ursache-Wirkungs-Zusammenhänge formuliert, die für die Ermittlung des Prozeßzustands genutzt werden können. Das Expertensystem muß aber in der

Lage sein die spezifischen Auswirkungen anhand von ermittelten Merkmalen bzw. Merkmalsänderungen zu erkennen. Eine vom Bediener angeforderte Off-line-Datenanalyse liefert spezifische Merkmale, die von dem Expertensystem aus den Prozeßdaten für die Überwachung ermittelt werden. Der Bediener muß die im Prozeßmodell beschriebenen Vorgänge bestimmten Merkmalen zuordnen. Dieses kausale Wissen muß durch die Formulierung von Regeln in dem Expertensystem implementiert werden.

Im folgenden soll ein repräsentativer Verlauf einer Kultivierung des im Kapitel 3 beschriebenen Beispielprozesses und die Implementierung des Wissens über einen ‚normalen‘ Verlauf im Expertensystem beschrieben werden. Bei dem Prozeß handelt es sich um die Kultivierung eines *Acremonium chrysogenum* Stamms in einem 30 Liter Rührkesselreaktor unter Fed-batch-Bedingungen mit einem auf Maisquellwasser (CSL) basierendem Komplexmedium zur Produktion von Cephalosporin-C. Die wesentlichen Einflußgrößen, die zum charakteristischen Verlauf von Prozeßgrößen führen, sind zum einen das Verhalten der Mikroorganismen sowie zum anderen verfahrenstechnische Operationen, beispielsweise regelnde Eingriffe. Die verschiedenen Einflüsse überlagern sich häufig, wobei sie sich gegenseitig aufheben oder verstärken können.

Im folgenden soll die Beschreibung des Prozeßverlaufs einzelner Prozeßgrößen auf die vier wichtigsten Größen pH, pO_2 , O_2 im Abgas und CO_2 im Abgas beschränkt werden.

Bei Betrachtung eines repräsentativen Verlaufs der pH-Werte, der in Abbildung 4-44 dargestellt ist, ist erkennbar, daß er zu Beginn der Kultivierung ca. 6,6 beträgt. Dieser bleibt während die Mikroorganismen wachsen zunächst über eine Zeitdauer von ca. 25 h konstant, um dann innerhalb von 10 Stunden stark abzufallen. Die Ursache für diesen Abfall hängt mit der Umsetzung von in dem Medium gelöstem Ammonium zusammen. Je mehr von dieser basischen Komponente dem Medium durch die Mikroorganismen entzogen wird, desto saurer wird das Medium. Anfänglich wirkt sich dieser Ammonium-Verbrauch nur gering auf den pH-Wert aus, da die Pufferkapazität des Mediums den pH-Wert konstant hält. Unterschreitet der pH den Wert 5,8 so setzt die pH-Regelung ein die – durch Zugabe einer Ammonium Lösung – den pH-Wert auf einen konstanten Wert von 5,8 regelt. Dies ist insbesondere im Hinblick auf die pH-Wert abhängige Hydrolysegeschwindigkeit des Produktes bedeutsam.

Soll der charakteristische Verlauf der pH-Werte von dem Expertensystem überwacht werden, so muß der Benutzer in einer Off-line-Sitzung zunächst überprüfen, ob die von dem Expertensystem ermittelten Merkmale ausgeprägt genug sind, um eine Erkennung zu gewährleisten. In

der Abbildung 4-44 ist das Ergebnis der Mustererkennung für diesen repräsentativen Kultivierungsverlauf dargestellt. Ab der 25. bis zur 30. Stunde wird dem pH-Verlauf ein abfallendes Referenzmuster (DECR) zugeordnet, welches den Verlauf charakterisiert. Anschließend ermittelt die Mustererkennung das konstante Referenzmuster (STABLE) als das Muster, daß nun den Verlauf der pH-Werte der Kultivierung am besten widerspiegelt. Die Ermittlung des Merkmals ‚STABLE‘ kann also von dem Expertensystem genutzt werden, um den Anfang der pH-Regelung zu erkennen.

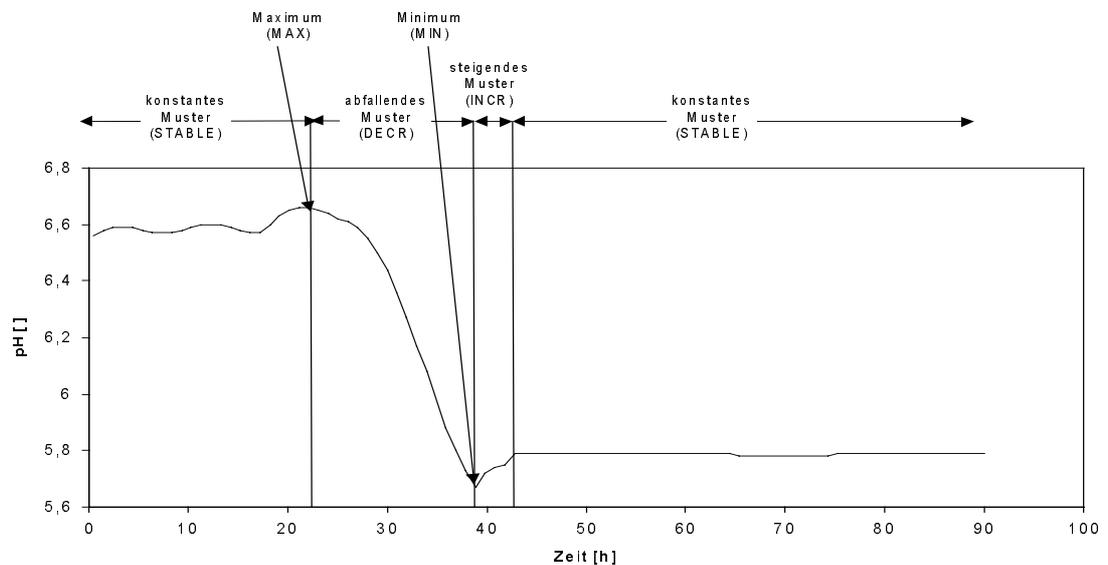


Abbildung 4-44 Verlauf der pH-Werte einer repräsentativen Kultivierung und Darstellung der Ergebnisse der Mustererkennung; Änderungen des zugeordneten Musters dienen als Merkmal für die Erkennung von spezifischen Prozeßsituationen

Der Verlauf der pO_2 -Werte der Kultivierung ist dem der pH-Werte sehr ähnlich und liegt zu Beginn bei 100 %. Innerhalb der ersten 30 Stunden sinkt der pO_2 -Wert kontinuierlich auf einen Wert von etwa 60 % ab, wie in der Abbildung 4-45 zu ersehen ist. Durch Wachstum und Stoffwechselfvorgänge der Kultur wird der im Nährmedium gelöste Sauerstoff verbraucht, so daß der pO_2 -Wert sinkt. Um eine Sauerstofflimitierung während der Kultivierung zu verhindern, setzt ab einem pO_2 -Wert von 60 % eine Regelung des pO_2 ein. Die Regelung basiert auf einer Zudosierung von reinem Sauerstoff zu der Preßluft, die für die Begasung verwendet wird. Daraus resultiert ein – bis auf leichte Schwankungen – konstanter pO_2 -Wert. Bei der Regelung wird der Gesamtvolumenstrom der Begasung konstant gehalten. Da nur eine Sauerstofflimitierung, nicht aber eine Überversorgung, sich negativ auf die Produktion auswirkt,

wird ein Wiederansteigen des pO_2 über den kritischen Wert von 60 % Luftsättigung durch den Regler toleriert, falls der Sauerstoffbedarf kurzfristig absinkt.

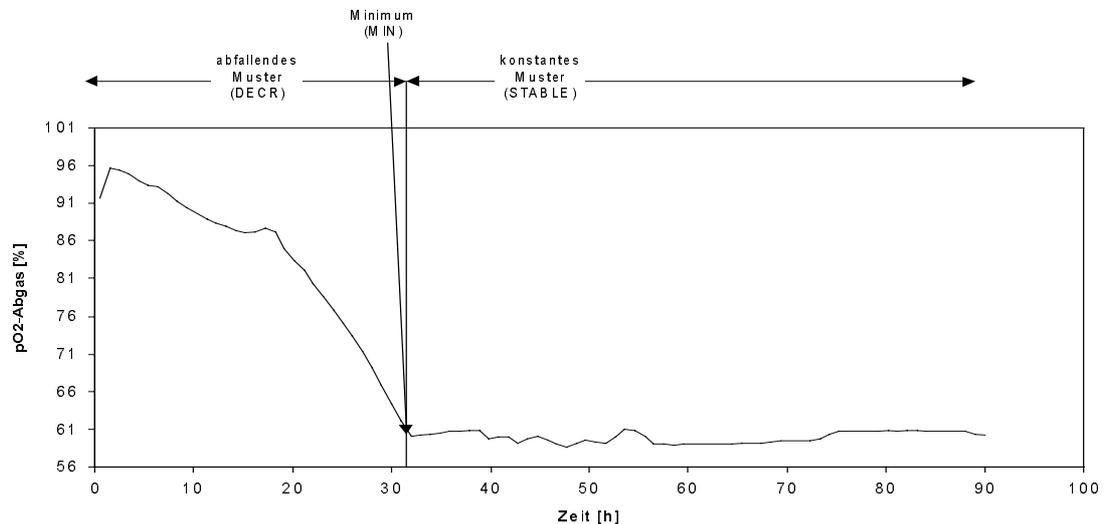


Abbildung 4-45 Verlauf der pO_2 -Werte einer repräsentativen Kultivierung und Darstellung der Ergebnisse der Mustererkennung

In der Abbildung 4-45 sind die Ergebnisse der Mustererkennung für den Verlauf der pO_2 -Werte eingetragen. Die Zuordnung des stetigen Referenzmusters (STABLE) kann im Falle des pO_2 Verlaufs als Merkmal verwendet werden, an denen der Beginn der Sauerstoffregelung erkannt werden kann.

In direktem Zusammenhang mit dem pO_2 im Nährmedium steht der Sauerstoffgehalt im Abgas. Die Abnahme des pO_2 durch die Stoffwechselaktivität der Kultur wird durch den Eintrag von Sauerstoff aufgrund der Begasung ausgeglichen. Die Größe der Differenz des O_2 -Gehalts zwischen Zu- und Abluft gibt die vom Medium aufgenommene O_2 -Menge wieder und läßt sich direkt mit der Aktivität der Mikroorganismen korrelieren. Sobald aber die Sauerstoffregelung einsetzt, wird der Begasungsluft reiner Sauerstoff zugesetzt. Dies führt zu einem erhöhten Anteil Sauerstoff in der Begasung und somit zunächst zu einem Anstieg des O_2 Gehaltes im Abgas, bis ein relativ konstanter Wert erhalten wird. Der Anstieg des O_2 im Abgas nach 58 Stunden auf ca. 31 % ist ein Hinweis darauf, daß in der zweiten Wachstumsphase der Sauerstoffbedarf der Kultivierung sinkt, und die Regelung einen Überschuß an Sauerstoff in der Be-

gasungsluft bewirkt. Abbildung 4-46 zeigt den Verlauf der O_2 -Werte sowie die Ergebnisse der Mustererkennung für diesen Prozeßverlauf.

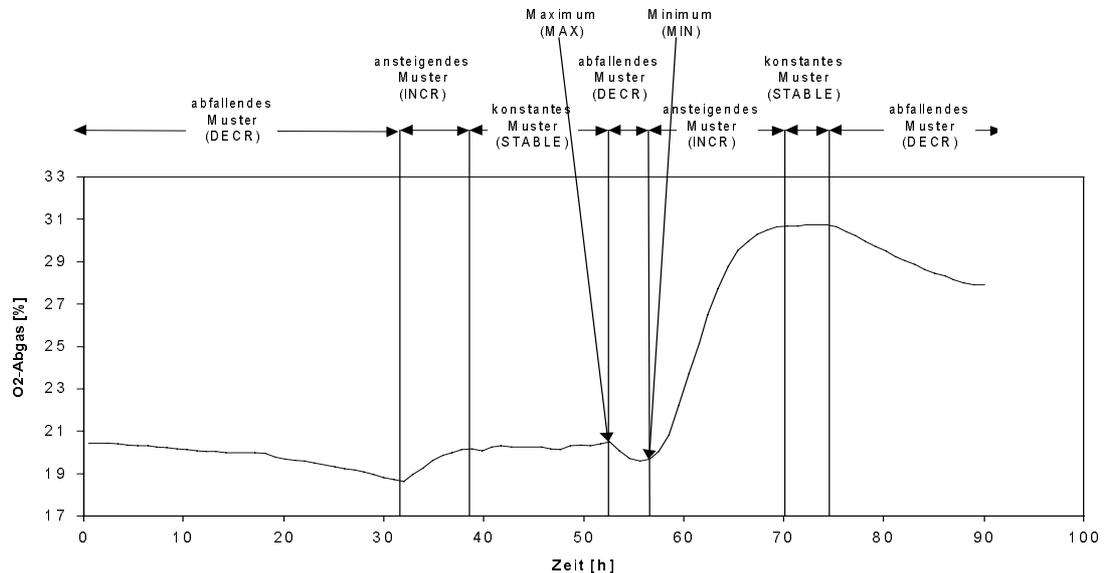


Abbildung 4-46 Verlauf der O_2 -Werte einer repräsentativen Kultivierung und Darstellung der Ergebnisse der Mustererkennung

Auch der CO_2 Gehalt des Abgases wird maßgeblich durch den Stoffwechsel der Mikroorganismen beeinflusst. In Abbildung 4-47 ist der Verlauf der CO_2 -Werte der repräsentativen Kultivierung dargestellt. Die bei den Stoffwechselfvorgängen umgesetzten Nährstoffe werden teilweise als Kohlendioxid an das Medium abgegeben. Da die Löslichkeit von CO_2 in dem Medium begrenzt ist, wird ein Teil davon an die Begasungsluft abgegeben und führt zu einem erhöhten Anteil von CO_2 im Abgas. Da in der anfänglichen Wachstumsphase der Mikroorganismen die Stoffwechselftätigkeit stetig zunimmt, führt dies auch zu einer stetigen Zunahme des CO_2 Gehalts im Abgas.

Nach etwa 50 Stunden erreicht der CO_2 -Gehalt im Abgas einen Maximalwert von etwa 3 %. Der anschließende Abfall des CO_2 läßt sich durch den vollständigen Verbrauch leicht umsetzbaren C-Quellen erklären. Die hierdurch verursachte Limitierung führt zu einer Phase reduzierter Stoffwechselfaktivität, in der weniger CO_2 produziert wird. Durch eine Stoffwechselfumstellung auf andere verfügbare C-haltige Nährstoffe ist nach etwa 8 Stunden ein erneuter Anstieg im CO_2 -Gehalt zu verzeichnen. Dieses Phänomen wird als Diauxie bezeichnet.

Nach einer erneuten Wachstumsphase – die mit einem stetigen Anstieg im CO_2 -Gehalt der Abluft einher geht – stellt sich ein Gleichgewicht zwischen der Absterbe- und Wachstumsrate der Mikroorganismen ein. In diesem als stationäre Phase bezeichneten Abschnitt der Kultivierung bleibt die Stoffwechselaktivität und somit auch der CO_2 -Gehalt in der Abluft relativ konstant.

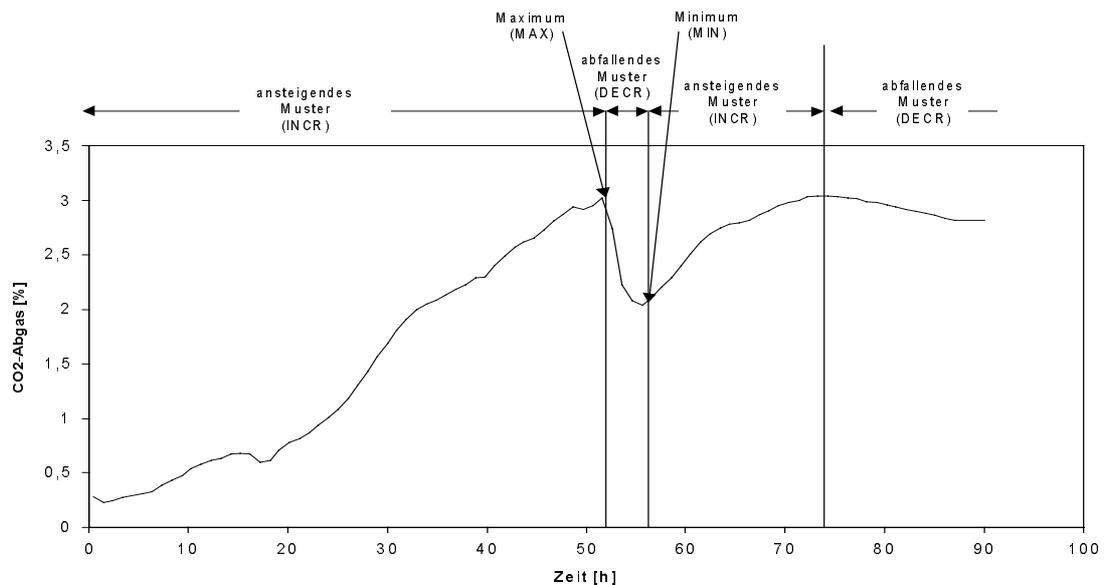


Abbildung 4-47 Verlauf der CO_2 -Werte einer repräsentativen Kultivierung und Darstellung der Ergebnisse der Mustererkennung

Das zuverlässige Erkennen des Prozesszustands ist von großer Bedeutung bei der Überwachung und Regelung von Bioprocessen. Häufig werden Regelstrategien gewählt, die sich stark an den spezifischen Phasenverlauf der Kultivierung anlehnen. So wurde bei dem Beispielprozess eine Fütterstrategie durchgeführt, bei der nach der Diauxie die Zufütterung von Soja-Öl initiiert wurde. Um diese Fütterstrategie zu automatisieren, muß das Expertensystem zunächst den Phasenverlauf des Prozesses feststellen und überwachen können.

Tabelle 4-15 zeigt zusammengefaßt die charakteristischen Muster, die zur Erkennung der beschriebenen Erscheinungen in den Prozeßgrößen genutzt werden können. Als Phase wird die Prozeßdauer zwischen zwei auftretenden Prozeßerscheinungen bezeichnet. Mit dem 'normalen Muster' wird das Referenzmuster bezeichnet, das vorwiegend in der Phase von der Mustererkennung ermittelt wird. Ändert sich das charakteristische Muster eines Signalverlaufs, ohne das ein Phasenwechsel definiert wird, so ist dies in der Tabelle in der Zeile 'Wechsel in Phase' gekennzeichnet.

Tabelle 4-15 Unterteilung des Prozeßverlaufs in Phasenabschnitte, die sich mit Hilfe der Mustererkennung erkennen lassen

	CO ₂	O ₂	PO ₂	PH
Phase 1	Erste Wachstumsphase			
Norm. Muster	INCR	DECR	DECR	STABLE
Wechsel in Phase	-----	-----	-----	-----
Muster Ende	-----	INCR	STABLE	DECR
Phase 2	Beginn der pO₂-Regelung			
Norm. Muster	INCR	INCR	STABLE	DECR
Wechsel in Phase	-----	STABLE	-----	-----
Muster Ende	-----	-----	-----	INCR
Phase 3	Beginn der pH-Regelung			
Norm. Muster	INCR	STABLE	STABLE	STABLE
Wechsel in Phase	-----	-----	-----	-----
Muster Ende	MAX	-----	-----	-----
Phase 4	Diauxie			
Norm. Muster	DECR	DECR	STABLE	STABLE
Wechsel in Phase	-----	-----	-----	-----
Muster Ende	MIN	MIN	-----	-----
Phase 5	Zweite Wachstumsphase			
Norm. Muster	INCR	INCR	STABLE	STABLE
Wechsel in Phase	-----	-----	-----	-----
Muster Ende	-----	-----	-----	-----

Die Muster, die charakteristisch für ein Prozeßphänomen sind und somit das Ende einer Phase signalisieren, sind in der Tabelle unter 'Muster Ende' aufgeführt.

Um mit dem Expertensystem den Prozeßzustand zu überwachen, sind in dem System Regeln implementiert, die das Auftreten der Änderungen dieser Muster überprüfen. In Abbildung 4-48 sind einige dieser prozeßspezifischen Regeln schematisch dargestellt, die den Phasenverlauf der repräsentativen Kultivierung überwachen. Die Pfeile zeigen eine von mehreren möglichen Wegen der Inferenzabfolge.

Da das Auftreten bestimmter Merkmale nicht in einen Inferenzzyklus fallen muß, werden Hypothesen formuliert, die in späteren Inferenzzyklen bestätigt werden. So wird in dem Beispiel, das in Abbildung 4-48 dargestellt ist, bei dem Phasenübergang von Phase 4 zu Phase 5 zunächst ein Minimum in dem Signalverlauf der O₂-Werte erkannt und als Schlußfolgerung eine Hypothese erstellt, daß der Phasenwechsel begonnen hat. Wird in einem späteren Inferenzzyklus ein Minimum in dem CO₂-Abgas festgestellt, so führt dies zur Bestätigung der Hypothese.

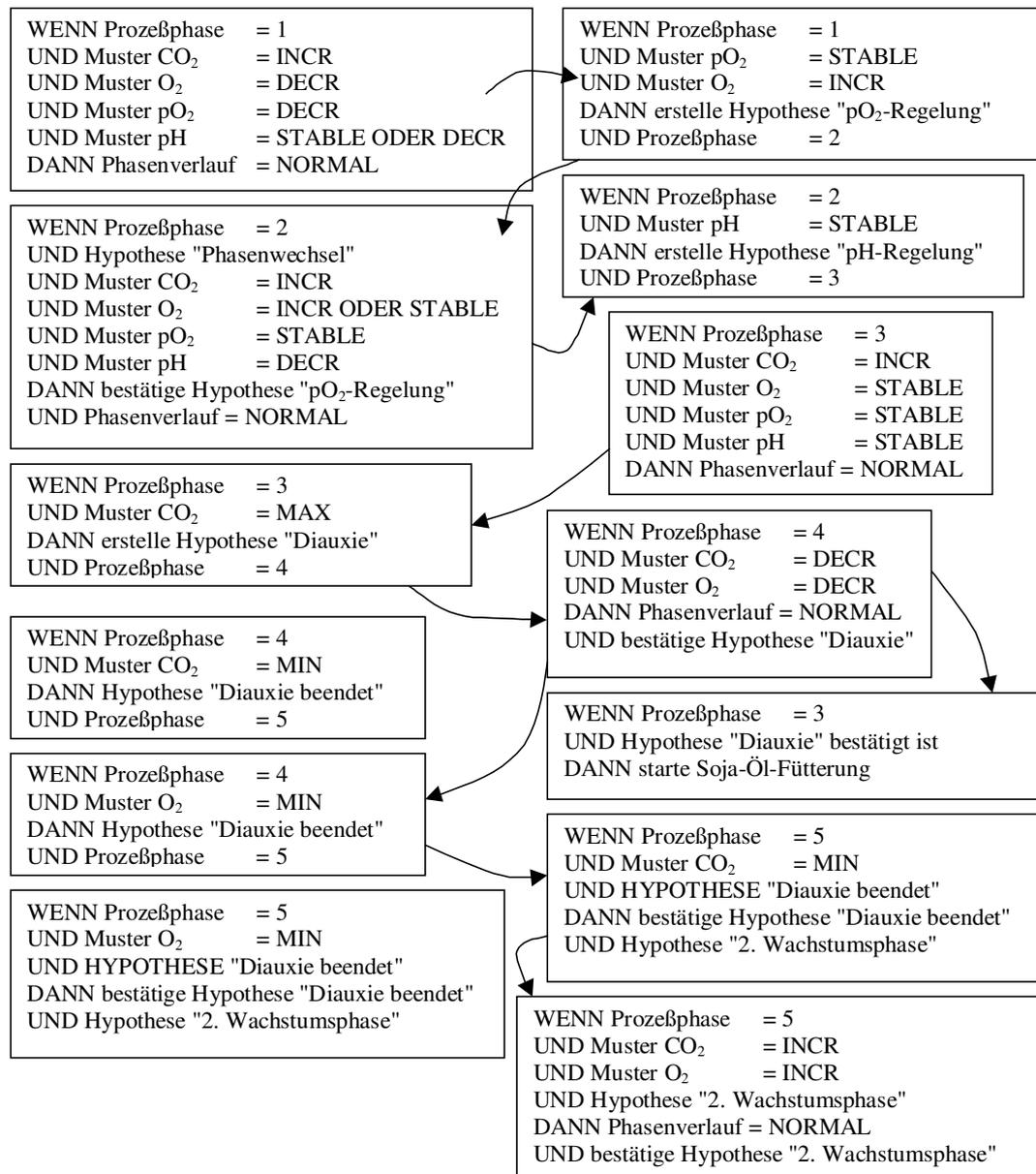


Abbildung 4-48 Schematische Darstellung einiger Regeln zur Phasenerkennung

Denkbar ist auch die Möglichkeit, bei der die Muster in umgekehrter Reihenfolge auftreten. Dieser Fall muß ebenfalls vom Benutzer – durch die Implementierung entsprechender Regeln – berücksichtigt werden. Die Zugabe von Soja-Öl wird von einer zusätzlichen Regel aus gestartet, die anspricht sobald die Hypothese "Diauxie" bestätigt wurde. Beim On-line-Einsatz des Expertensystems wurde eine Benachrichtigung an den Bediener gesendet.

Zusätzlich zur Überwachung der sequentiellen Abfolge von erwarteten Mustern wird mit separaten Regeln das zeitliche Auftreten der Phasenwechsel überwacht. Hierbei sind für die einzelnen Phasenübergänge Zeitintervalle definiert, in denen der Übergang stattgefunden haben soll-

te. Den Anfangszeitpunkt des Phasenwechsels markiert den Zeitpunkt, an dem die erste Merkmalsänderung festgestellt wurde. Nachdem alle Merkmalsänderungen aufgetreten sind, gilt der Phasenwechsel als abgeschlossen. Abweichungen von diesem normalen Phasenverlauf deuten auf einen Störfall hin. Hierbei können generell folgende Fälle eintreten:

1. Der Anfang des Phasenwechsels tritt verspätet auf,
2. Der Phasenwechsel findet zu früh statt und
3. Der Phasenwechsel kann nicht erkannt werden.

Fall 1) und 2) deuten auf Störfälle im biologischen System hin. Fall 1) kann z. B. auftreten, wenn sich der Stoffwechsel der Mikroorganismen nach dem Überimpfen nur sehr langsam an das Medium gewöhnt und es dadurch zu einer verlängerten ersten Phase kommt. Fall 2) tritt z. B. bei Kontaminationen auf. Diese treten häufig am Anfang eines Prozesses auf, wenn ein schnell wachsender Keim in Konkurrenz zu dem erwünschten Organismus tritt. Durch das starke Wachstums des infizierenden Keims ist, das Auftreten der ersten Phase stark verkürzt.

Fall 3) deutet in der Regel auf eine Sensorstörung hin, da nur für diesen Fall die Auswirkung auf den Verlauf einzelner Meßgrößen beschränkt bleibt.

Beispiel einer Störfalldiagnose

Für den Verlauf der Prozeßgrößen und der zugehörigen Merkmale werden im Vorfeld der Überwachung Datenanalysen durchgeführt und der erwartete zeitliche Verlauf dieser Merkmale durch obere und untere Intervallgrenzen vorgegeben. Verlassen die Merkmale im Verlauf der Kultivierung den durch die Intervallgrenzen vorgegebenen Konfidenzbereich, so deutet dieses auf einen anormalen Zustand des Systems hin, der näher untersucht werden muß.

In Abbildung 4-49 ist der Verlauf der pO_2 -Werte graphisch dargestellt, bei dem ein Störfall aufgetreten ist. Das auffälligste Symptom des Störfalls ist, daß der Signalverlauf starken Schwankungen unterliegt. Dieses Symptom macht sich anhand des numerischen Merkmal Rauschen bemerkbar. In Abbildung 4-49 ist der Verlauf dieses Merkmals gezeigt, wobei zu sehen ist, daß die obere Grenze des Konfidenzbereichs nach 32 Stunden der Kultivierungszeit überschritten wird. Über Regeln wird die Schlußfolgerung generiert, daß das Rauschen der pO_2 -Werte ‚HOCH‘ ist werden weitere Regeln zur näheren Diagnose aktiviert. Als mögliche Ursachen für diesen anormalen Verlauf des Signals kommen zum einen eine Störung des Sensors in Frage und zum anderen eine nicht optimal eingestellter Regler. Zur Unterscheidung dieser beiden möglichen Ursachen werden weitere Merkmale betrachtet.

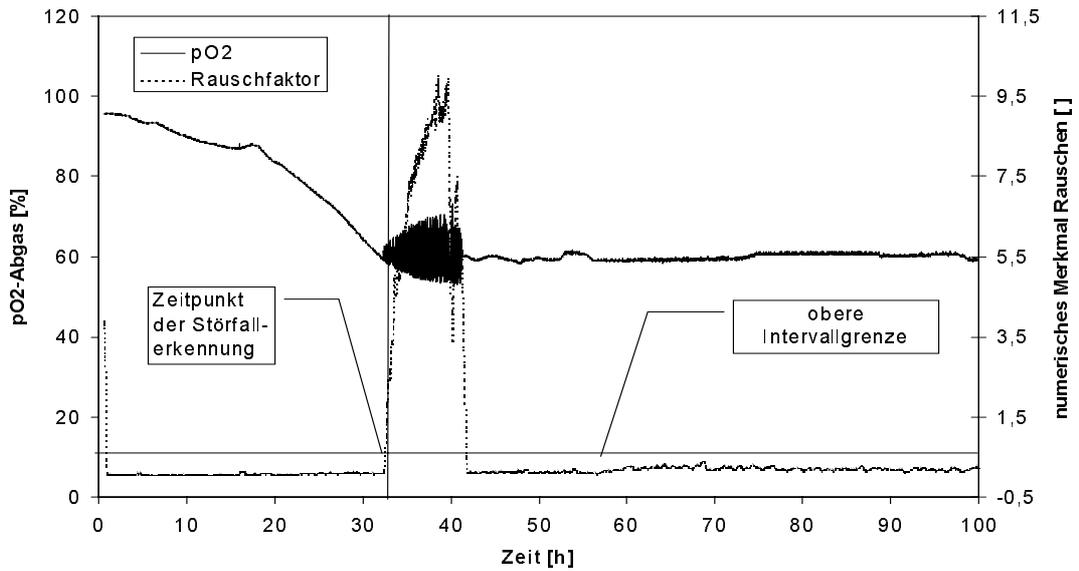


Abbildung 4-49 Verlauf der pO₂-Werte mit auftretender Störung und dem Merkmal 'Rauschen'

Liegt die Ursache der Störung in der Regelung, so macht sich dies auch an dem Verlauf der O₂-Abgas-Werte bemerkbar, da von der Regelung zugegebene Sauerstoff nicht vollständig vom Medium aufgenommen wird und dadurch ebenfalls zu Schwankungen dieser Werte führt. Für den gegebenen Fall trifft dies – wie in der Abbildung 4-50 zu sehen ist – zu, da auch bei den O₂-Werten das Merkmal Rauschen die obere Intervallgrenze überschreitet und das Symbol ‚HOCH‘ zugeordnet bekommen.

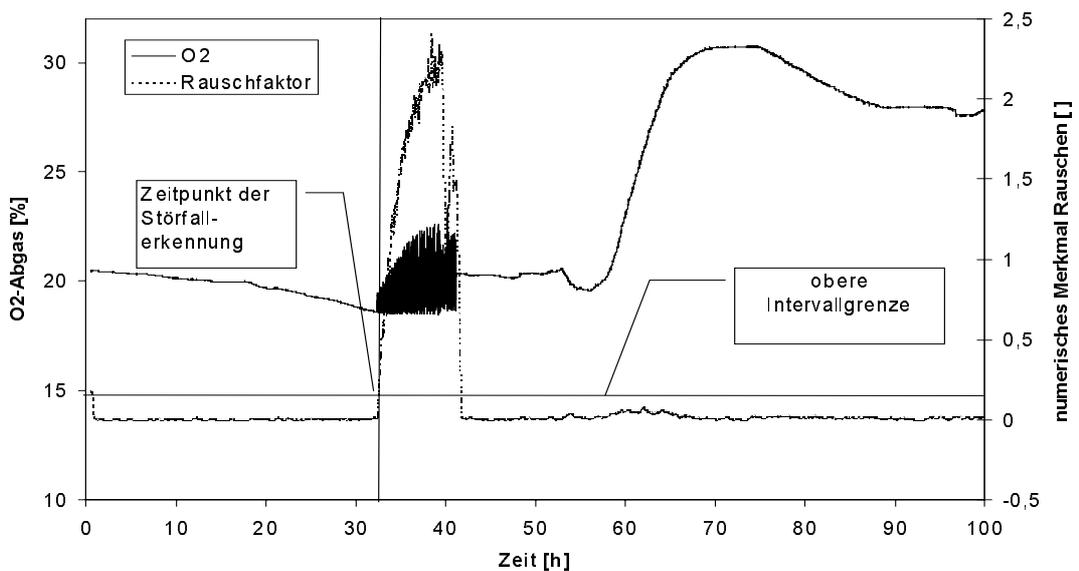


Abbildung 4-50 Verlauf der O₂-Abgas-Werte und der Kennzahl 'Rauschen'

Ein weiteres Kriterium, das auf eine nicht optimale Reglereinstellung als Ursache für das Rauschen hindeutet, ist die Zeit, an dem der Störfall auftritt, da diese häufig am Beginn der Regelung erscheinen. Nach dem vorliegenden Prozeßmodell ist dies der Beginn der zweiten Phase bzw. das Ende der ersten Phase. Die Regeln zur Diagnose eines Rauschens der pO_2 -Werte können also, wie in Abbildung 4-51 dargestellt, formuliert sein.

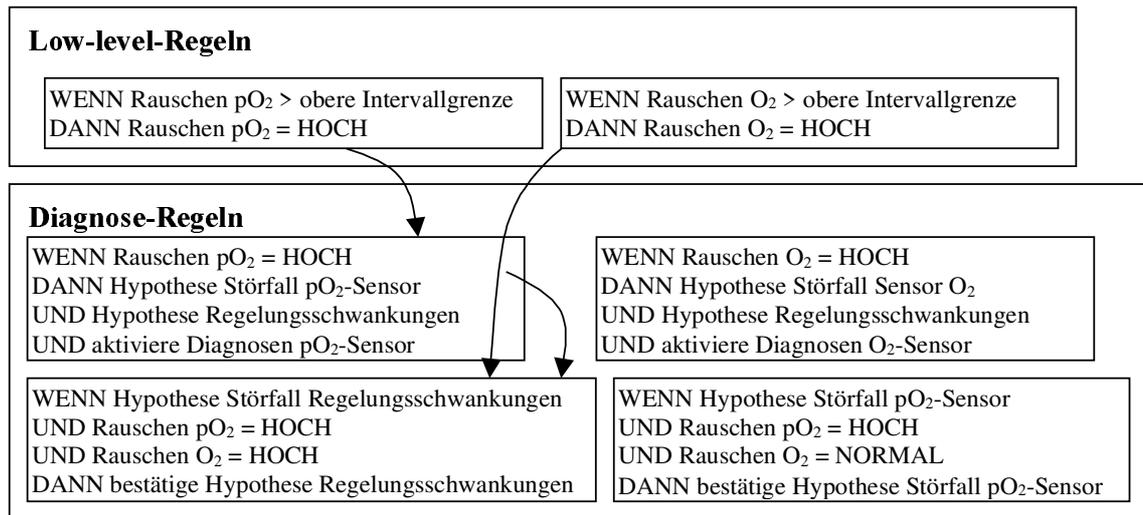


Abbildung 4-51 Schematische Darstellung der Regeln zur Erkennung einer Regelungsschwankung von pO_2

Mit dem Expertensystem konnten im On-line-Einsatz sowie in der Simulation zahlreiche Störfälle und spezielle Prozeßzustände erkannt werden. Ein wichtiges Ziel bei diesen Untersuchungen war möglichst allgemeine Merkmale für die Erkennung der Zustände zu finden. Durch eine allgemeine Formulierung des Wissens – unter Verwendung der von dem Expertensystem erzeugten symbolischen Merkmale – ist es möglich, die notwendigen Änderungen bei der Anpassung des Expertensystems an einen neuen Prozeß zu verringern. Im günstigsten Fall sollte die Anpassung der Konfidenzbereiche ausreichen, um bei einem neuen Prozeß zumindest die Symptome der häufigsten Störfälle zu erkennen. Die größte Schwierigkeit bei der Anpassung stellt die Erkennung des Prozeßzustands dar, da dieser meistens sehr spezifisch ist. Das hierfür notwendige heuristische Modell muß immer neu erstellt und somit müssen auch die Regeln erneut vom Benutzer formuliert werden. Kommen in den verschiedenen Prozeßmodellen aber vergleichbare Zustände – wie z. B. eine Diauxie – vor, so kann zumindest die gleiche Strategie zur Erkennung genutzt werden. Die nachfolgenden Tabelle 4-16, 4-17, 4-18 und 4-19 stellen die bei dem Beispielprozeß untersuchten Zustände und Störfälle dar. Die Tabellen enthalten nicht

das vollständige Wissen, sondern sollen einen Überblick geben, in wie weit das Wissen zur Erkennung dieser Zustände allgemein formuliert werden kann.

Tabelle 4-16 Zustände des biologischen Systems, die mit dem Expertensystem an dem Beispielprozess untersucht wurden

Zustand	Bemerkungen	Objekt	Attribut	Wert
Kontamination	Bakter. Infektion	Prozeß	Phase	1
		CO ₂	Wert	HOCH
		O ₂	Wert	NIEDRIG
		PO ₂	Wert	NIEDRIG
kurze Lag-Phase		Prozeß	Phase	1
		CO ₂	Wert	HOCH
		O ₂	Wert	NIEDRIG
		Prozeß	Verlauf	NORMAL
lange Lag-Phase		Prozeß	Phase	1
		CO ₂	Wert	NIEDRIG
		O ₂	Wert	HOCH
		Prozeß	Verlauf	NORMAL
Limitierung	Keine pO ₂ -Regelung	CO ₂	Passendes Referenzmuster	MAX
		O ₂	Passendes Referenzmuster	MIN
		PO ₂	Passendes Referenzmuster	MIN
Limitierung	pO ₂ -Regelung (Sauerstoffzudosierung)	CO ₂	Passendes Referenzmuster	MAX

Tabelle 4-16 beschreibt unterschiedliche biologische Erscheinungen und ihre Auswirkungen auf den Verlauf wichtiger Prozeßgrößen. Die Auswirkungen hängen stark von der spezifischen Prozeßführung ab. So wirken sich Limitierungen auf den Verlauf der CO₂-, O₂- und pO₂-Werte aus. Wenn hingegen der pO₂-Wert durch Sauerstoffzudosierung zur Luft der Begasung geregelt wird, so wirkt sich eine Limitierung nur auf den Verlauf der CO₂-Werte aus. Tabelle-4-17 beschreibt anormale Prozeßzustände die allgemein formuliert werden können. Zum einen handelt es sich dabei um Richtwerte, die allgemein für Bioprosesse unter bestimmten Regulationsbedingungen festgelegt werden können. So darf der pO₂-Wert – für den Fall das er nicht geregelt wird – beispielsweise nicht 100 % überschreiten. Zum anderen handelt es sich um das ungewöhnliche Verläufe von Prozeßgrößen, die mit der allgemeinen Vorstellung von physikalisch-chemischen Wirkungen nicht erklärt werden können. So sollte der pH-Wert nicht steigen, wenn Base zu dem Prozeß zugegeben wird.

Tabelle 4-17 Zusammenhänge zwischen einzelnen Größen, die mit dem Expertensystem an dem Beispielprozeß untersucht wurden

Zusammenhang	Bemerkungen	Objekt	Attribut	Wert
Ammoniumverbrauch	Wenn pH nicht geregelt wird	pH	TREND	FALLEND
Ammoniumverbrauch	Wenn pH geregelt wird	Base	Zugabe	AN
Säuren Produktion	Wenn pH nicht geregelt	pH	TREND	STEIGEND
Säuren Produktion	Wenn pH geregelt	Säure	Zugabe	AN
Verlauf CO ₂ /O ₂ ungewöhnlich	Keine pO ₂ -Regelung über Sauerstoffzudosierung	CO ₂	TREND	STEIGEND
		O ₂	TREND	STEIGEND
Verlauf CO ₂ /O ₂ ungewöhnlich	Keine pO ₂ -Regelung über Sauerstoffzudosierung	CO ₂	TREND	FALLEND
		O ₂	TREND	FALLEND
Wert O ₂	Keine pO ₂ -Regelung über Sauerstoffzudosierung	O ₂	Wert	>25 %
Wert pO ₂		pO ₂	Wert	>100 %
Verlauf O ₂ /pO ₂ ungewöhnlich	Keine pO ₂ -Regelung über Sauerstoffzudosierung	O ₂	TREND	STEIGEND
		pO ₂	TREND	FALLEND
Verlauf O ₂ /pO ₂ ungewöhnlich	Keine pO ₂ -Regelung über Sauerstoffzudosierung	O ₂	TREND	FALLEND
		pO ₂	TREND	STEIGEND

Tabelle 4-18 Zusammenhänge, die durch Störfälle entstehen und die mit dem Expertensystem an dem Beispielprozeß untersucht wurden

Störung	Bemerkungen	Objekt	Attribut	Wert
Sensorstörung	Allgemeine Formulierung	Meßwert	Wert	HOCH
		alle anderen Prozeßgrößen	Wert	NORMAL
Sensorstörung	Allgemeine Formulierung	Meßwert	Wert	NIEDRIG
		alle anderen Prozeßgrößen	Wert	NORMAL
Sensorausfall	Schnelle Änderung von der Meßgröße und Verlassen des Konfidenzbereichs	Meßwert	Meßwertabstand	HOCH
		Meßwert	Wert	HOCH
Sensorausfall	Schnelle Änderung von der Meßgröße und Verlassen des Konfidenzbereichs	Meßwert	Meßwertabstand	HOCH
		Meßwert	Wert	NIEDRIG
Spikes	Kurzfristige schnelle Änderung	Meßwert	Meßwertabstand	HOCH
		Meßwert	Wert	NORMAL
Rauschen		Meßwert	Rauschfaktor	HOCH
Sensordrift		Meßwert	TREND	HOCH
		Meßwert	Wert	HOCH
Sensordrift		Meßwert	TREND	NIEDRIG
		Meßwert	Wert	NIEDRIG

In der Tabelle 4-18 handelt es sich um eine allgemeine Formulierung der Merkmale, die für die Erkennung von Störungen genutzt werden können. So kann bei einer Sensorstörung die möglichen Störungsursachen schnell eingeschränkt werden, wenn man zunächst überprüft, ob auch unnormale Verläufe bei anderen Prozeßgrößen erkannt werden. Ist dies nicht der Fall, so kann in der Regel darauf geschlossen werden, daß eine Sensorstörung vorliegt. In solch einem Fall könnten dann sensorspezifische Regeln zwecks einer genaueren Diagnose aktiviert werden.

Wird eine Prozeßgröße geregelt, so stellen sich die Verhältnisse komplizierte dar, da immer Zusammenhänge zwischen Stell- und Regelgröße bestehen, die allgemeine Formulierungen erschweren. In Tabelle 4-19 sind besondere Zusammenhänge aufgelistet, an denen sich unter bestimmten Bedingungen Störungen der Regelung erkennen lassen.

Tabelle 4-19 Störfälle, die an geregelten Größen vorkommen und die mit dem Expertensystem an dem Beispielprozeß untersucht wurden

Störung	Objekt	Attribut	Wert
Einschwingen des pO ₂ -Reglers	pO ₂	Rauschfaktor	HOCH
	O ₂	Rauschfaktor	HOCH
Fehler pH-Regelung	pH	Wert	NIEDRIG
	Base	Zugabe	AN
Fehler pH-Regelung	pH	Wert	NIEDRIG
	Base	Zugabe	AUS
Fehler pH-Regelung	pH	Wert	NIEDRIG
	Säure	Zugabe	AUS
Fehler pH-Regelung	pH	Wert	NIEDRIG
	Säure	Zugabe	AN
Fehler Rührerdrehzahl	RPM	Wert	HOCH
	pO ₂	Wert	HOCH
Fehler Rührerdrehzahl	RPM	Wert	NIEDRIG
	pO ₂	Wert	NIEDRIG

4.5.2 Anpassung des Expertensystems

Das entwickelte des Expertensystem soll nicht nur an einem speziellen Bioprozeß einsetzbar sein, sondern es sollte die Möglichkeit geschaffen werden, das Expertensystem flexibel an neue Prozesse anpassen zu können. Im folgenden wurde untersucht, inwieweit das Expertensystem für die Überwachung eines anderen Prozesses eingesetzt werden kann. Ein Aspekt der Unter-

suchung war, welche Teile der Wissensbasis für die Überwachung eines anderen Prozesses verändert werden mußten.

Die Vorgehensweise, bei der das Expertensystem an einen anderen Prozeß angepaßt wird, beinhaltet folgende Schritte:

- Erstellung und Konfiguration der Überwachungsobjekte,
- Erstellung und Konfiguration von Objekten zur Verwaltung historischer Daten,
- Durchführung von Datenanalysen zur Akquisition von prozeßspezifischem Wissen,
- Akquisition von methodischem Wissen zur Überwachung des Prozeßzustands und von Störfällen.

Zur Durchführung dieser Schritte wurden die in Kapitel 4.4.5.4 vorgestellten Werkzeuge zur Wissensakquisition verwendet. Dabei wurde die Anpassung des Expertensystems für Daten einer Hefe-Kultivierung untersucht. Im folgenden soll an einigen Beispielen die erhaltenen Ergebnisse näher beschrieben werden.

Kultivierung von Hefe [Kirchner, 1997]

Der untersuchte Prozeß unterschied sich in mehreren Punkten von dem Beispielprozeß, für den das Expertensystem entwickelt wurde. Zum einen wurden nur die wichtigsten Prozeßgrößen O_2 , CO_2 , pO_2 , pH, T und die Rührerdrehzahl on line gemessen und von dem Prozeßleitsystem als drei Minutenmittelwerte zur Verfügung gestellt. Durch diese relativ kleine Meßfrequenz wurden Einflüsse durch Rauschen bereits heraus gefiltert. Für die Untersuchung stand keine On-line-Anbindung zur Verfügung. Somit wurden alle Untersuchungen als Simulation durchgeführt. Hierfür standen drei Kultivierungen zur Verfügung, deren Verlauf sehr reproduzierbar waren. Mit Hilfe zweier Datensätze wurde das Expertensystem trainiert und im Simulationslauf der Verlauf der dritten Kultivierung untersucht. An den untersuchten Prozeßverläufen traten zwei Störfälle auf. Zum einen eine verkürzte Lag-Phase und zum anderen ein falsch kalibrierter pO_2 -Sensor. Das Expertensystem sollte nach der Anpassung, ohne Veränderung des Konzepts und nur mit den schon implementierten Methoden in der Lage sein, den Prozeßzustand zu ermitteln und diese Störfälle zu erkennen.

Überwachung und Ermittlung des Prozeßzustands

Um mit dem Expertensystem – nach dem gleichen Konzept wie beim Beispielprozeß – den Prozeßzustand des neuen Prozeßverlaufs ermitteln und überwachen zu können, müssen folgende Schritte durchgeführt werden:

- Allgemeine Formulierung eines heuristischen Prozeßmodells,
- Off-line-Mustererkennung und Datenanalyse vorhandener Daten,
- Analyse des heuristischen Prozeßmodells und der Ergebnissen der Mustererkennung und
- Formulierung von Regeln zur Überwachung des Prozeßverlaufs.

Im ersten Schritt müssen mit Hilfe eines Experten, die Prozeßdaten gesichtet und ein heuristisches Prozeßmodell erstellt werden.

Dieses unterscheidet sich von dem Prozeßmodell, das für die Überwachung des Beispielprozesses eingesetzt wurde. Im folgenden soll auf die wesentlichen Unterschiede und ihre Auswirkung auf das Prozeßmodell eingegangen werden.

Betrachtet man den in Abbildung 4-52 dargestellten Verlauf des pH-Wert, so stellt man fest, daß dieser über die gesamte Kultivierungsdauer durch eine Regelung konstant auf pH 4 gehalten wird. Demnach äußern sich keine – durch die Aktivität der Mikroorganismen hervorgerufenen – Prozeßphänomene direkt in dem zeitlichen Verlauf dieser Größe. Eine Abweichung von diesem konstanten Verlauf deutet hingegen direkt auf einen Störfall hin.

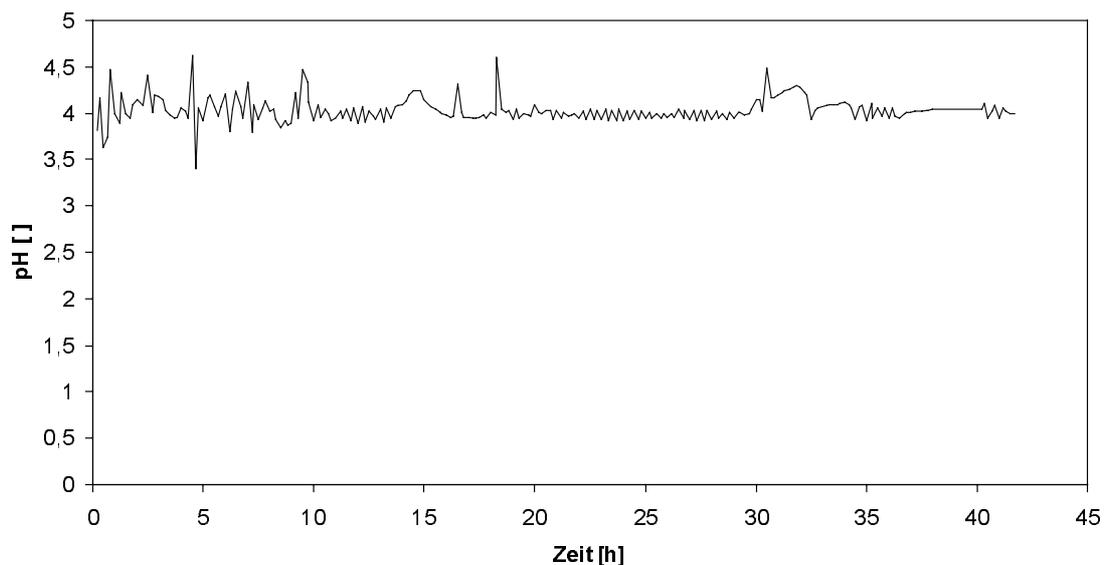


Abbildung 4-52 Verlauf der pH-Werte einer repräsentativen Hefe-Kultivierung

Der pO_2 -Wert wird bei dem betrachteten Prozeß nicht geregelt. Im Gegensatz zu der Pilz-Kultivierung äußert sich dies darin, daß der pO_2 -Wert über den nahezu ganzen Verlauf der Kultivierung kontinuierlich bis auf 50 % fällt, wie in der Abbildung 4-53 deutlich wird. Der am Ende der Kultivierung zu verzeichnende starke Anstieg der pO_2 -Werte ist auf das Absterben der Mikroorganismen zurückzuführen. In Abbildung 4-53 sind die Ergebnisse der Mustererkennung für diesen repräsentativen Kultivierungsverlauf dargestellt. Hier wird für die ersten 30 Stunden der Kultivierungszeit das abfallende Referenzmuster (DECR) zugeordnet, um den Verlauf dieser Prozeßgröße zu charakterisieren. Der Anstieg der pO_2 -Werte nach 30 Stunden führt zu einem Musterwechsel. Es wird nun das steigende Referenzmuster (INCR) als charakteristisches Muster ermittelt. Das Auftreten dieses Merkmals kann demnach zur Erkennung der Absterbephase verwendet werden.

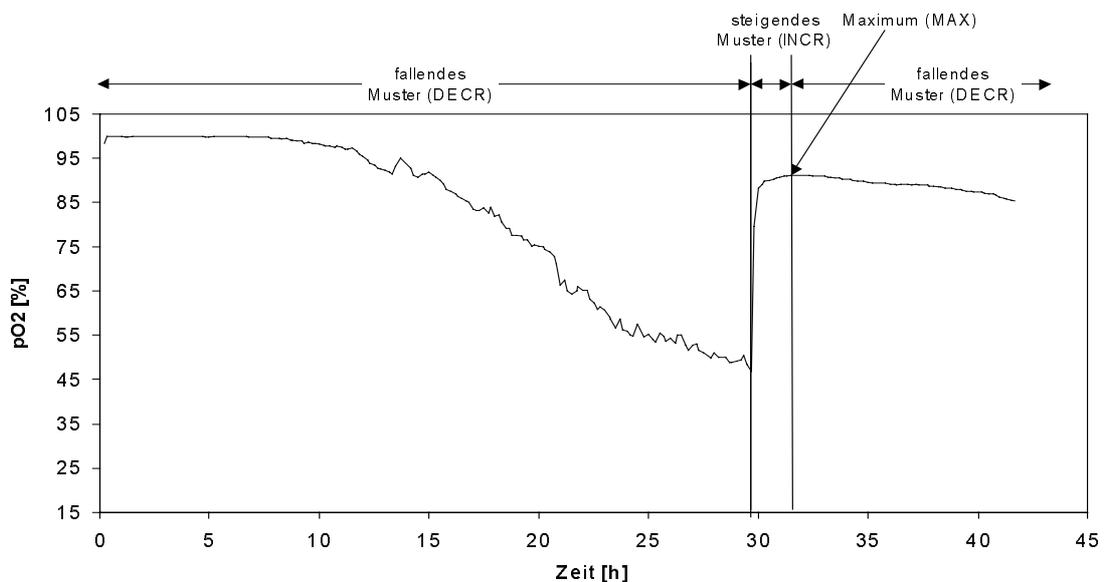


Abbildung 4-53 Verlauf der pO_2 -Werte einer repräsentativen Hefe-Kultivierung sowie die Ergebnisse der Mustererkennung

Der Verlauf des O_2 im Abgas ist, da keine Regelung der pO_2 -Werte stattfindet, durch das biologische Verhalten der Mikroorganismen gekennzeichnet. Zu Beginn der Kultivierung findet durch das starke Wachstum und dem entsprechend großen Sauerstoffverbrauch durch die Mikroorganismen ein kontinuierlicher Abfall des O_2 Gehalts auf 19 % in der Abluft statt. Der nach etwa 13 Stunden der Kultivierung zu verzeichnende Anstieg der O_2 -Werte im Abgas ist auf eine Limitierung der leicht umzusetzenden Stoffe, wie Glucose, zurückzuführen. Nach einer kurzen Zeit – in der sich der Stoffwechsel der Mikroorganismen auf ein anderes Substrat

umstellt – findet erneut eine Zunahme der Stoffwechselaktivität statt, die zu einem weiteren Absinken des O_2 im Abgas führt. Der O_2 im Abgas durchläuft demnach zu Beginn der Limitierung ein Minimum um dann nach der Stoffwechselumstellung ein Maximum zu durchlaufen. Bei Betrachtung der von der von der Mustererkennung für den Kultivierungsverlauf des O_2 im Abgas ermittelten charakteristischen Referenzmuster, die in Abbildung 4-54 dargestellt sind, so ist erkennbar, daß sich dieser charakteristische Verlauf der Prozeßgröße in der Abfolge abfallendes Referenzmuster (DECR), Referenzmuster Minimum (MIN), ansteigendes Referenzmuster (INCR), Referenzmuster Maximum (MAX) und abfallendes Referenzmuster (DECR) widerspiegelt. Mit den von der Mustererkennung ermittelten Merkmale ist es also möglich, den charakteristischen Verlauf des O_2 im Abgas zu überwachen und die Limitierung zu erkennen. Die im Verlauf der Kultivierung auftretende Absterbephase wirkt sich auf den Verlauf des O_2 im Abgas dadurch aus, daß dieser stark ansteigt. Diese Erscheinung bewirkt, daß die Mustererkennung ab 30 Stunden der Kultivierung das ansteigenden Referenzmuster (INCR) als passendes Muster ermittelt.

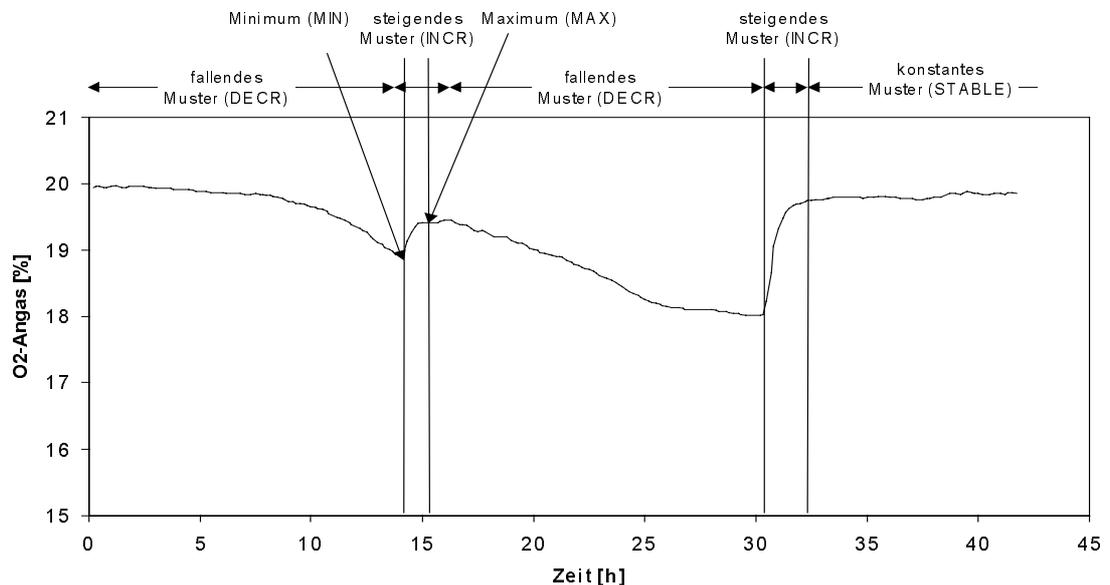


Abbildung 4-54 Verlauf des O_2 im Abgas einer repräsentativen Hefe-Kultivierung sowie die Ergebnisse der Mustererkennung

Analog zu der Pilzkultivierung führt auch das anfängliche Wachstum der Hefe-Kultivierung zu einem stetigen Anstieg des CO_2 im Abgas. Die nach ca. 13 Stunden auftretende Limitierung führt dazu, daß der CO_2 -Wert ein Maximum durchläuft und von etwa 3 % auf 0,5 % sinkt.

Nach dem sich der Stoffwechsel der Mikroorganismen umgestellt hat, steigt der CO₂-Wert stetig bis ca. 28 Stunden ein konstanter Wert von 1 % gehalten wird. Die anschließende Absterbephase bewirkt, daß der CO₂ Wert nach ca. 30 Stunden auf unter 0,2 % absinkt. In der Abbildung 4-55 sind die Ergebnisse der Mustererkennung für den Verlauf der CO₂-Werte dargestellt. Man sieht, daß jede Prozeßerscheinung mit einer Änderung des zugeordneten Referenzmusters assoziiert werden kann. Diese Merkmale können demnach für die Überwachung des Prozeßablaufs verwendet werden.

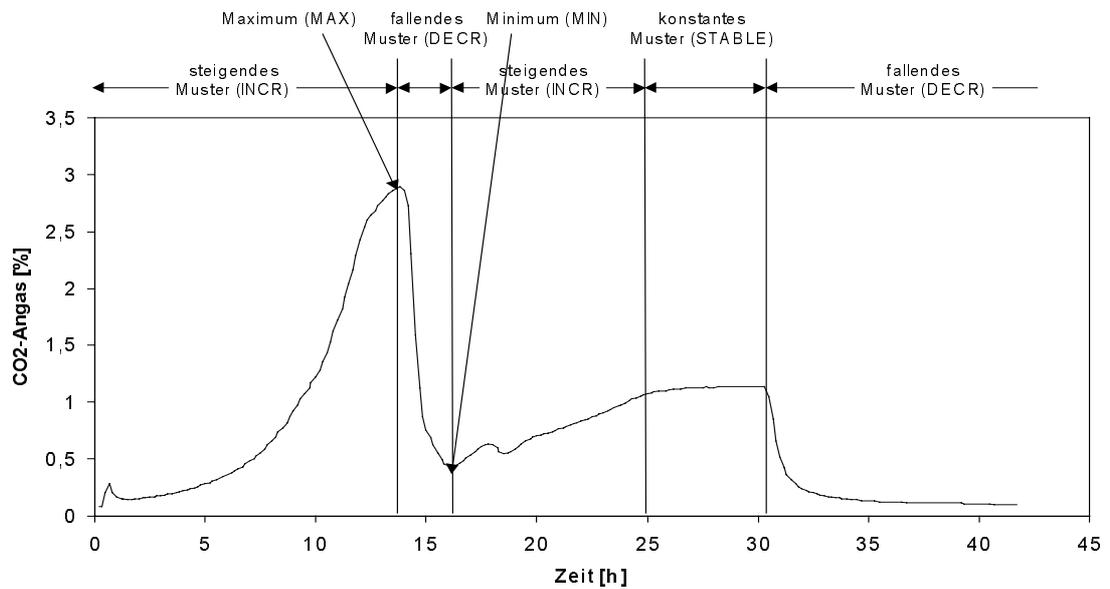


Abbildung 4-55 Verlauf des CO₂ im Abgas einer repräsentativen Hefe-Kultivierung sowie die Ergebnisse der Mustererkennung

Tabelle 4-20 zeigt die aus dem Prozeßmodell und den Ergebnissen der Mustererkennung gewonnenen Zusammenhänge, die von dem Expertensystem zur Prozeßüberwachung verwendet werden können.

Tabelle 4-20 Unterteilung des Prozeßverlaufs in Phasenabschnitte mit Hilfe charakteristischer Merkmale der Mustererkennung

	CO ₂	O ₂	PO ₂
Phase 1	Erste Wachstumsphase		
Norm. Muster	INCR	DECR	DECR
Wechsel in Phase	-----	-----	-----
Muster Ende	MAX	MIN	-----
Phase 2	Diauxie		
Norm. Muster	DECR	INCR	DECR
Wechsel in Phase	-----	-----	-----
Muster Ende	MIN	MAX	-----

Fortsetzung	Tabelle 4-20		
Phase 3	Zweite Wachstumsphase		
Norm. Muster	INCR	DECR	DECR
Wechsel in Phase	-----	-----	-----
Muster Ende	STABLE	-----	-----
Phase 4	Stationäre Phase		
Norm. Muster	STABLE	DECR	DECR
Wechsel in Phase	-----	-----	-----
Muster Ende	DECR	INCR	INCR
Phase 5	Absterbephase		
Norm. Muster	DECR	INCR	INCR
Wechsel in Phase	-----	-----	-----
Muster Ende	-----	-----	-----

Die entsprechenden Regeln, die diese Zusammenhänge für die Überwachung nutzen sind in der Abbildung 4-56 dargestellt.

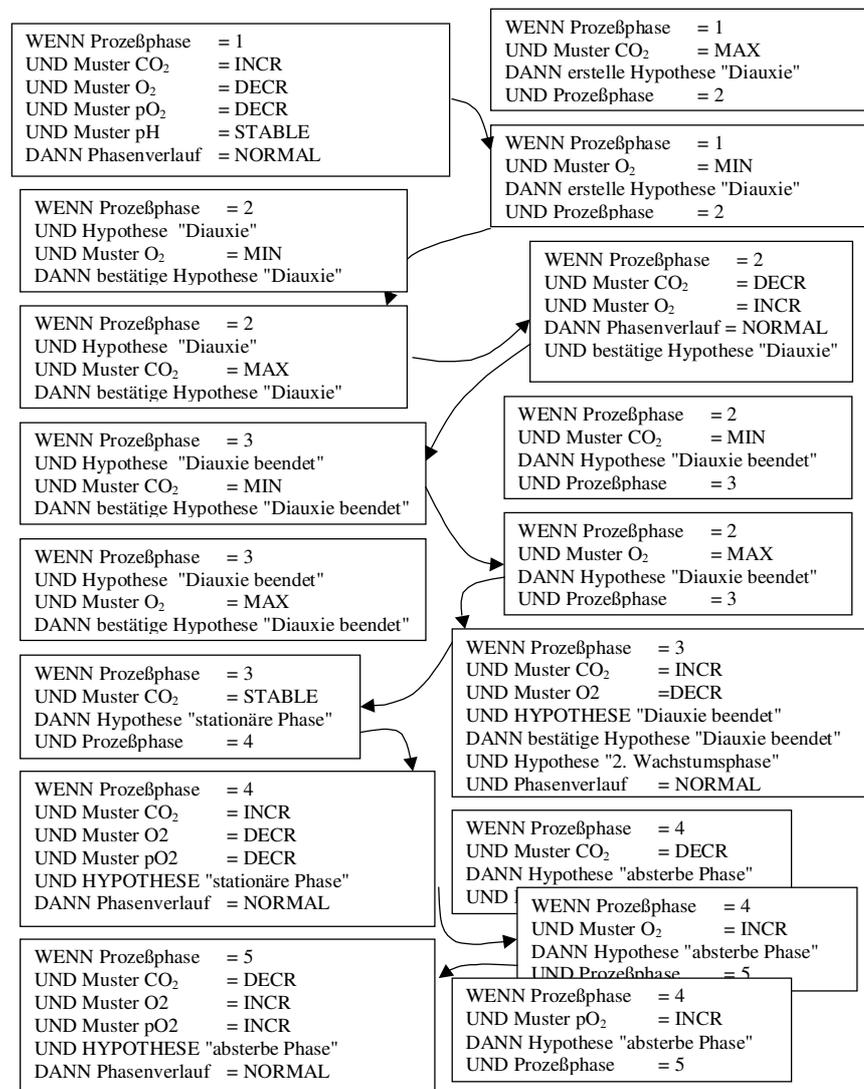


Abbildung 4-56 Schematische Darstellung einiger Regeln zur Überwachung des Phasenverlaufs der Kultivierung

Erkennung einer verlängerten Lag-Phase

Zu den Aufgaben des Expertensystems zählen neben der Bestimmung des Prozeßzustands auch die Erkennung von Abweichungen im Prozeßverlauf. Eine Abweichung die beim Prozeßverlauf der Hefe-Kultivierungen auftrat war z. B. eine verkürzte Lag-Phase. Dieser leicht veränderte Prozeßverlauf führte dazu, daß der gesamte Verlauf und die dazu gehörigen Muster zeitlich früher auftraten. Bei der Musteranalyse der Referenzkultivierungen werden Informationen über das zeitliche Auftreten der Muster und der Absolutwert am Anfang und am Ende eines charakteristischen Signalabschnitts berechnet und als Faktenwissen im Expertensystem gespeichert.

Dieses Wissen über den erwarteten zeitlichen Verlauf der Kultivierung kann genutzt werden, um derartige Abweichungen zu erkennen. Bei dem betrachteten Beispiel führt diese Abweichung zu folgenden Symptomen, die frühzeitig von dem System erkannt werden:

- die CO₂-Werte überschreiten den Konfidenzbereich in der ersten Wachstumsphase,
- die O₂-Werte unterschreiten den Konfidenzbereich in der ersten Wachstumsphase.

Werden diese Symptome erkannt, kann zunächst nur eine Hypothese aufgestellt werden, daß die Ursache für diese Symptome eine verkürzte Lag-Phase ist. Die Hypothese kann erst dann bestätigt werden, wenn Merkmale erkannt werden, die auf einen – bis auf die zeitliche Verschiebung – ‚normalen‘ Phasenverlauf hinweisen:

- ein Maximum der CO₂-Werte
- und ein Minimum der O₂-Werte.

Nachdem die Hypothese „verkürzte Lag-Phase“ formuliert wurde, wird die Größenordnung der zeitlichen Verschiebung gegenüber den Referenzkultivierungen berechnet. Hierfür wird für den zeitlichen Verlauf der CO₂-Werte in der Wachstumsphase folgender Zusammenhang angenommen:

$$CO_2(t) = Ke^{mt} \quad \text{Formel 4-6}$$

Nach Umformen der Gleichung erhält man:

$$\ln CO_2(t) = mt + \ln K \quad \text{Formel 4-7}$$

Die Formel 4-7 entspricht einer linearen Gleichung mit 2 Unbekannten wie in Formel 4-8 dargestellt, so daß die Konstanten **K** und **m** nach linearer Regression von dem natürlichen Logarithmus der CO₂-Werte gegen die Zeit erhalten werden:

$$y(t) = mt + b$$

$$\text{mit } y(t) = CO_2(t)$$

$$\text{und } b = \ln K$$

Formel 4-8

Zur Bestimmung wie groß die zeitliche Verschiebung der Kultivierung ist, wird nach Formel 4-9 der Zeitpunkt berechnet an dem der CO_2 -Wert das Maximum der Referenzkultivierungen \max_{ref} erreicht:

$$t_{\max ref} = \frac{\ln \max_{ref} - \ln K}{m}$$

Formel 4-9

Eine definitive Bestimmung der zeitlichen Verschiebung kann erst nach Erreichen der 1. Diau-
xie gegeben werden. In diesem Fall muß eine Anpassung der Konfidenzbereiche an dieses ver-
änderte zeitliche Verhalten der Kultivierung vorgenommen werden.

Erweiterung der Wissensbasis für einen neuen Störfall

Wie bei der Implementierung des Wissens zur Überwachung des Prozeßverlaufs findet die Er-
weiterung der Wissensbasis zur Erkennung eines neuen Störfalls in mehreren Schritten statt.
Zunächst müssen Zusammenhänge zwischen der Ursache des Störfalls und Symptomen in den
Meßsignalen gefunden werden. Dabei ist es wichtig, daß dieses heuristische Fehlermodell hin-
reichend von anderen Ursachen ähnlicher Symptome abgegrenzt werden kann. Liegt ein sol-
ches heuristisches Fehlermodell vor, so muß im nächsten Schritt anhand der Datenanalyse
überprüft werden, ob die Symptome von dem Expertensystem erkannt werden. Dies gilt insbe-
sondere, wenn symbolische Eigenschaften des Signalverlaufs in dem Fehlermodell enthalten
sind. Die Beurteilung ob ein Wert als 'HOCH' eingeschätzt wird, hängt im wesentlichen von
dem Konfidenzbereich ab. Überschreitet das numerische Merkmal bei dem Störfall nicht den
Konfidenzbereich, so genügt dieses Symptom allein nicht, um eine Diagnose zu initiieren. Es
muß zur Erkennung des Störfalls mindestens ein numerisches Merkmal seinen Konfidenzbe-
reich verlassen, da ansonsten keine Low-level-Regel eine Diagnose initiieren kann. Sind hinrei-
chend viele Symptome und mindestens eine Low-level-Regel vorhanden, so kann im nächsten
Schritt die Diagnoseregeln – unter Zuhilfenahme der Wissenserwerbskomponente – erstellt wer-
den. Als Kategorie der Diagnoseregeln werden die Kategorien der Low-level-Regeln eingetra-
gen.

Der Störfall der in diesem Beispiel betrachtet wird, ist eine Fehlkalibration des pO_2 -Sensors.
Zur Kalibrierung des pO_2 -Sensors wird normalerweise die obere Meßgrenze in einer vollstän-

dig mit Sauerstoff gesättigten Lösung auf 100 % kalibriert. Erfolgt diese Kalibration in einer nicht mit Sauerstoff gesättigten Lösung so wird während der Kultivierung ein zu hoher Wert angezeigt. Hat der Sensor nur einen Meßbereich von 0 % -100 %, dann wird solange ein Wert von 100 % angezeigt, wie der pO_2 -Wert des Mediums über dem pO_2 -Wert liegt, bei dem kalibriert wurde. Daraus resultieren folgende Symptome, an denen der Störfall sich äußert:

- Der pO_2 -Wert ist HOCH,
- Der Verlauf der pO_2 -Werte wird mit dem Referenzmuster STABLE am besten wiedergegeben,
- der pO_2 -Wert wiederholt sich und
- der pO_2 -Wert ist 100 %.

In Abbildung 4-57 ist der Verlauf der pO_2 -Werte für zwei Hefe-Kultivierungen dargestellt. Man erkennt, wie die Fehlkalibrierung des pO_2 -Sensors sich an den beschriebenen Symptomen äußert.

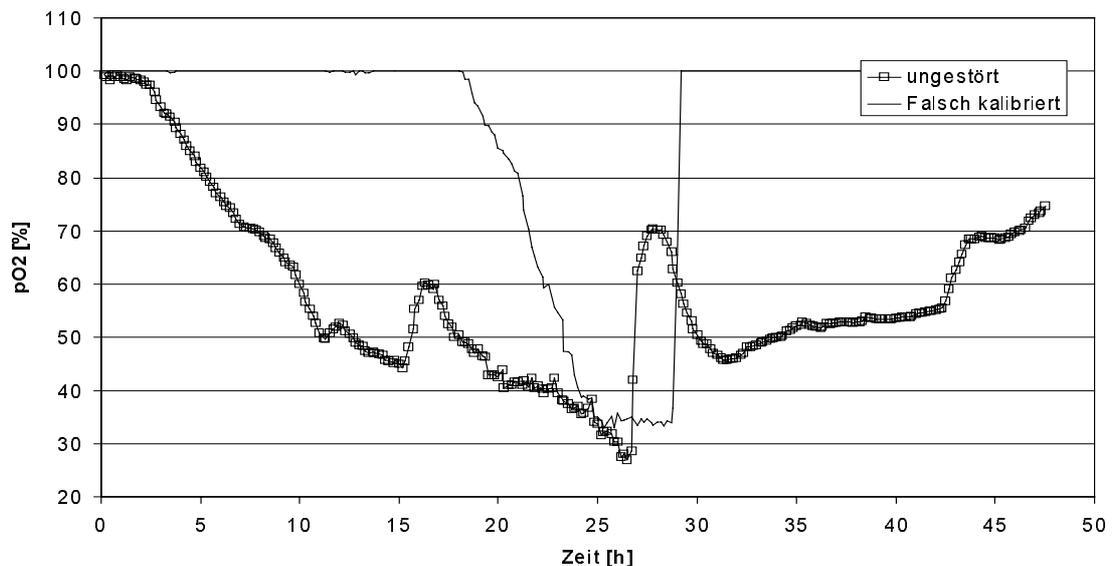


Abbildung 4-57 Darstellung des pO_2 -Verlaufs zweier Hefe-Kultivierungen.

Die ersten drei Symptome werden unabhängig voneinander von Low-level-Regeln erkannt, die schon in dem Expertensystem implementiert sind. Die Low-level-Regeln aktivieren weitere Diagnose-Regeln, die den anormalen Zustand weiter untersuchen. Die Zuordnung, welche Diagnose-Regeln aktiviert werden, erfolgt über die Kategorie der Regel. Tabelle 4-21 zeigt die in dem Expertensystem vorhandenen Kategorien, welche von Low-level-Regeln aktiviert werden, wenn eines der Symptome auftritt.

Tabelle 4-21 Kategorien, die von vorhandenen Low-Level-Regeln aktiviert werden, wenn ein Symptom des Störfalls auftritt

Symptom	Aktivierte Kategorien
PO ₂ -Wert = HOCH	Diagnose-Sensor-defekt, Diagnose-Phasenverlauf
PO ₂ Wertwiederholung = WAHR	Diagnose-Sensor-defekt
PO ₂ Muster = STABLE	Diagnose-Stable

Um den Störfall seiner Ursache zuordnen zu können, muß eine Regel in der Wissensbasis implementiert werden, die das gleichzeitige Auftreten der Symptome überprüft. Darüber hinaus muß der Regel mindestens eine Kategorie zugeordnet werden, damit sie von einer der Low-level-Regeln aktiviert wird. Die sprachlich formulierte Regel lautet:

WENN pO₂-Wert = 100 %
 UND Wertwiederholung pO₂ = WAHR
 UND Phase = 1
 DANN schlußfolgere „pO₂-Sensor falsch kalibriert!“
 UND benachrichtige den Benutzer

Dieser Regel werden die Kategorie „Diagnose-Sensor-defekt“ und „Diagnose-Stable“ zugeordnet sowie der Fokus auf den Objekt „Messwert-pO₂“. Diese Regel läßt sich in dem Expertensystem mittels der Benutzerschnittstelle vollständig dialogorientiert implementieren. Bei der Erstellung der Regel können weitere Eigenschaften festgelegt werden. So bietet es sich an die Regel – nachdem der Störfall festgestellt wurde – zu deaktivieren, um mehrfache Benachrichtigungen des Bedieners zu vermeiden. Abbildung 4-58 zeigt die in G2 mit Hilfe der Benutzerschnittstelle erstellten Regel. An der Regel wird die Notwendigkeit der Wissenserhebungs-komponente deutlich, da es für einen ungeübten Benutzer des Expertensystems schwierig ist, manuell solche Regeln zu erstellen.

a rule	
Options	not invocable via backward chaining, not invocable via forward chaining, may not cause data seeking, may not cause forward chaining
Notes	RULE-XXX-46: OK
User restrictions	none
Names	none
Tracing and breakpoints	default
if the sub-ws-active of the diagnose-regel-ws superior to the workspace of this rule = the symbol true and messwert-po2 = 100.0 and the wert-symbol of messwert-po2 = the symbol wert-wiederholung and the prozessphase of aktueller-prozess = 1 then start message-auf-scroll(this rule, "Prozessphase = 1, pO2-Wert = 100 % und wiederholt sich. pO2-Sensor falsch kalibriert!", "") and start deactivate-rule (the workspace of this rule, 10.0)	
Scan interval	none
Focal classes	none
Focal objects	messwert-po2
Categories	diag-wert-wiederholung, diagnose-stable, diagnose-sensor-defekt
Rule priority	6
Depth first backward chaining precedence	1
Timeout for rule completion	use default

Abbildung 4-58 Eigenschaftstabelle der Regel zur Erkennung des Störfalls

4.6 Zusammenfassung

Die On-line-Überwachung stellt eine große Herausforderung bei der Automatisierung komplexer Prozesse dar. Vordergründiges Ziel solcher Systeme ist u.a. eine erhöhte Betriebssicherheit und damit verbunden geringere Ausfallzeiten der Anlagen. Besonders für die Überwachung ist aber eine große Menge Erfahrungswissen notwendig. Expertensysteme bieten die Möglichkeit, dieses Erfahrungswissen zu speichern und zu verarbeiten. Die Speicherung des fachspezifischen Wissens bietet zusätzliche Vorteile. So findet in einem Expertensystem auch eine Konservierung des Wissens statt und das Expertensystem kann so u. a. zu Schulungszwecken eingesetzt werden. Die Bioprozeßtechnik, die das Wissen mehrerer Fachgebiete umfaßt (u. a. Verfahrenstechnik, Meßtechnik, Chemie und Mikrobiologie), bietet vielfältige Möglichkeiten für ihren Einsatz.

Das Ziel dieser Arbeit war die Entwicklung eines verallgemeinerbaren Expertensystems zur Überwachung von Bioprozessen. Dazu wurde unter Verwendung der Expertensystemshell G2

ein hybrides System entwickelt, das ein numerisches und ein wissensverarbeitendes Modul enthält. Diese hybride Kombination von numerischen und wissensverarbeitenden Methoden hat sich für die On-line-Überwachung von Bioprozessen bestens bewährt. Um innerhalb des wissensbasierten Moduls eine dem menschlichen Experten gerechte Terminologie zu verwenden, werden zunächst die numerischen Merkmale der On-line-Prozeßgrößen anhand von Richtwerten, Konfidenzbereichen und einer integrierten Mustererkennung klassifiziert. Die Resultate dieser Klassifikation werden in dem System mit Symbolen repräsentiert. Low-level-Regeln reagieren auf Abweichungen vom normalen Zustand und aktivieren Diagnose-Regeln, die den anormalen Zustand näher untersuchen und Ursachen ergründen sollen. Durch diese Fokussierung konnte die Antwortzeit des Expertensystems erheblich verkürzt und ein Echtzeitverhalten garantiert werden.

In dem System werden die Schlußfolgerungen anhand von Meldungen auf einer Bildlaufleiste sowie anhand von Graphiken erläutert. Durch eine zusätzliche – auf dem Internet basierende – Visualisierung der Schlußfolgerungen und Graphiken konnte die Verfügbarkeit des Ergebnisse des Expertensystems bedeutend erhöht werden.

Das Expertensystem wurde sowohl im On- als auch im Off-line-Betrieb bei der Überwachung von *Acremonium chrysogenum* eingesetzt. Das System war hierbei in der Lage zuverlässig den Prozeßverlauf zu ermitteln und zahlreiche Störfälle zu erkennen.

Für die Anpassung des Expertensystems an andere Prozesse werden dem Benutzer zahlreiche Werkzeuge zur Verfügung gestellt. Mit induktiven Methoden werden in der Off-line-Datenanalyse Richtwerte und Konfidenzbereiche berechnet, die es dem System erlauben die numerischen Merkmale des Systems qualitativ zu beurteilen.

Die integrierte Mustererkennung unterteilt die Signale der Prozeßgrößen in charakteristische Abschnitte. Mit den Merkmalen, welche die Mustererkennung erzeugt, kann zuverlässig der typische Phasenverlauf von Kultivierungen überwacht werden. Hierfür muß aber zunächst vom Bediener ein heuristisches Prozeßmodell mit Hilfe von Regeln in der Wissensbasis implementiert werden. Die Erstellung von Regeln ist in dem Expertensystem vollständig dialoggeführt möglich, so daß ein Benutzer schnell syntaktisch korrekte Regeln erstellen kann. Um die Transparenz der Wissensbasis zu erhöhen, wurde eine Erklärungskomponente geschaffen, die Zusammenhänge zwischen Regeln darstellt und somit dem Bediener hilft, semantisch falsche Regeln innerhalb des komplexen Diagnosenetzwerks zu entdecken.

Das Expertensystem wurde bewußt so konzipiert, daß es unproblematisch an neue Prozesse angepaßt werden kann. Um dieses Konzept zu überprüfen, wurde das Expertensystem für die Überwachung einer Hefe-Kultivierung eingesetzt. Bei diesem Prozeß konnten sowohl der Phasenverlauf als auch neue Störfälle richtig erkannt werden.

5 Anwendung der Ergebnisse der Clusteranalyse zur Überwachung von Bioprocessen

In Kapitel 4 wurde gezeigt, daß sich die Kombination von numerischen und wissensbasierten Methoden zur Überwachung von Bioprocessen nutzen läßt. Eine große Herausforderung ist dabei die maschinelle Klassifizierung der On-line-Daten. Je leistungsfähiger die Klassifizierungsmethode ist, desto souveräner und spezifischer ist auch die Erkennung des Prozeßzustands und damit auch die Überwachung. In diesem Kapitel werden Untersuchungen vorgestellt, bei denen die Ergebnisse der Clusteranalyse zur On-line-Zustandsermittlung und Überwachung von Bioprocessen eingesetzt werden. Zum tieferen Verständnis sollen zunächst die Grundlagen dieser Klassifizierungsmethode näher erläutert werden. Für eine tiefergehende Einführung in die Methode der Clusteranalyse sei hier auf die Arbeiten von Guthke hingewiesen [1991].

5.1 Grundlagen der Clusteranalyse

Für die Klassifikation von Kultivierungsdaten ohne Modellannahmen gilt der Abstand d_{ij} zwischen zwei Kultivierungen i und j eines gegebenen Signals als die Summe der Abweichungen der Prozeßgröße $p_i(k\Delta t)$ zu festen Zeitpunkten ($k\Delta t$).

$$d_{ij} = \frac{1}{n} \sum_{k=1}^n |p_i(k\Delta t) - p_j(k\Delta t)| \quad \text{Formel 5-1}$$

Für die hierarchische Klassifizierung werden agglomerative Methoden mit einfacher, vollständiger und mittlerer Verknüpfung gewählt [Guthke und Roßmann 1991]. Es werden aufeinanderfolgend immer zwei Kultivierungen bzw. Gruppen zu einer neuen, höheren Gruppe zusammen gefaßt, falls ihr Abstand kleiner ist als der Abstand zwischen allen anderen Kombinationen. Als Ergebnis dieser Agglomeration werden Dendogramme erhalten, welche die Kultivierungen in Klassen unterteilen. Jede Klasse i ist hierbei durch den zeitlichen Verlauf des Mittelwerts $c_i(t)$ und der dazu gehörigen Standardabweichung definiert. Abbildung 5-1 zeigt die Verläufe der Mittelwerte dreier Kultivierungsklassen für die Kohlendioxid Produktionsrate

eines industriellen Prozesses zur Produktion eines Antibiotikums. Die Werte sind aus Gründen der Vertraulichkeit normiert. Die gestrichelten Linien zeigen den Bereich der zweifachen Standardabweichung für die Klasse 2.

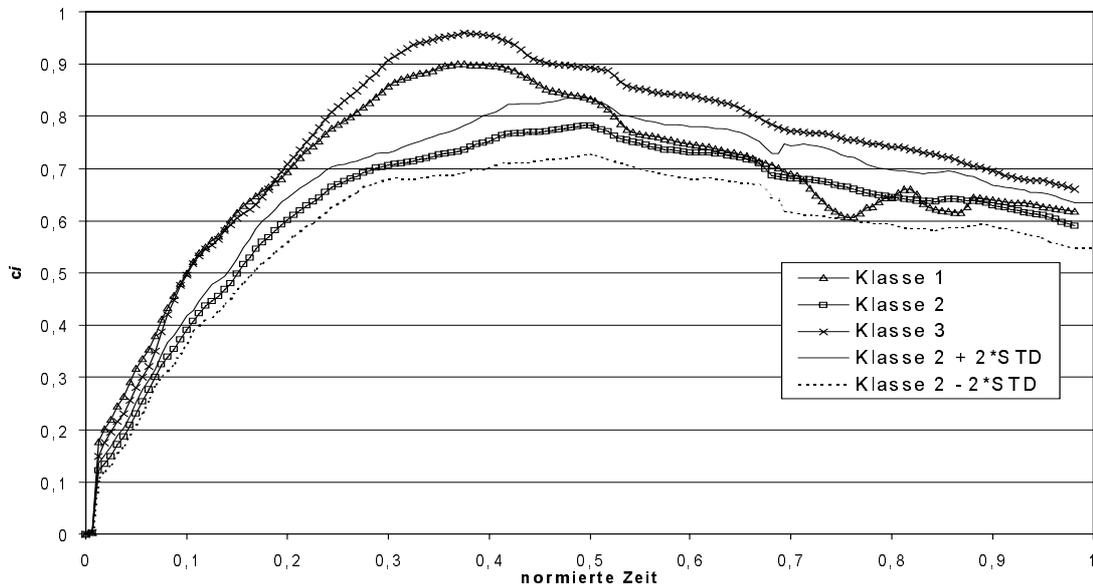


Abbildung 5-1 Mittelwerte (durchgezogene Linien) von drei Klassen sowie den Bereich der zweifachen Standardabweichungen (gestrichelte Linien) für die Klasse 2

5.2 On-line-Klassifizierung

Um on line eine Klassifikation von Kultivierungsdaten vorzunehmen, wird der Abstand d_i zwischen den On-line-Daten p und den definierten Klassen c_i berechnet. Der Verlauf der Daten wird der Klasse zugeordnet, die den kleinsten Abstand zu den On-line-Daten hat.

$$d_i(n\Delta t) = \frac{1}{n} \sum_{k=1}^n |p(k\Delta t) - c_i(k\Delta t)| \tag{Formel 5-2}$$

Es ist wichtig ein Maß festzulegen, nach dem beurteilt werden kann, wann die on line gemessenen Daten keiner Klasse zugeordnet werden können. Um dieses Maß zu bestimmen wird die Abstandsfunktion D_{ij} für eine Klasse gegen alle anderen Klassen berechnet.

$$D_{ij}(n\Delta t) = \frac{1}{n} \sum_{k=1}^n |c_i(k\Delta t) - c_j(k\Delta t)| \tag{Formel 5-3}$$

Abbildung 5-2 zeigt die Abstandsfunktionen $D_{ij}(n\Delta t)$ für jede Kombination der drei Klassen. Der maximal erreichte Wert Abstand D_{max} gibt einen Toleranzbereich für den Verlauf des online ermittelten Abstands an. Ist der Mittelwert des Abstands d_i zwischen den On-line-Daten und jeder Klassen größer als D_{max} , so ist dies ein Hinweis darauf, daß eine neue Klasse definiert werden muß oder ein Störfall vorliegt.

Der Verlauf des Abstands verdeutlicht des Weiteren Ähnlichkeiten zwischen einzelnen Klassen. In diesem Fall ähneln sich Klasse 1 und Klasse 3 am meisten, da ihr Abstand D_{13} am niedrigsten verläuft. Für die On-line-Klassifikation ist es von Bedeutung einen Zeitpunkt festzusetzen, an dem die endgültige Klassenzugehörigkeit ermittelt ist. Der Zeitpunkt des Maximums aller Klassen zeigt sich hierbei als guter Richtwert.

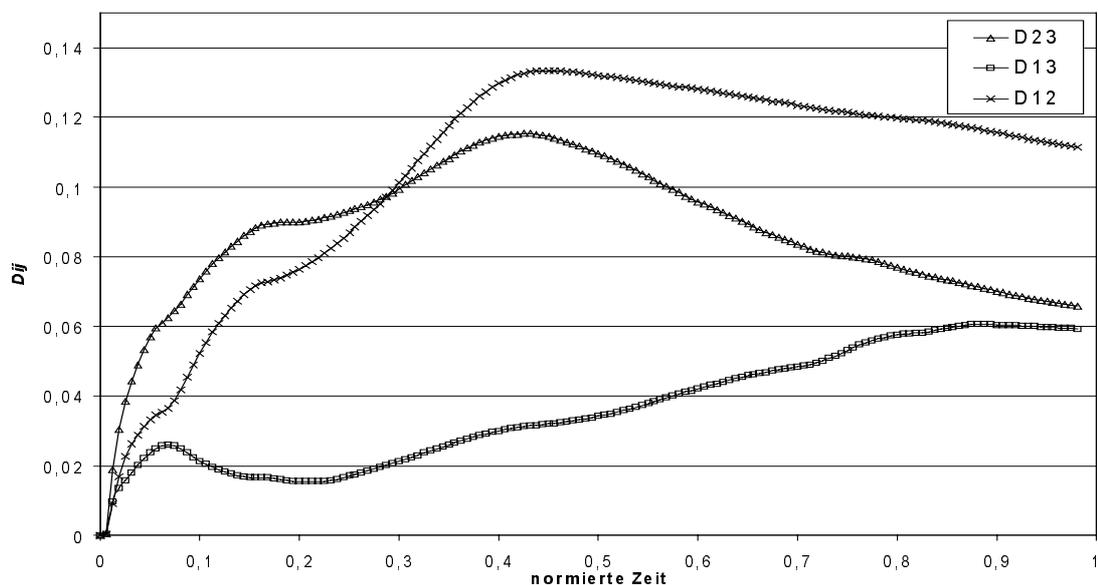


Abbildung 5-2 Abstandsfunktion D_{ij} für jede Kombination von Klassen

5.3 Modellvorhersage

Unter Verwendung des Abstands d_i und den M Klassen, die bei der Clusteranalyse wurden, kann ein Modell für die Vorhersage formuliert werden. Prinzipiell kann der Meßwert $p(t)$ zum gegebenen Zeitpunkt t durch die Multiplikation des Werts einer Klasse zu diesem Zeitpunkt $c_i(t)$ mit einem Faktor $f_i(t)$ ermittelt werden:

$$p(t) = f_i(t)c_i(t)$$

Formel 5-4

Der Faktor $f_i(t)$ kann für eine Vorhersage von p für den Zeitpunkt $t + n\Delta t$ verwendet werden. Um die Information aller Klassen auszunutzen, wird ein Faktor berechnet, der alle Klassen unter Verwendung eines Wichtungsfaktors α_i anteilmäßig berücksichtigt:

$$f(t) = \frac{p(t) \sum_{i=1}^M \alpha_i(t)}{\sum_{i=1}^M \alpha_i(t) c_i(t)} \quad \text{Formel 5-5}$$

Die Vorhersage für den Zeitpunkt $t + \Delta t$ wird aus den Werten der Klasse zum Zeitpunkt $t + \Delta t$ und dem gewichteten Faktor zum Zeitpunkt t berechnet:

$$p_{pred}(t + n\Delta t) = \bar{f}(t) \frac{\sum_{i=1}^M \alpha_i(t) c_i(t + n\Delta t)}{\sum_{i=1}^M \alpha_i(t)} \quad \text{Formel 5-6}$$

Um eine glättende Wirkung auf die Kurve zu erreichen, wird der mittlere Faktor \bar{f} der letzten N Faktoren berechnet. N ist ein freier Parameter und wurde auf einen Wert zwischen 1 und 5 gesetzt.

$$\bar{f}(t) = \sum_{n=0}^{N-1} \frac{f(t - n\Delta t)}{N} \quad \text{Formel 5-7}$$

Das Gewicht α_i wird nach einem heuristischen Ansatz, unter Verwendung des minimalen Abstands d_{min} ermittelt:

$$d_{min} = \min_i \{d_i(t)\} \quad \text{Formel 5-8}$$

Für die Klasse mit dem kleinsten Abstand – demnach wo $d_i = d_{min}$ gilt – wird das Gewicht auf 1 gesetzt. Das Gewicht aller anderen Klassen wird nach Formel 5-9 berechnet:

$$\alpha_i = \begin{cases} 1 - \frac{|d_i - d_{min}|}{d_{min}} & \text{mit } d_i < 2d_{min} \\ 0 & \text{mit } d_i \geq 2d_{min} \end{cases} \quad \text{Formel 5-9}$$

5.4 Implementierung

Unter Berücksichtigung der beschriebenen Vorgehensweise wurde mit der Expertensystemshell G2 (GENSYM, Cambridge, USA) ein Modul erstellt, das die Klassifikationsergebnisse zur Zustandsermittlung und Überwachung von Bioprocessen verwendet. Für jede Prozeßvariable wird in der Wissensbasis ein Objekt angelegt, in dem die Klassendefinitionen als Attribute gespeichert sind. Sobald ein Objekt über die Schnittstelle einen neuen Wert erhält, werden Regeln und Prozeduren aktiviert, die alle Abstandsfunktionen, Gewichte und einen neuen Faktor für das Modell berechnen sowie die Signalverläufe auf Fehler überprüft. Alle Schlußfolgerungen des Moduls werden dem Benutzer auf einer Bildlaufleiste mitgeteilt und der Verlauf aller wichtigen Variablen in einer Graphik dargestellt.

5.5 Ergebnisse

In der Abbildung 5-3 ist der Verlauf der Kohlenstoffdioxidproduktionsrate p einer Kultivierung sowie den drei Klassen c_i dargestellt. In Abbildung 5-4 weist der Abstand d_i zwischen den Daten und den jeweiligen Klassen darauf hin, daß der betrachtete Verlauf der Kohlenstoffproduktionsrate der betrachteten Kultivierung dem Verlauf der Klassen 1 und 3 am ähnlichsten ist.

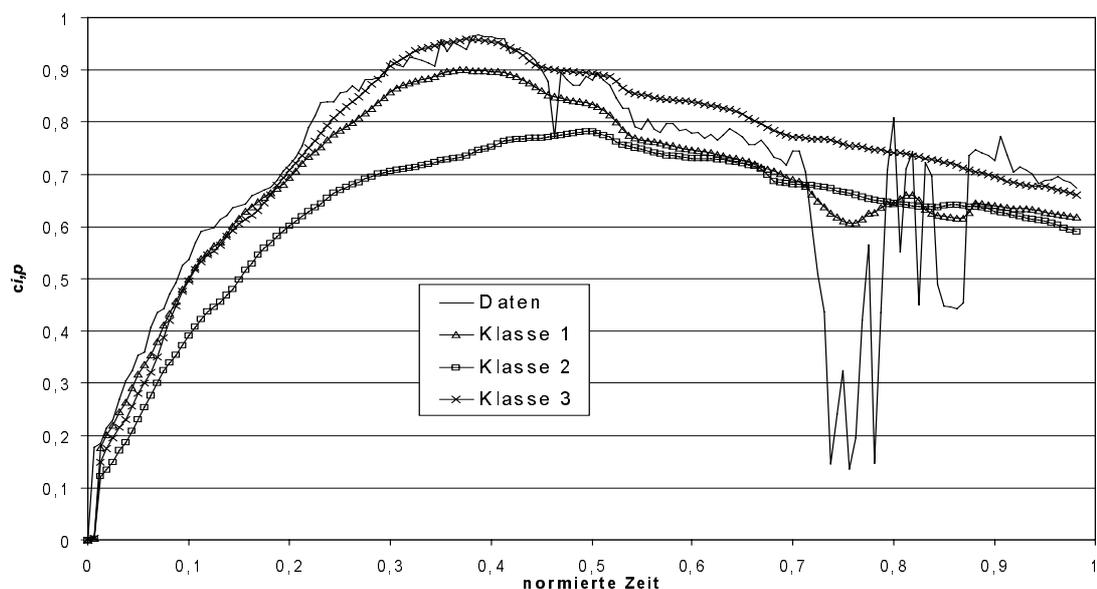


Abbildung 5-3 Kultivierungsdaten der CO₂-Produktionsrate sowie drei Klassen

Die in Abbildung 5-5 dargestellten Gewichte verdeutlichen dies, da das Gewicht α_2 der Klasse 2 auf Null fällt. Dies bedeutet, daß die Klasse 2 nicht um die Klassifikation konkurriert. Bis 0,2 der normierten Zeit kann eine eindeutige Klassifikation nicht vorgenommen werden, da die Abstandsfunktionen d_1 und d_3 der Klassen 1 und 3 sich zu wenig unterscheiden. Zwischen 0,3 und 0,7 der normierten Zeit hat die Klasse 3 den kleinsten Abstand. Klasse 1 bleibt aber konkurrenzfähig, da sein Gewicht nur auf einen Wert von ca. 0,5 sinkt. Nach ca. 0,68 der normierten Zeit tritt ein Störfall auf, der sich durch Spikes in den Daten äußert. Der Störfall wird entdeckt, da die Daten den Bereich der zweifachen Standardabweichung der Klassen überschreitet. Solange die Daten außerhalb des Bereichs verlaufen, werden sie als ungültig angesehen und die Gewichte und Faktoren für das Model werden nicht neu berechnet.

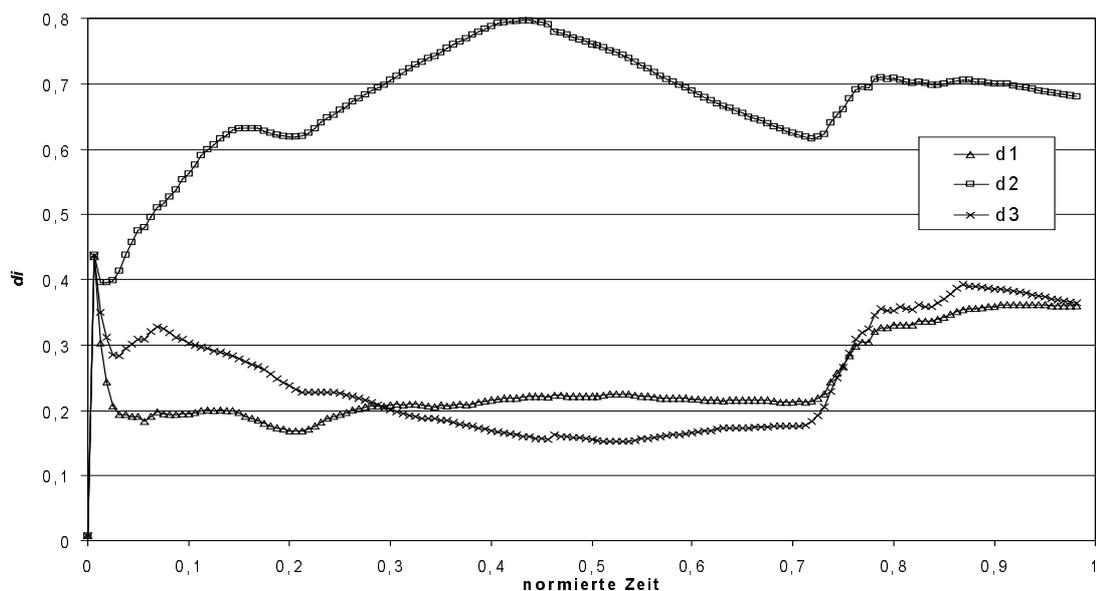


Abbildung 5-4 Verlauf der Abstandsfunktion d_i zwischen den Kultivierungsdaten und den drei Klassen

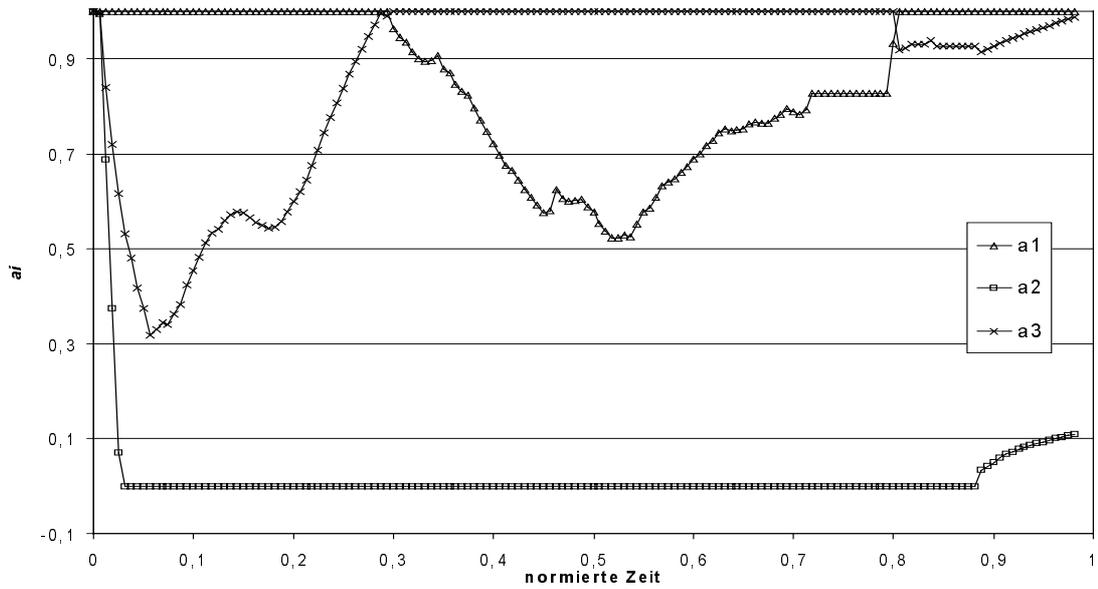


Abbildung 5-5 Gewichtsfunktion a_i für die Klassen 1-3 und den Kultivierungsdaten

Andererseits werden die Abstände neu berechnet und in Abbildung 5-4 verdeutlicht, daß die Folgen der Störung in einer Verminderung des Abstands zu der Klasse 1 resultieren. Dies wird nicht bei der Klassifizierung berücksichtigt, da die letzte Klassifizierung gültig bleibt solange ein Störfall vorliegt. Abbildung 5-6 zeigt den Verlauf der Daten zusammen mit dem vorhergesagten Verlauf. Die Spikes werden stark heraus gefiltert, das normale Rauschen hingegen wird nur leicht geglättet.

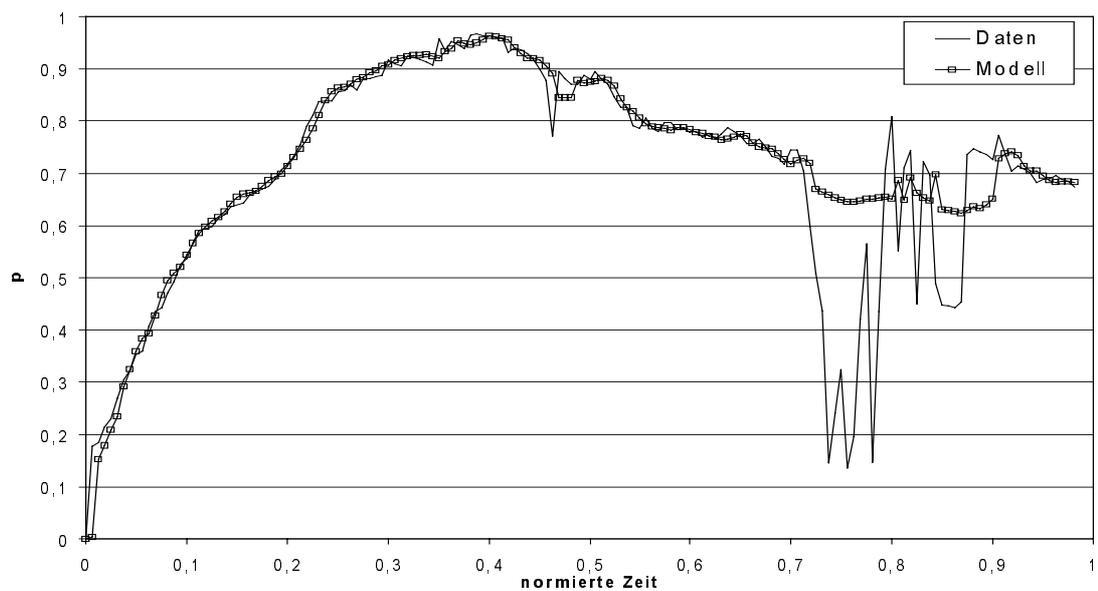


Abbildung 5-6 On-line-Daten und berechnete Vorhersage

Typische Störungen eines Signals resultieren am häufigsten in einzelnen Spikes oder in einer Drift des Signals außerhalb eines Bereichs. Mit den gegebenen Klassendefinitionen kann eine Störung des Signals zuverlässig erkannt werden. Für alle neuen Daten wird überprüft, ob sie sich innerhalb der zweifachen Standardabweichung aller Klassen befinden. Ist dies nicht der Fall, werden sie als fehlerhaft angesehen. Dies ist für Spikes besonders wichtig. Im Fall eines Drifts oder eines vollständigen Ausfalls des Signals, wird der Fehler anhand des hohen mittleren Abstands erkannt. In dem nächsten Beispiel wird ein Totalausfall eines Sensors betrachtet, bei dem der Wert des Signals um 0,5 der normierten Zeit auf ca. 0,1 fällt. Abbildung 5-7 zeigt den Verlauf der Daten. Sobald der Wert unterhalb der zweifachen Standardabweichung liegt, werden sowohl die Gewichte a_i als auch der Faktor f nicht neu berechnet und der Wert der Vorhersage wird mit dem Faktor $f(t_f)$ berechnet. Sobald der Störfall behoben ist und die Daten innerhalb des Bereichs der zweifachen Standardabweichung einer Klasse liegen werden die Gewichte erneut berechnet.

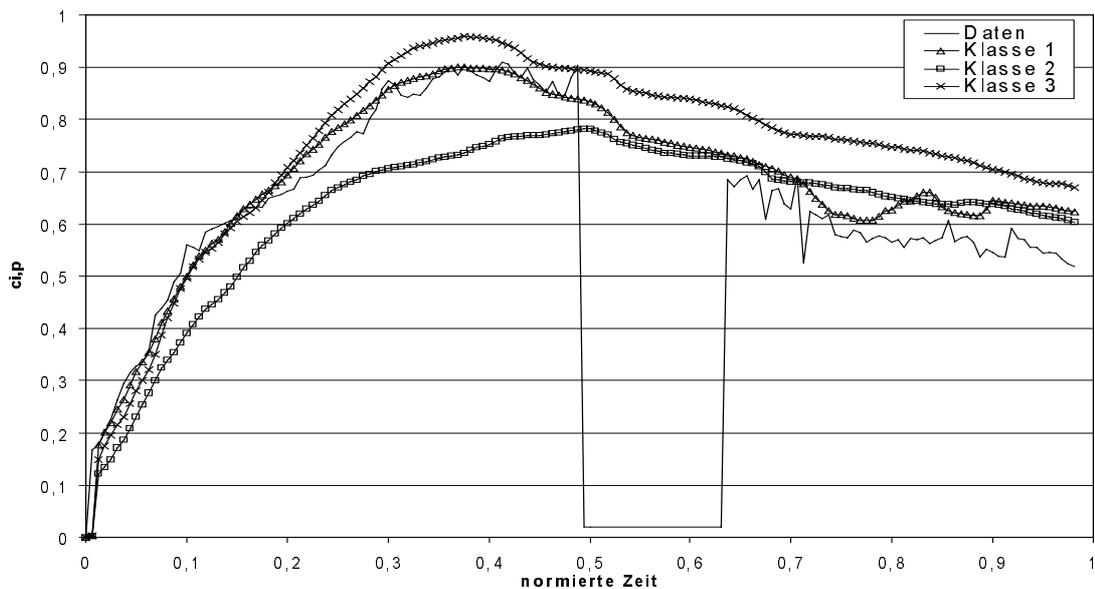


Abbildung 5-7 On-line-Klassifizierung einer Kultivierung mit drei Klassen

Dies wirkt sich in der Abbildung 5-8 durch einen großen Sprung im Verlauf der Gewichte um ca. 0,64 der normierten Zeit aus. In der Abbildung 5-9 sind die Daten zusammen mit der Vorhersage dargestellt. Es ist erkennbar, daß für die Zeitdauer der Störung eine zuverlässige Vorhersage für den Verlauf ermittelt wird.

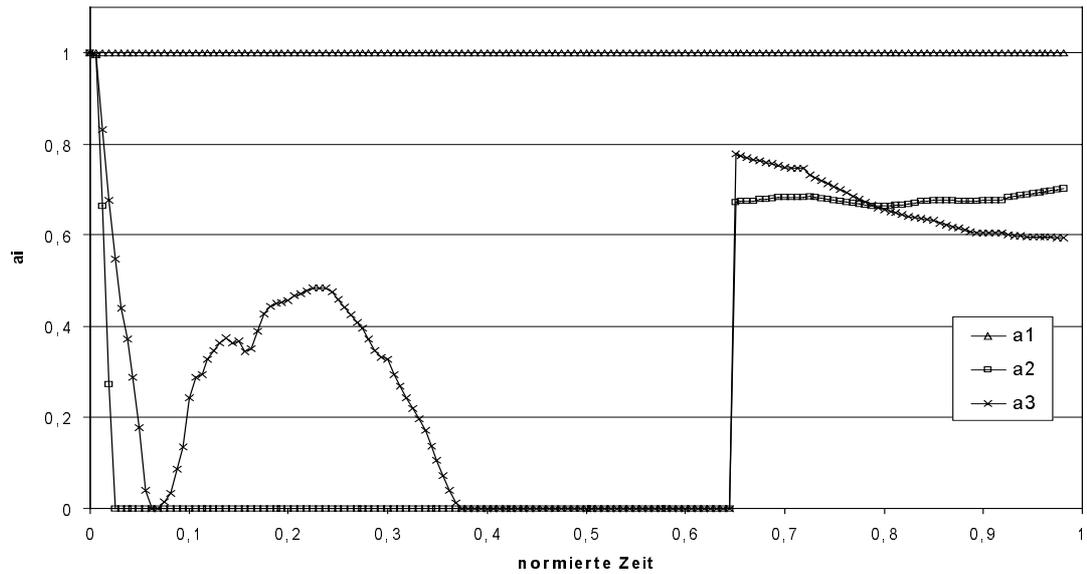


Abbildung 5-8 Gewichtsfunktionen a_i der Kultivierung gegen die Klassen 1-3

In der Abbildung 5-10 wird der Verlauf der Abstände dargestellt. Nach dem ersten Auftreten der Störung steigen die Abstände aller Klassen stark an. Dies resultiert in einem starken Anstieg des mittleren Abstands. Nach ca. 0,55 der normierten Zeit übersteigt dieser den Wert des absoluten Abstands D_{max} :

$$D_{max} = \max_{i \neq j} \{ \sup_t (D_{ij}(t)) \} \tag{Formel 5-10}$$

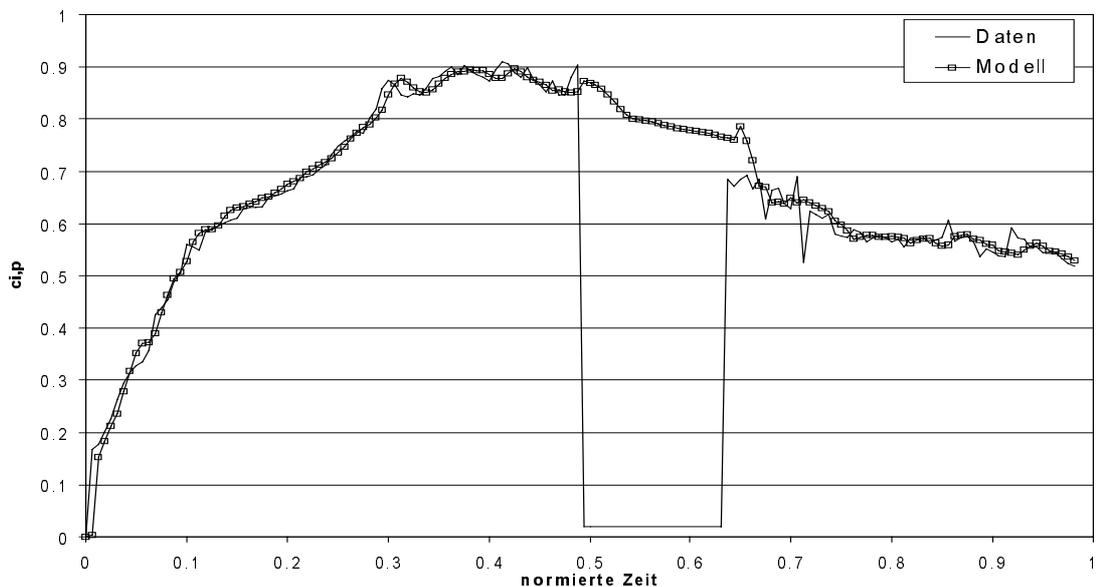
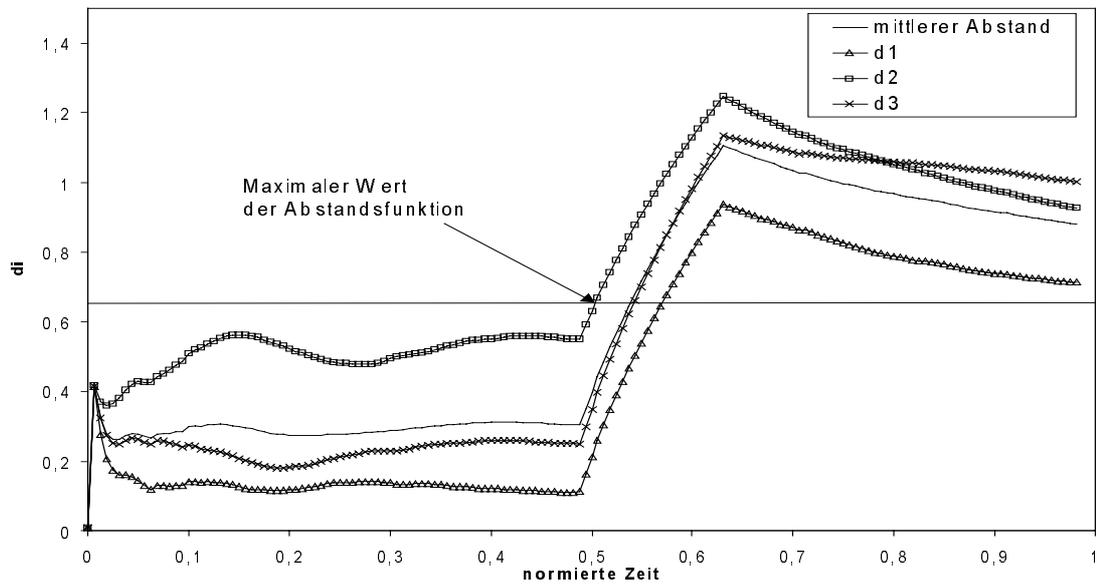


Abbildung 5-9 Modell für den erwarteten Verlauf der Kultivierung


 Abbildung 5-10 Abstandsfunktion d_i für die Kultivierung gegen die Klassen 1 – 3

Als Folge wird eine Warnung vom System an den Bediener ausgegeben. Der Störfall ist nach ca. 0,64 der normierten Zeit behoben und die Gewichte sowie der Faktor werden neu berechnet. Dies wird anhand des großen Sprungs innerhalb der Gewichte in Abbildung 5-8 deutlich.

5.6 Zusammenfassung

In den vorangegangenen Abschnitten wurde eine Methode vorgestellt, die es erlaubt die Ergebnisse einer Clusteranalyse archivierter Kultivierungsdaten für die Störfallanalyse, Fehlerbehandlung und On-line-Zustandsermittlung zu nutzen. Dabei zeigte sich, daß schon in einem frühen Stadium der Kultivierung zuverlässige Aussagen über die Klassenzugehörigkeit gemacht werden können. Als Ergebnis der Clusteranalyse liegen eine bestimmte Anzahl von Klassen vor. Diese Klassen sind charakterisiert durch den Verlauf des Mittelwertes der gemessenen Variable und dem Verlauf der Standardabweichung aller Kultivierungsverläufe, die den einzelnen Klassen angehören. In dem vorgestellten G2-Modul konnten neben der On-line-Klassifizierung auch fehlerhafte Daten erkannt werden. Hierbei werden Daten als fehlerhaft angesehen, wenn sie sich außerhalb eines Bereichs der zweifachen Standardabweichung aller

Klassen befinden. Im Falle eines Störfalls kann mittels der Clusteranalysen eine zuverlässige Modellvorhersage über den Verlauf des Meßsignals gegeben werden.

Die Ergebnisse der Clusteranalyse stellen eine Möglichkeit dar, den Verlauf einer Prozeßgröße noch genauer zu klassifizieren als mit den – im Kapitel 4 vorgestellten – Konfidenzbereichen. Die Clusteranalyse als Methode zur Datenanalyse eignet sich besonders, wenn zahlreiche Kultivierungsverläufe zur Analyse vorliegen. Es muß aber betont werden, daß in den Klassendefinitionen – die bei der Analyse erhalten werden – kein Wissen über die Qualität der Kultivierungen enthalten ist, sondern sie nur zur Erzeugung von spezifischen Merkmalen genutzt werden kann. Die Interpretation einer bestimmten Klassenzugehörigkeit muß durch Regeln erfolgen, die bei einer Wissensakquisition gewonnen werden.

6 Ein Expertensystem zur Wissensverarbeitung in der Stahlerzeugung

Ein wesentliches Problem regelbasierter Expertensysteme stellt die fehlende Übersichtlichkeit einer großen Regelmenge dar. Das in Kapitel 4 beschriebene Expertensystem zur Überwachung von Bioprozessen stellt einige Ansätze zur Lösung dieser Probleme vor; so z. B. mit der Möglichkeit nach Fehlern in der Wissensbasis mit Hilfe der Erklärungskomponente zu suchen. In diesem Kapitel soll ein anderer, vollständig graphisch orientierter Ansatz zur Wissensimplementierung und -visualisierung vorgestellt werden. Das Expertensystem in dem dieses Konzept Verwendung findet ist für die On-line-Ermittlung von Verfahrensvorschriften in einem Stahlwerk der Preussag AG erstellt worden und stellt ein Teilprojekt eines vollständig automatisierten Planungssystems dar.

6.1 Problemstellung

Zur Herstellung des Stahlwerkendprodukts, der Bramme bzw. dem Riegel, muß eine Reihe von Verfahrensvorschriften eingehalten werden. Jede Verfahrensvorschrift beschreibt dabei einen Produktionsaspekt auf dem Weg vom Rohstahl zum Endprodukt. Beispiele für Verfahrensvorschriften sind die Gießgeschwindigkeit, die Beruhigungsart, die Vakuumbehandlungsart u.a.m. Mit Hilfe von Verfahrensregeln wird die in der konkreten Situation gültige Verfahrensvorschrift ausgewählt. Die Situation wird allgemein beschrieben durch

- Ober- und Untergrenzen der Elemente in der Kataloganalyse,
- Zielwerte von Elementen in zusätzlich ermittelten Analysen,
- Zusatzinformation über die Stahlgüte wie Güteschlüssel,
- Schneidvorgaben wie Länge und Breite,
- die Abfolge von mehreren geplanten Schmelzen,
- bereits ermittelte Verfahrensvorschriften und
- Daten aus dem laufenden Prozeß.

Jede Verfahrensregel beinhaltet die Erkenntnisse langjähriger, betrieblicher Erfahrung.

Ziel der Arbeiten war es ein On-line-Expertensystem zu entwickeln, in dem das Wissen der Verfahrensregeln eines bestehenden Automationssystems – das mit herkömmlichen Programmier-techniken realisiert ist – übernommen werden soll.

6.2 Material und Methoden

Das Expertensystem wurde mit der Expertensystemshell G2 von Gensym erstellt. Sie ist als Client/Server Anwendung konzipiert, wobei der Client sich von PC's unter dem Betriebssystem Windows NT 4.0 (Microsoft) und Telewindows (Gensym) direkt auf einen Serverprozeß zugreifen kann. Abbildung 6-1 zeigt den Aufbau des Systems.

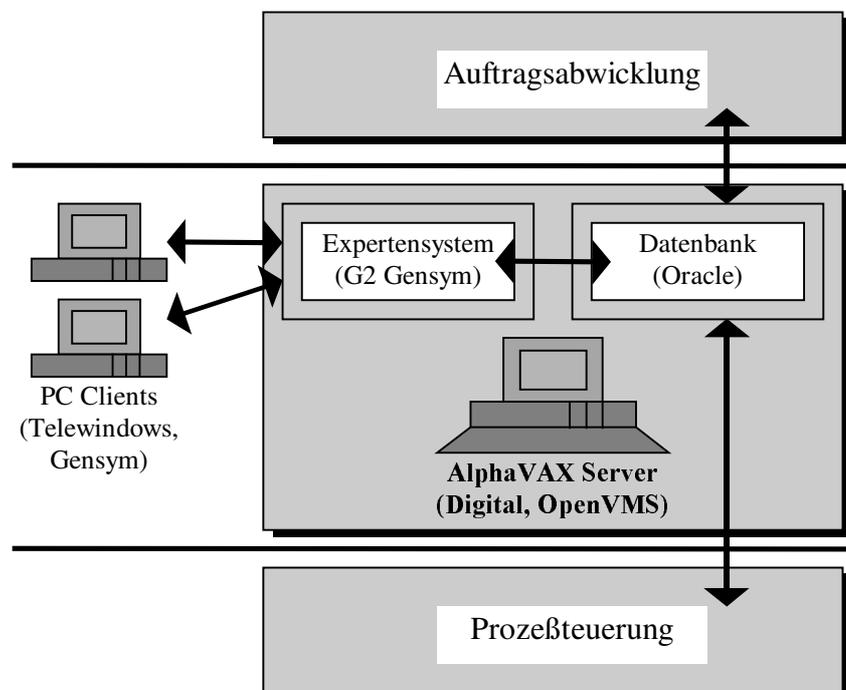


Abbildung 6-1 Aufbau des Systems

Der Serverprozeß läuft auf einer Digital AlphaVAX mit 512 MB RAM und 2GB Massenspeicher. Über eine kommerziell erhältliche Schnittstelle hat das Expertensystem direkten Zugriff auf die zentrale Datenbank, die auf dem gleichen Rechner implementiert ist. In der Datenbank, die mit ORACLE 7 realisiert wurde, laufen alle Informationen der Auftragsverwaltung und des Leitsystems zusammen.

6.3 Problemlösung

Im Zuge der Automatisierung des Stahlerzeugungsprozesses ist es von Nöten, möglichst viel Erfahrungswissen des Ingenieurs in eine für den Computer verständliche Form zu bringen. Durch die ständigen Veränderungen der Anforderung an den Prozeß, stellt sich das Erfahrungswissen der Experten als sehr dynamisch dar. Es muß deshalb in einem System zur Automatisierung immer die Möglichkeit gegeben sein weiteres Wissen in das System einzubringen. In einem früheren Ansatz zur Automatisierung wurde das Wissen der Ingenieure in einem herkömmlichen Programm (MIDAS-Regelwerk) realisiert. Diese Implementierung hat jedoch zahlreiche Nachteile, die durch den Einsatz von Expertensystemtechnologien entscheidend verbessert werden. Tabelle 6-1 werden die Nachteile des bisher verwendeten Systems den Vorteilen des implementierten Expertensystems gegenübergestellt.

Tabelle 6-1: Nachteile des herkömmlichen System und Vorteile des Expertensystems

<i>Bisheriges System</i>	<i>Expertensystem</i>
Vermischung von Expertenwissen und Wissen über den Programmablaufs	Klare Trennung von Expertenwissen und Wissen über den Programmablaufs
Keine graphische Benutzeroberfläche	Vollständige graphische Benutzeroberfläche
Umständliches Einbringen von neuen Regeln	Einfache Erstellung von Regeln mit 'Drag und Drop' Techniken
Keine Analyse- und Testwerkzeuge	Graphische Visualisierung der Schlußfolgerungen
	Konsistenzprüfung des Regelwerks

Um neue Regeln in das bisherige MIDAS Regelwerk einzubringen, mußte der Experte diese mit einer beschränkten Syntax in ein Terminal eingeben. Hierbei gab es keine Überprüfung der Syntax und keine Möglichkeit die eingegebene Regel zu testen. Des Weiteren war die Position der Regel innerhalb des Regelwerks von entscheidender Bedeutung während der Verarbeitung, da während der Verarbeitung in jedem Zyklus nur eine Regel ausgeführt wurde. Es war dementsprechend notwendig, daß der Experte sowohl mit der Syntax als auch mit dem Programmablauf vertraut war. Darüber hinaus gab es keine Möglichkeit die Logik einer neuen Regel zu prüfen.

Mit Hilfe der Expertensystemshell G2 wurde eine Anwendung geschaffen, die mittels einer graphischen Benutzeroberfläche dem Benutzer ermöglicht neues Wissen in das System einzubringen. Das Wissen wird in dem System übersichtlich dargestellt und es gibt weitreichende Möglichkeiten, die Wissensbasis auf Inkonsistenzen und Lücken zu überprüfen. Darüber hinaus ist das Expertensystem als Client/Server-Anwendung konzipiert, was das Verwalten der Wissensbasis von mehreren Benutzern gleichzeitig an unterschiedlichen Terminals erlaubt.

Die wesentlichen Aufgaben, die von dem Expertensystem unterstützt bzw. durchgeführt werden, sind die Verarbeitung von Anfragen an das Expertensystem (Inferenzkomponente), die Implementierung von neuem Wissen (Wissenserwerbskomponente), das Testen und Überprüfen einer neuen Wissensbasis (Erklärungskomponente) und das Freischalten einer neuen Wissensbasis. Im folgenden sollen die Funktionen und Arbeitsweisen des Expertensystems näher beschrieben werden.

6.3.1 Inferenzkomponente

In der Inferenzkomponente sind alle Funktionen zusammengefaßt, die für die Auswertung der Regeln notwendig sind. Der Anstoß für die Auswertung kann hierbei unterschiedlich erfolgen. Im On-line-Einsatz erhält das Expertensystem von einem externen Programm die Anfrage, für bestimmte Bedingungen die verfahrenstechnischen Entscheidung zu treffen. Der Austausch der Daten erfolgt hierbei über die Datenbank. Die Inferenzkomponente soll hierbei für jeden Arbeitsbereich genau eine Schlußfolgerungen liefern. Wenn bei keiner Regel eines Arbeitsbereiches die Prämisse erfüllt ist, so wird die Fehlermeldung 'Kein Ergebnis' erzeugt; sind hingegen die Bedingungen zweier Regeln erfüllt, so wird die Fehlermeldung 'Konflikt' zurückgegeben. In diesen Fällen ist die Wissensbasis nicht in der Lage eine sinnvolle Entscheidung zu treffen. Ein Bediener muß demnach die Entscheidung fällen und manuell eingeben. Um diese Situation näher analysieren und die Wissensbasis für diesen Fall erweitern zu können, werden alle Datensätze, bei denen Unregelmäßigkeiten auftreten, in der Datenbank separat gespeichert. In der Erklärungskomponente können die Daten für die Analyse der Wissensbasis verwendet werden. Abbildung 6-2 zeigt den Verarbeitungsablauf einer Anfrage und die Fehlerbehandlung der Inferenzkomponente.

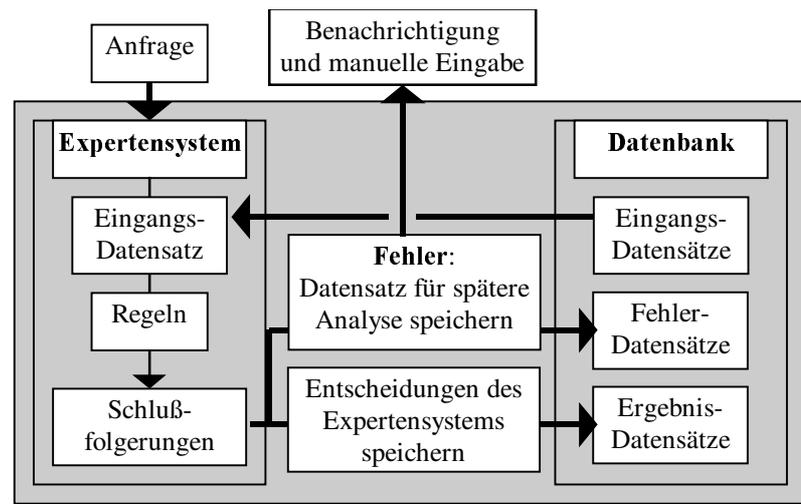


Abbildung 6-2: Inferenzkomponente und Fehlerbehandlung

6.3.2 Implementierung von neuen Regeln

Die Wissenserwerbskomponente stellt alle Funktionen zu Verfügung, die benötigt werden, um neues Wissen zu akquirieren. Die Komponente sollte dem Anspruch genügen, einfach und ohne umfangreiche Programmierkenntnisse bedienbar zu sein. Des Weiteren sollte das Wissen übersichtlich und transparent dargestellt werden. Abbildung 6-3 zeigt einen Bildschirmausschnitt der Benutzeroberfläche des Expertensystems. Alle Regeln werden den betreffenden unterschiedlichen verfahrenstechnischen Entscheidungen zugeordnet und auf Arbeitsbereichen sortiert. In der Abbildung 6-3 ist der Arbeitsbereich der Regeln dargestellt, die für die Auswahl eines Stopfens – der beim Abstich Verwendung findet – zuständig sind. Jede Regelposition besteht aus einem Startblock, mehreren Logikblöcken und ein oder mehreren Ergebnisblöcken. Durch die Art der Verbindung werden die logischen Zusammenhänge dargestellt. So bedeutet die Reihenschaltung von zwei Blöcken eine UND Verknüpfung und die Parallelschaltung eine ODER Verknüpfung. Die auf den Paletten vorliegenden Logikblöcke, die zur Erstellung von neuen Regeln dienen, sind nach Kategorien geordnet. Generell können fünf verschiedene Blöcke unterschieden werden. In der Tabelle 6-2 Beispiele von Logikblocktypen dargestellt.

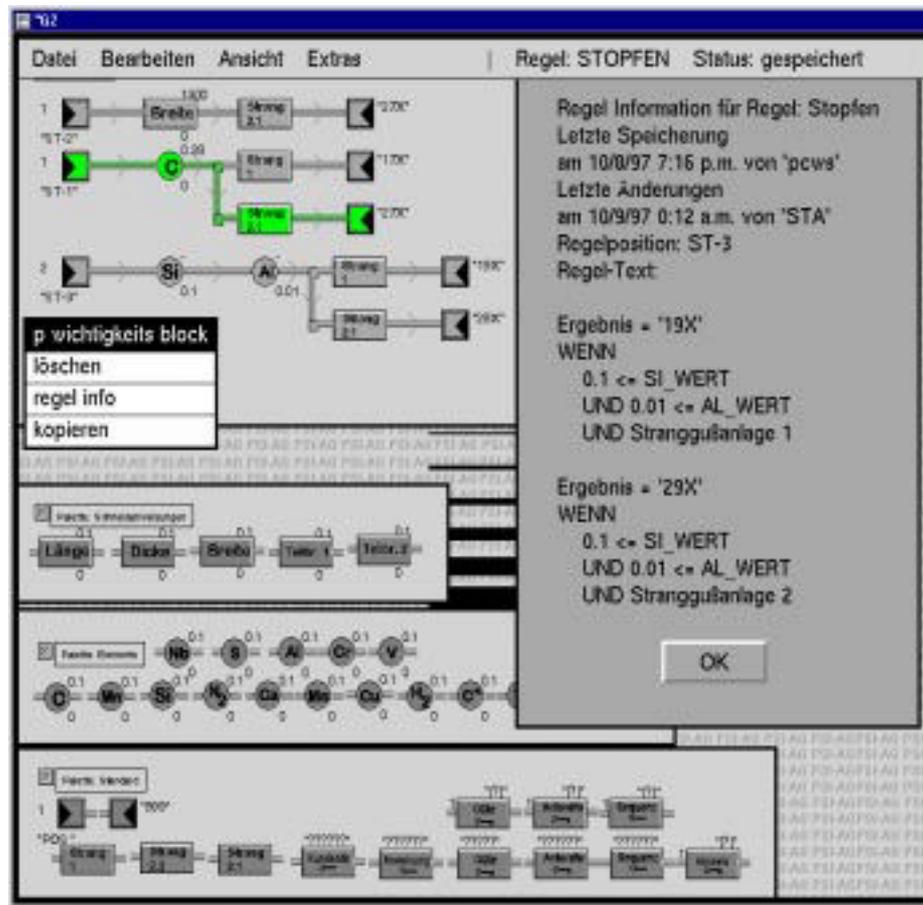


Abbildung 6-3: Bildschirmausschnitt der Wissenserwerbskomponente

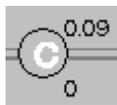
Tabelle 6-2: Beispiele von Logikblocktypen

Blocktyp	Beispiel	Blocktyp	Beispiel
Startblock	2  "ST-3"	Mengentextblock	1 
Ergebnisblock	 "29X"	Textvergleichblock	
Intervallblock		Intervallblock (Invertiert)	

Ein Startblock kennzeichnet den Anfang einer neuen Regel. Um in Konfliktfällen auf einzelne Regeln verweisen zu können, muß jeder Startblock mit einem eindeutigen Namen versehen werden. Durch eine Wichtigkeit, die am Startblock angegeben wird, können Konflikte zwischen einzelnen Regeln gelöst werden. Die Inferenzmaschine entscheidet sich im Konfliktfall für die Regel mit der größten Priorität. Je geringer die Wichtigkeit ist, desto größer ist die Priorität.

Ein Ergebnisblock beinhaltet die Schlußfolgerung der Regel. Beim Erstellen von Regeln muß darauf geachtet werden, daß jeder Ergebnisblock nur mit einem Startblock verbunden ist. Umgekehrt können aber mehrere Ergebnisblöcke mit einem Startblock verbunden sein. In der Abbildung liefert die erste Regel als Auswahl für den Stopfen das Ergebnis 27X.

Intervallblöcke bezeichnen alle Logikblöcke, die Zahlenintervalle bestimmter Werte abfragen. So existieren für alle Elemente der Analyse entsprechende Intervallblöcke, die graphisch voneinander unterscheidbar sind. Liegt der Wert der Analyse für das entsprechende Element innerhalb des Intervalls, ist der Logikblock wahr. Ein Intervallblock läßt sich zusätzlich invertieren, wodurch der Ausdruck negiert wird. Der negierte Intervallblock



bedeutet demnach:

WENN NICHT ($0,0 \% < \text{Kohlenstoffgehalt der Analyse} < 0,09 \%$)

Ein Textvergleichsblock wird verwendet, um bestimmte Schlüssel abzufragen. Die Qualität eines Stahls wird z. B. als fünfstelliger Güteschlüssel angegeben. Hierbei hat jede Stelle des Schlüssels eine spezifische Bedeutung, die abgefragt werden kann. Die Fragezeichen stellen hierbei Platzhalter dar. Der Ausdruck des Textvergleichsblocks



bedeutet:

WENN die erste Stelle der Anforderungsnummer eine '4'

UND die vierte Stelle eine '3'

UND die fünfte Stelle eine '3'

UND die sechste Stelle eine '6' ist

Mit einem Mengentextblock kann das Vorhandensein einer bestimmten Menge von Angaben in einem Schlüssel gesucht werden. Die Zahl links neben den Block zeigt an welche Stelle des Schlüssels untersucht werden soll. Der Logikblock ist wahr, wenn eines der Zeichen in den eckigen Klammern mit der Stelle des Schlüssels übereinstimmt. So bedeutet der Mengentextblock, der in der Tabelle 6-2 dargestellt ist:



bedeutet:

WENN die erste Stelle des Hinweisschlüssels ein 'M'

ODER ein 'U'

ODER ein 'Z' ist.

Will der Benutzer eine neue Regel erstellen, so wählt er zuerst aus dem Menü den Entscheidungsbereich aus, für den eine neue Regel erstellt werden soll. Von den angezeigten Paletten müssen zur Erstellung einer Regel ein Startblock und mindestens ein Ergebnisblock mit Hilfe der Maus auf die Arbeitsfläche gezogen werden. Zwischen dem Startblock und dem Ergebnisblock werden im folgenden alle Blöcke – die für die Regel benötigt werden – positioniert und miteinander verbunden. Jeder Block besitzt zunächst eine offene Ein- und Ausgangsverbindung. Durch Ziehen dieser offenen Verbindungen mit Hilfe der Maus können zwei Blöcken miteinander verbunden und eine logische Verknüpfung erstellt werden. Sind alle Blöcke miteinander verbunden, so können ihre Werte direkt editiert werden. Eine automatische Sortierfunktion ordnet die Verbindungen und Abstände der Blöcke und platziert die Regel nach ihrer Wichtigkeit sortiert innerhalb der vorhandenen Regeln ein. In Abbildung 6-4 ist eine Regel für die Auswahl des Zusatzdamm dargestellt.

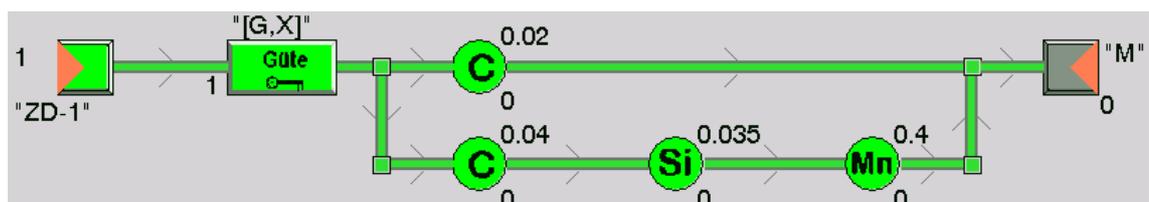


Abbildung 6-4 Beispiel einer Regel

Sprachlich formuliert lautet die Regel:

WENN die erste Stelle der Güte ein G ODER ein X ist

UND

($C < 0,02$

ODER

($C < 0,04$ UND $Si < 0,035$ UND $Mn < 0,4$))

DANN

verwende Zusatzdamm ,M'

6.3.3 Testen einer neuen Wissensbasis

Die Erklärungskomponente beinhaltet alle Funktionen, die notwendig sind, um die Schlußfolgerungen des Expertensystems zu verdeutlichen und auf Inkonsistenzen zu überprüfen. Die Überprüfung einer Wissensbasis erfolgt nachdem neue Regeln in der Wissensbasis implementiert worden sind und bevor sie für externe Anfragen freigegeben werden. Ruft man die Erklärungskomponente auf, werden zunächst Testdatensätze aus der Datenbank in das Expertensystem geladen. Diese beinhalten einen repräsentativen Durchschnitt der verwendeten Güten und Analysen sowie alle Datensätze, die schon einmal in dem Expertensystem zu Fehlentscheidungen geführt haben. Abbildung 6-5 zeigt das Menü der Erklärungskomponente. Wesentlicher Bestandteil dieses Menüs ist hierbei ein Datensatzbrowser in denen einzelne Testdatensätze ausgesucht und die Ergebnisse für jede Verfahrensvorschrift angezeigt werden.

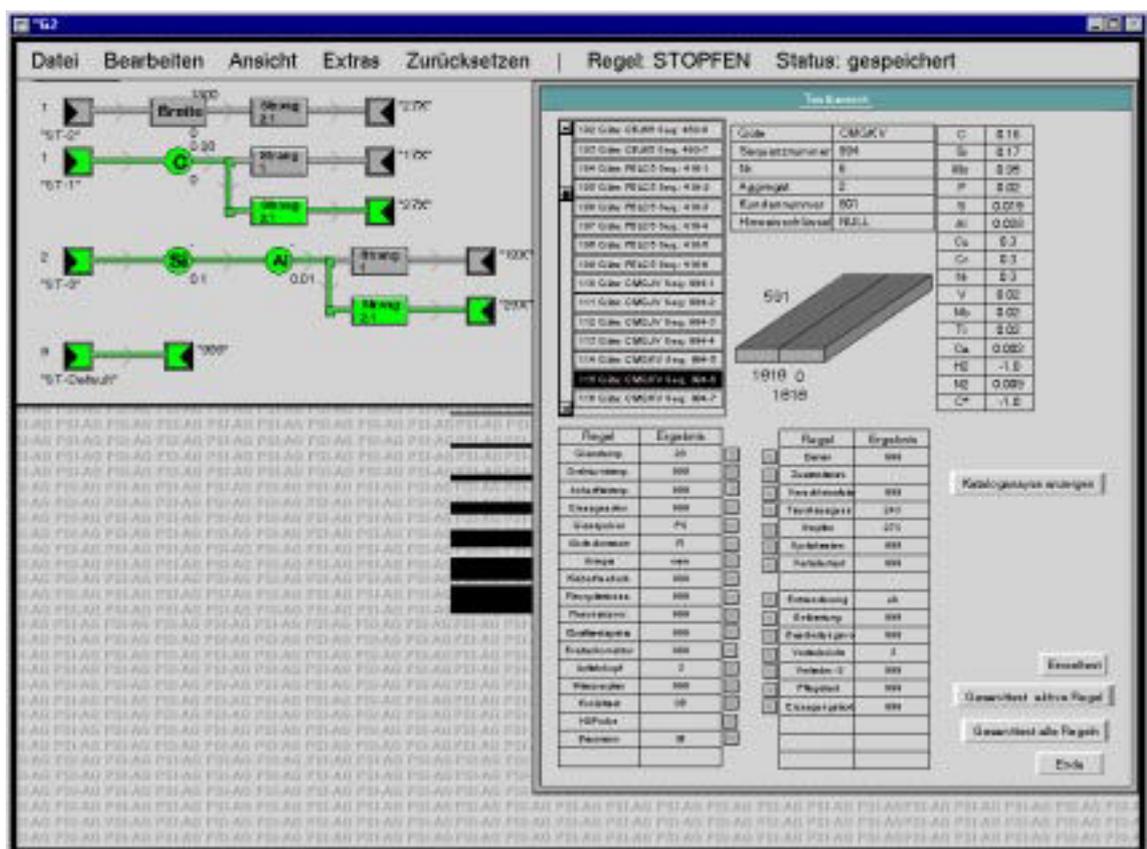


Abbildung 6-5: Erklärungskomponente

Jeder Testdatensatz besteht aus einem Ist-Zustand, der u. a. durch die Elementaranalyse charakterisiert ist und dem Soll-Zustand, der durch die Güte gekennzeichnet ist. Aus Gründen der

Übersichtlichkeit wird nur ein Güteschlüssel angezeigt. Mit der Funktion 'Kataloganalyse' werden aus der Datenbank die Intervalle der Elementaranalyse ermittelt, die ein Stahl einer bestimmten Güte einhalten muß. Abbildung 6-6 zeigt den Arbeitsbereich, der von dem System für die Kataloganalyse der Güte PBL5 angezeigt wird. Bereiche der Tabelle in denen eine negative oder das Symbol 'NONE' steht bedeuten, das keine Angaben für dieses Element in der Datenbank gefunden wurden. Es stehen im wesentlichen zwei verschiedene Tests zur Verfügung, um die Wissensbasis mit den Testdatensätzen zu prüfen. Mit dem Einzeltest, wird die Schlußfolgerung des Expertensystems für eine ausgewählte Kategorie und einem ausgewählten Testdatensatz erfragt und visuell dargestellt. Dabei werden alle Regeln, deren Bedingungsteile erfüllt sind, grün gefärbt. Ein Bediener kann anhand dieser Visualisierung genau erkennen, warum das Expertensystem eine spezifische Entscheidung fällt. Bei dem Einzeltest werden die Regeln des ausgesuchten Arbeitsbereiches auch auf fehlerhafte logische Strukturen z. B. zyklische Verbindungen überprüft. In dem Gesamttest werden alle Testdatensätze auf die vollständige Wissensbasis angewendet, um Konflikte und Fälle, in denen keine Entscheidung getroffen werden kann, aufzudecken. Die Ergebnisse dieses Gesamttests werden in der Datenbank gespeichert. Dadurch ergibt sich ein direkter Vergleich zwischen allen durchgeführten Gesamttests, so daß unterschiedliche Entscheidungen deutlich werden.

C	0.0	0.03	AlC	-1	NONE
Si	0.0	0.034	AlSi	-1.0	-1.0
Mn	0.23	0.27	AlMn	-1	NONE
P	0.0	0.012	AlP	-1.0	-1.0
S	0.0	0.01	AlS	0.0	0.0
Al	0.03	0.055	B	-1	NONE
Cu	0.0	0.03	CCN1	0.0	0.08
Cr	0.0	0.03	CCN2	-1.0	-1.0
Ni	0.0	0.04	CCN3	-1.0	-1.0
V	0.0	0.004	Co	-1	NONE
Nb	0.0	0.002	Fe	-1	NONE
Ti	0.0	0.002	H	-1	NONE
Mo	0.0	0.005	Ma	-1.0	-1.0
Ca	0.0	0.0	O2	-1	NONE
H2	-1.0	-1.0	Pb	-1	NONE
N2	0.004	0.006	Sb	-1	NONE
C*	-1.0	-1.0	SIPEq	-1.0	-1.0
			Sn	0.0	0.02
			TIEq1	-1.0	-1.0
			TIEq2	-1.0	-1.0
			VnBa	-1	NONE
			W	-1	NONE
			Zr	-1	NONE

Ende

Abbildung 6-6 Arbeitsbereich zum Anzeigen von Kataloganalysen

6.3.4 Aktivieren einer neuen Wissensbasis

Die Anforderungen an die Verfügbarkeit des Expertensystems sind sehr hoch. Durch den kontinuierlichen Betrieb der Stahlerzeugung ist das ständige Abfragen des Expertensystems notwendig. Um die zeitgleiche Akquisition von neuem Wissen und die Überprüfung einer neuen Wissensbasis mit der Verarbeitung von Anfragen zu ermöglichen, findet die Verarbeitung der Anfragen in einem separaten Prozeß statt. In dem Entwicklungsprozeß wird mit der Hilfe der Wissenserwerbskomponente und Erklärungskomponente eine neue Wissensbasis erstellt, die für den Prozeß freigegeben werden muß. In Abbildung 6-7 ist der Vorgang der Aktivierung einer neuen Wissensbasis dargestellt.

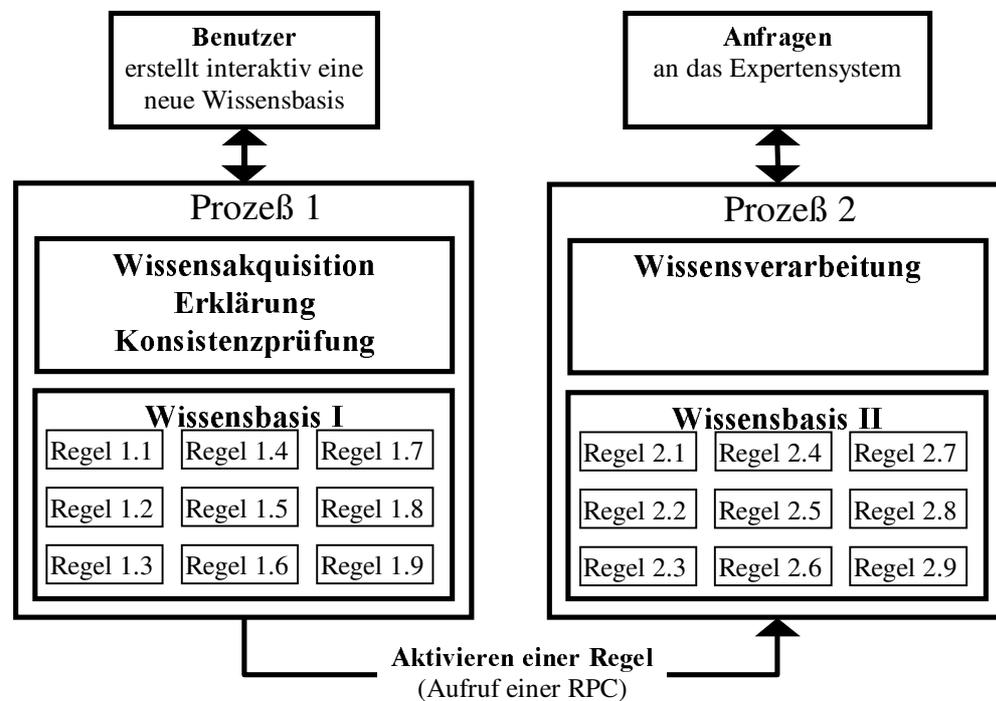


Abbildung 6-7: Aktivierung einer neuen Wissensbasis

Bei der Aktivierung einer Wissensbasis ruft der Prozeß I – über eine G2-Schnittstelle – in dem Prozeß II eine RPC (remote procedure call) auf, welche die entsprechend modifizierte Regeln austauscht. Dies geschieht während des Betriebs des Expertensystems, so daß keine Ausfallzeiten auftreten und alle Anfragen an das Expertensystem ohne Unterbrechung bearbeitet werden.

6.4 Zusammenfassung

Mit dem erstellten Expertensystem ist eine neuartige graphische Möglichkeit geschaffen worden, das Erfahrungswissen von Ingenieuren im Computer zu implementieren und zu visualisieren und somit für die Automation nutzbar zu machen. Darüber hinaus bietet das Expertensystem vielfältige Möglichkeiten, um das implementierte Wissen zu verwalten und zu erweitern. Bei der Wissensakquisition ergibt sich durch die vollständig graphische Benutzeroberfläche, die einfache 'Drag und Drop'-Techniken' nutzt der Vorteil, daß Ingenieure ohne besondere Programmierkenntnisse ihr Wissen selber in das Expertensystem einbringen können. Dies liefert somit einen wertvollen Beitrag, um den Flaschenhals "Wissensakquisition" in der Expertensystementwicklung zu beseitigen. In dem bestehenden Expertensystem wurde mit Hilfe der Benutzerschnittstelle ca. 200 Regeln implementiert, die im industriellen Einsatz ca. 32 unterschiedliche Verfahrensparameter ermitteln. Die in der Erklärungskomponente gegebenen Möglichkeiten, mit einer repräsentativen Anzahl von Testdatensätzen die Ergebnisse des Expertensystems zu visualisieren und auf Inkonsistenzen zu prüfen, stellt einen Beitrag zur Übersichtlichkeit komplexer Expertensysteme dar. Die Erfahrung mit dem Expertensystem im praktischen Betrieb zeigen, daß durch das System die Transparenz des Wissens wesentlich erhöht worden ist. So konnten zahlreiche Redundanzen in den Vorschriften des alten Systems entdeckt werden. Die Implementierung und das Testen von neuem Wissen in dem alten MIDAS-System stellte eine langwierige und komplizierte Tätigkeit dar. Aufgrund der neuen Möglichkeiten die Wissensbasis des Expertensystems einfach zu erweitern und zu testen ist die Motivation der Mitarbeiter gestiegen auf neue Situationen direkt zu reagieren. Dadurch bleibt die Wissensbasis ständig aktuell und stellt ein aktuelles Abbild des Wissens der beteiligten Mitarbeiter dar.

Die Domäne, für die das Expertensystem entwickelt wurde, charakterisiert sich durch viele jeweils abgeschlossene Entscheidungen, die sich nicht gegenseitig bedingen. In dem System sind z. Z. keine Regeln implementiert, welche die getroffenen Entscheidungen weiter interpretieren. Des Weiteren gibt es keine Möglichkeit auf das zeitliche Verhalten des Prozesses sowie auf die Ergebnisse vorheriger Inferenzzyklen zurückzugreifen. Dieses wäre aber für den Einsatz des Expertensystems zur Zustandsermittlung und Überwachung von Bioprozessen notwendig.

7 Schlußbetrachtung

Für die zuverlässige Überwachung, Störfalldiagnose und Zustandsermittlung von Bioprozessen wird viel Wissen zahlreicher Fachgebiete benötigt. Dieses Wissen wird üblicherweise von erfahrenen Prozeßbediener zur Verfügung gestellt. Um eine Prozeßüberwachung mit Expertensystemen zu automatisieren muß das Wissen zunächst gesammelt und in dem System implementiert werden. In dieser Arbeit wurde ein Expertensystem entwickelt, mit dem Bioprozesse überwacht werden können. Das Expertensystem ist so konzipiert, daß es flexibel und schnell für die Überwachung unterschiedlicher Prozesse angepaßt werden kann. Um dies zu erreichen, wurde eine leistungsfähige, graphische Benutzerschnittstelle implementiert, die den Bediener in allen Schritten der Wissensakquisition interaktiv unterstützt. Die vorliegende Arbeit bietet Lösungen, um die wichtigsten Schritte der Wissensakquisition erheblich zu vereinfachen. So ermöglicht die Analyse historischer Daten eine induktive Erhebung prozeßspezifischen Faktenwissens. Darüber hinaus zeigen Untersuchungen mit Ergebnissen der Clusteranalyse, daß sie direkt zur On-line-Klassifizierung, Fehlererkennung und Prozeßvorhersage genutzt werden können. Ein weiterer wichtiger Schritt der Wissensakquisition stellt die Implementierung von Prozeß- und Fehlermodellen durch den Benutzer dar. Hierfür wurde eine vollständig dialoggeführte Regelerstellung entwickelt, mit der fehlerfreie Regeln erstellt werden. Das Problem der fehlenden Übersichtlichkeit großer Regelmengen wurde durch die Entwicklung einer speziellen Erklärungskomponente gelöst. Diese Erklärungskomponente bietet Funktionen zum gezielten Durchsuchen einer Wissensbasis; so lassen sich beispielsweise die Regeln, die zu einer Schlußfolgerung geführt haben, direkt anzeigen.

Ein anderer Ansatz wurde hingegen bei der Entwicklung eines graphischen Moduls zur Wissensimplementierung und -visualisierung verfolgt. In diesem Modul bleibt die Übersichtlichkeit der Wissensbasis aufgrund der graphischen Darstellung der Regeln gewahrt. Ein zusätzliches Testmodul erlaubt darüber hinaus ein schnelle überprüfen der Wissensbasis auf Konsistenz. Die Erfahrungen im industriellen Betrieb haben hierbei gezeigt, daß die Transparenz der Wissensbasis und somit die Akzeptanz von Expertensystemen erheblich vergrößert wurde.

Beim Einsatz des Expertensystems zur Überwachung von Bioprocessen wurden Prozeßzustand und Störungen sowohl bei der Kultivierung von *Acremonium chrysogenum* als auch bei einer Hefe-Kultivierung zuverlässig erkannt. Hierdurch reduziert sich der erforderliche Aufwand des Bedienpersonals bei der Überwachung von Bioprocessen erheblich. Durch eine zusätzliche – auf dem Internet basierende – Visualisierung wird eine wesentlich höhere Verfügbarkeit der Ergebnisse des Expertensystems erreicht.

Speziell bei der Anpassung des Expertensystems an neue Prozesse hat sich gezeigt, daß die symbolische Repräsentation von Wissen erheblich dazu beiträgt den Aufwand bei einer Anpassung des Expertensystems für einen neuen Prozeß zu reduzieren. Das System sollte in zukünftigen Untersuchungen für weitere Prozesse mit anderen Kulturen, Prozeßabläufen und neuen Störfällen angepaßt werden, um die Verallgemeinerung weiter zu überprüfen und zu verbessern.

Zur Prozeßüberwachung werden immer häufiger komplexe Analysensysteme wie z. B. die FIA oder HPLC Anlagen an den Prozeß gekoppelt. Für die Überwachung dieser komplexen Analysensysteme werden zunehmend eigene Expertensysteme entwickelt und eingesetzt. Die Anbindung des Expertensystems an die Expertensysteme komplexer Analysensysteme wäre ein Ansatz zur integrativen Überwachung des gesamten Prozeßvorgangs. Neben der einheitlichen Darstellung und Visualisierung der Ergebnisse und Meldungen der einzelnen Systeme würde die zusätzliche Information weiterführende Diagnosen über den Prozeßzustand erlauben. Die Möglichkeiten zur Nutzung des globalen Informationsdienstes Internet bzw. WWW sind bislang nur wenig ausgelotet. Die im Rahmen dieser Arbeit durchgeführten Untersuchungen zum Einsatz des Internet zur Erhöhung der Verfügbarkeit des Systems stellen nur eine Möglichkeit dar. Denkbar sind beispielsweise zentrale Expertensysteme, die für spezielle Fachdomänen eine Fernüberwachung als Dienstleistung anbieten. Zur Zeit wird das Internet vornehmlich als reiner Informationspool genutzt, in dem kleinste Datenstrukturen ausgetauscht werden. Zunehmend finden aber Objekt-orientierte Ansätze – wie sie CORBA bietet – zum Austausch ganzer Objekte Verwendung. Würde ein standardisiertes Protokoll zum Austausch von allgemeinem Wissen existieren, so könnte sich das Internet zum Superexpertensystem entwickeln, das für Anfragen jeder Fachdomäne nicht nur Faktenwissen sondern auch Methodenwissen zur Verfügung stellt.

Erfahrungen mit der Shell

Generell kann gesagt werden, daß moderne Expertensystementwicklungstools vielfältige Möglichkeiten bieten schnell robuste Anwendungen zu entwickeln. Mit G2 im speziellen ist es möglich vollständige Client/Server-Umgebungen zu schaffen, die zahlreiche Funktionalitäten eines Prozeßleitrechners beinhalten. Bei allen Vorzügen die solche Entwicklungsumgebungen zur Verfügung stellen muß betont werden, daß eine erhebliche Einarbeitungszeit notwendig ist. Bei der Vielfalt der Möglichkeiten eine Problemlösung zu programmieren, ist viel Erfahrung notwendig, um innerhalb der Shell die optimale Methode zu generieren.

G2-Applikationen können ohne weitere Probleme auf unterschiedlichen Plattformen betrieben werden. So wurde innerhalb des Projekts G2 auf den Betriebssystemen UNIX und VMS betrieben.

Mit G2 erhält man schnell lauffähige Anwendungen, die nach der 'rapid-prototyping'-Methode entwickelt werden. Dem unerfahrenen G2-Anwender bieten sich hierdurch die Möglichkeit schnell die Entwicklungsumgebung kennenzulernen und zu überprüfen, ob die geplanten Methoden erfolgreich eingesetzt werden können. Nach dieser Lernphase ist es ratsam den Prototyp zu verwerfen und die Anwendung von Grund auf neu zu erstellen, um zu vermeiden, daß sich zu viele 'Altlasten' in der Anwendung anzusammeln.

8 Abkürzungsverzeichnis

CORBA	Common Object Request Broker Architecture
DEC	Digital Equipment Corporation, Massachusetts, USA
FIA	Flow Injection Analysis
GB	Giga Byte
HPLC	High Performance Liquid Chromatography
HTML	Hypertext Markup Language
LAN	Local Area Network
MB	Mega Byte
PC	(IBM kompatibler) Personal Computer
RAM	Random Access Memory
TCP/IP	Transmission Control Protocol / Internet Protocol
VMS	Virtual Memory System
WWW	World Wide Web

9 Literaturverzeichnis

- Aldridge, S. (1994) Chemical Engineering Meeting Spotlights New Technologies for Bioprocess Control, Sonderdruck aus: Genetic Engng. News, 14, (18)
- Alford, J.; Fowler, G. L.; Higgs, R. E.; Clapp, D. L.; Huber, F. M. (1992) Development of Real-Time Expert-System Applications for the On-line Analysis of Fermentation Data, in: Ladick, M.; Rose, A. (eds.) Harvesting Biotechnology for the 21st century, Amer. Chem. Soc., Seite 375-379
- Aarts, R. J.; Suviranta, A.; Rauman-Aalto, P.; Linko, P. (1990) An Expert System in Enzyme Produktion Control, Food Biotechnol. 4, (1), 301-315
- Aynsley, M.; Hofland, A.; Morris.; Montague, G.; Di Massimo, C. (1993) Artificial Intelligence and the Supervision of Bioprocesses, Advances in Biochemical Engineering Biotechnology, 48, 1-27
- Benez, H. (1995) Multimedia für die Automatisierungstechnik, atp, 37 (10), 79-85
- Boelow, T.; Mühlenfeld, E.; Nicolaou, M. (1996) Fallbasierte Automatisierung zeitvarianter Prozesse am Beispiel von Verbrennungsanlagen, Bildverarbeitung und Auge-Hand-Koordination, atp, 38, (12), 50-66
- Brandt, J. (1994) Wissensbasierte Überwachung und Fehlerdiagnose komplexer Prozeß-analysensysteme Dissertation, Universität Hannover
- Brandt, J.; Hitzmann, B. (1994) Knowledge-based fault detection and diagnosis in flow-injection analysis, Anal. Chim. Acta, 291, 317-347
- Brosilow, C.; Sterling, L.; Renard, F.-X. (1993) Knowledge Verification in Expert Systems Combining Declaritive and Procedural Representations, Computers chem. Engng., 17 (11), 1067-1090
- Buchanan, B.G.; Barstow, D.; Bechtal, R.; Bennet, J.; Clancey, W. (1983) Constructing an Expert System, in: Waterman, D.; Lenat, D. (eds.) Building Expert Systems, Hayes-Roth, London, Seite 127-167
- Cauvin, S.; Braunschweig, B.; Galtier, P.; Glaize, Y. (1993) Model-Based Diagnosis for Continous Process Supervision: The Alexip Experience, Engng. Applic. Artif. Intell., 6 (4), 333-343
- Charwat, H.J. (1996) Farbkonzept für die Prozeßführung mit Bildschirmen, atp, Teil 1: 38 (5), 50-53, Teil 2: 38 (6), 58-65

- Chynoweth, D. P.; Svoronos, S. A.; Lyberatos, G.; Harman, J. L.; Pullammanappallil, P.; Owens, J.M.; Peck, M. J. (1994) Real-Time Expert System Control of Anaerobic Digestion, *Wat. Sci.*, 30 (12), 21-29
- Clapp, K. P.; Ruel, G. J. (1991) Expert Systems in Bioprocessing, *BioPharm* 4 (2), 28-35
- Cooney, C.L.; O'Connor, C.L.; Sanchez-Riera, F. (1988) An Expert System for Intelligent Supervisory Control of Fermentation Processes, in: *Proc. 8th International Biotechnology Symposium, Paris*
- Davidson, V. J. (1994) Expert Systems in Process Control, *Food Research Intern.*, 27 (2), 121-128
- Davis, J. F.; Kannan, S.; Ramanathan, P. (1993) Use Knowledge-Based-System Programming Toolkits to Improve Plant Troubleshooting, *Chem. Engng. Prog.*, June, 55-58
- Fogarty, T.C.; Pan, W.; Sharpe P.K.; Miles, R.G. (1994) A Comparative Study of Artificial Neural Networks and Rule-based Techniques in the Development of a Computer-aided Control System, *Engng. Applic. Artif. Intell.*, 1 (1), 53-58
- Fraleigh, S. P.; Finch, F. E.; Stanley, G. M. (1992) Integrated Dataflow and Sequential Control in a Graphical Diagnostic Language, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) *On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA*, Seite 101-108
- Fraleigh, N.; Lauer, M. (1994) Werkzeuge für das Prozeßmanagement, Sonderdruck aus: *Design & Elektronik – Zukunftstechnologien, MagnaMedia AG, Haar, Deutschland*
- Gang, R.; Jicheng, W.; Shuqing, W. (1992) Expert System for Supervisory Control of an Ethanol Oxidizing Plant, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) *On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA*, Seite 293-297
- Gehlen, S.; H. Tolle; Kruzig, J.; Friedl, P. (1992) Integration of Expert Systems and Neural Networks for the Control of Fermentation Processes, in: Karim, M. N.; Stephanopoulos, G. (eds.) *IFAC Mod. and Contr. of Biotechnol. Proc.*, Colorado, USA, 10, 320-323
- Gehlen, S. (1993) Untersuchungen zur wissensbasierten und lernenden Prozeßführung in der Biotechnologie, *Fortschritt Berichte VDI Verlag*, 20, 87
- Giron-Sierra, J. M.; Gomez-Pulido, J. A.; Andres-Toro, B. (1994) Simulation Based Real-Time Monitoring of an Experimental Bioprocess using an Expert System Shell, in: Hamza, M. H. (ed.) *Proc. of the IASTED Int. Conf. Appl. Model. and Simulation, Anaheim, USA*, 209-203
- Gollmer, K.; Posten, C. (1995) Pattern Recognition for Phase Detection in Bioprocesses, in: Munak, A.; Schügerl, K. (eds.) *Proc. 6th Intern. Conf. on Comp. Applic. in Biotechnol.*, Garmisch-Partenkirchen, Germany, Seite 41-46

- Gomersall, R.C.; Löhn, A.; Stenns, M.; Hitzmann, B. (1996) Visualisierung von Prozeßdaten über das World Wide Web, Chem.-Ing.-Tech., 68 (9), 1311-1313
- Gomersall, R.C.; Hitzmann, B.; Guthke R.(1997) Fermentation Fault Management and On-line State Estimation using Methods of Cluster Analysis, Bioprocess Eng., 17, 69-73
- Grumann, S. (1997) Untersuchungen zur Integration von Expertenwissen in ein System zur Modellierung, Überwachung und Regelung von bioverfahrenstechnischen Prozessen I, Abschlußbericht, Braunschweig
- Guthke, R.; Roßmann, R. (1991) Fermentation Analysis by clustering, Bioprocess Eng., 6, 157-161
- Guthke, R. (1992) Learning of Rules from Fermentation Data, in: Karim, M. N.; Stephanopoulos, G. (eds.) IFAC Mod. and Contr. of Biotechnol. Proc., Colorado, USA, 10, Seite 157-161
- Guthke, R.; Ludwig, B. (1994) Generation of Rules for Expert Systems by Statistical Methods, Acta Biothechnol., 14 (1), 13-26
- Haferkamp, B. (1995) Transitioning a Fermentation Process Control System to G2, Gensym Users Society Meeting, Cambridge, USA
- Halme, A.; Karim N. (1989) Expert System Approach to Recognize the State of Fermentation and to Diagnose Faults in Bioreactors, in: Fish, N.M.; Fox, R.I.; Thornhill, N.F. (eds.) Computer Applications in Fermentation Technology: Proc. 4th Int. Conf. on Computer Applications in Fermentation Technology, Elsevier, Univ. of Cambridge, UK
- Halme, A. (1991) Combining Symbolic and Numerical Information in Modelling the State of Biotechnological Processes, in: Proc. European Control Conference, ECC 91, Grenoble, France
- Harmon, P. (ed.) (1993) The market for expert systems building tools, Intell. Software Strategies, 9 (6)
- Hitzmann, B. (1989) SUPERVISOR, ein Realzeit-Expertensystem für die Prozeßführung einer *E. coli*-Kultivierung, VDI-Verlag, Düsseldorf
- Hitzmann, B. (1992) Expertensysteme zur Automation von Bioprozessen, Chem.-Ing.-Tech., 64 (8), 708-714
- Hitzmann, B.; Lübbert, A.; Schügerl, K. (1992) An Expert System Approach for the Control of a Bioprocess, Biotech. and Bioeng., 39, 33-43
- Hoff, M. (1997) Ein Expertensystem zur On-line-Überwachung von HPLC-Systemen, Dissertation, Universität Hannover
- Ignova, M.; Paul, G. C.; Glassey, J.; Ward, A. C.; Montague, G. A.; Thomas, C. R.; Karim, M. N. (1996) Towards Intelligent Process Supervision: Industrial Penicillin Fermentation Case Study, Computers chem. Engng., 20 (Suppl.), 545-550

- Isermann, R. (1996) Modellgestützte Überwachung und Fehlerdiagnose Technischer Systeme, atp, Teil 1: 38 (5), 9-20, Teil 2: 38 (6), 48-57
- Isermann, R. (1992) Wissensbasierte Fehlerdiagnose technischer Prozesse, at, 36 (11), 421-426
- Jones, A. V.; Rodd, M. G. (1993) Problems with Expert Systems in Real-Time Control, Engng. Applic. Artif. Intell., 6 (6), 499-506
- Kaldorf, S. (1992) Echtzeit-Expertensysteme in der chemischen Industrie, Intelligente Software Technologien, 2, 13-17
- Kantrowitz, M., [Http://www.cs.cmu.edu/Groups/AI/html/faqs/ai/top.html](http://www.cs.cmu.edu/Groups/AI/html/faqs/ai/top.html)
- Karbach, W.; Linster, M. (1990) Wissensakquisition für Expertensysteme, Carl Hanser Verlag, München
- King, R. (1992) Plant and Sensor Failure Detection by Multiple Non-Linear Filters, in: Karim, M. N.; Stephanopoulos, G. (eds.) IFAC Mod. and Contr. Of Biotechnol. Proc., Colorado, USA, 10, Seite 129-134
- Kirchner, U. (1997) Anwendungen der 2D-Fluoreszenzspektroskopie zur Bioprozeßkontrolle, Diplomarbeit, Universität Hannover
- Konstantinov, K. B.; Yoshida, T. (1992) A Method for On-line Reasoning about the Time-Profiles of Process Variables, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA, 133-138
- Konstantinov, K. B.; Yoshida, T. (1992) Knowledge-Based Control of Fermentation, Biotechnol. and Bioengng., 39, 479-486
- Konstantinov, K. B.; Yoshida, T. (1990) On-line Monitoring of Representative Structural Variables in Fed-Batch Cultivation of Recombinant *Escherichia coli*. for Phenylalanine Production, J. of Ferm. and Bioeng., 70 (6), 420-426
- Kuesten, M.; McLellan, P. (1994) Expert System Shells – Selecting The Most Appropriate Development Enviroment, Food Reasearch Intern., 27, 101-110
- Larsson, J. E. (1992) Model-based Measurement Validation using MFM, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA, Seite 127-132
- Lee, I. B.; Minseok, K.; Juneyoung, J.; Chang, K. S. (1992) Rule-Based Expert System for Diagnosis Energy Distribution in Steel Plant, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA, Seite 281-286
- Locher, G.; Duff, M.; Stephanopoulos, G. (1995) Pattern Recognition Methods for Fermentation Database Mining, in: Munak, A.; Schügerl, K. (eds.) Proc. 6th Intern. Conf. on Comp. Applic. in Biotechnol., Garmisch-Partenkirchen, Germany, Seite 195-198

- Löhn, A. (1997) Wissensbasierte Systeme zur Methodenentwicklung und zur Analyse von Kalibrationen in der Fließinjektionsanalyse, Dissertation, Universität Hannover
- Ludewig, D. (1998) Expertensystem zur Entwicklung von Prozeßmodellen für biotechnische Prozesse, Dissertation, Universität Hannover
- Mandenius, C. F.; Polakovic, M. (1994) Multipurpose use of a Knowledge-Based Simulation Module applied to a Bioprocess, *Chem. Biochem. Eng. Q.*, 8 (4) 189-193
- Mason, A. (1993) Decisions Abound, Success is Manifest, *Proc. Ind. Journ.*, 1, 43-45
- Morris, A.J.; Montague, G. A.; Tham, M. T.; Aynsley, M.; Massimo, D.; Lant, P. (1991) Towards Improved Process Supervision, in: *Chemical Process Control (CPC IV)*, South Padre Island, Texas, USA
- Nakajima, M.; von Numers, C.; Yada, H.; Siimes, T.; Pokkinen, M.; Endo, I.; Linko, P. (1994) An On-line Advisory Control System for the Lactic Acid Fermentation Process, *Appl. Microbiol. Biotechnol.*, 42, 204-211
- Nicolov, L.; Konstantinov, K.; Yoshida, T. (1994) Influence of Information Noise on Inference Procedures of Fuzzy Expert Systems for Fermentation Control and Management, *Bioproc. Engng.*, 10 (4), 167-172
- Numers, v. C.; Nakajima, M.; Siimes, T. (1994) A Knowledge Based System using Fuzzy Inference for Supervisory Control of Bioprocesses, *Jour. of Biotechnol.*, 34 (2), 109-118
- Peters, H. (1996) Anwendung künstlicher neuronaler Netze in Überwachungssystemen, *atp*, 38 (11), 37-47
- Poletykin, A. (1994) A Method for Computer-aided Design of Rule-based Systems for Plant Process Diagnosis, *Engng. Applic. Artif. Intell.*, 7 (6), 665-675
- Posten, C. (1992) Phase Detection and Classification in Biotechnological Processes, *Process Control and Quality*, 2, 385-393
- Posten, C.; Gollmer, K. (1995) Pattern Recognition for Phase Detection in Bioprocesses, in: Munack, A.; Schügerl, K. (eds.), *Proc. 6th Intern. Conf. on Comp. Applic. in Biotechnol.*, Garmisch-Partenkirchen, Germany, Seite 41-46
- Roca, E.; Flores, J.; Rodriguez, I.; Cameselle, C.; Nunez, M. J.; Lema, J. M. (1996) Knowledge Based Control applied to Fixed Bed Pulsed Bioreactors, *Bioproc. Engng.*, 14 (3), 113-118
- Rodd, M. G.; Verbuggen, H. B.; Krijgsman A. J. (1992) Artificial Intelligence in Real-time Control, *Engng. Applic. Artif. Intell.*, 5 (5), 385-399
- Rojas-Guzman, C.; Kramer, M. A. (1993) Comparison of Belief Networks and Rule-based Expert Systems for Fault Diagnosis of Chemical Processes, *Engng. Applic. Artif. Intell.*, 6 (3), 191-202

- Rowan, D. A. (1992) Beyond Falcon: Industrial Applications of Knowledge-Based Systems, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA, Seite 249-251
- Saelid, S.; Mjaavatten, A.; Fjalestad, K. (1992) An Object oriented Operator Support System based on Process Models and an Expert System Shell, Europ., Comput. Chem. Eng., 16 (Suppl.), 97-108
- Saner, U.; Stephanopoulos, G. (1992) Application of Pattern Recognition Techniques to Fermentation Data Analysis, in: Karim, M. N.; Stephanopoulos, G. (eds.) IFAC Mod. and Contr. Of Biotechnol. Proc., Colorado, USA, 10, Seite 123-128
- Schmidt, G.; Rathert, H.; Perschl, F. (1996) Ein Überwachungssystem zur Fehlerdiagnose in verfahrenstechnischen Anlagen, atp, 38 (1), 45-51
- Seeger, M. (1997) Untersuchungen zur Integration von Expertenwissen in ein System zur Modellierung, Überwachung und Regelung von bioverfahrenstechnischen Prozessen II, Abschlußbericht, Braunschweig
- Simutis, R.; Havlik, I.; Lübbert, A. (1993) Fuzzy-Aided Neural Network for Real-Time State Estimation and Process Prediction in a Production-Scale Beer Fermentation, J. of Biotech., 27, 203-215
- Stanley, G. M.; Fraleigh, S. P.; Finch, F. E. (1992) Integrating Dataflow and Sequential Control in a Graphical Diagnostic Language, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA, Seite 101-108
- Stender, J. (1989) Wissenserhebung und -strukturierung in Expertensystemen, Markt- und Technik-Verlag, Haar bei München
- Stephanopoulos, G.; Bakshi, B. R. (1992) Temporal Representation of Process Trends for Diagnosis and Control, in: Karim, M. N.; Stephanopoulos, G. (eds.) IFAC Mod. and Contr. of Biotechnol. Proc., Colorado, USA, 10, Seite 109-112
- Stephanopoulos, G.; Locher, G.; Duff, M. (1995) Pattern Recognition Methods for Fermentation Database Mining, in: Munack, A.; Schügerl, K. (eds.) Proc. 6th Intern. Conf. On Comp. Applic. in Biotechnol., Garmisch-Partenkirchen, Germany, Seite 195-198
- Steyer, J. P.; Queinnec, I.; Capit, F.; Pourciel, J. B. (1996) Qualitative Rules as a way to Handle the Biological State of a Fermentation Process : An Industrial Application, in: Gentil, S. (ed.) Europ. Journ. of Automat., 30 (2-3), 381-398
- Suzuki, T.; Sakino, Y.; Nakajima, M.; Asama, H.; Fujii, T.; Sato, K.; Kaetsu, H.; Endo, I. (1997) A Novel Man-Machine Interface for a Bioprocess Expert System Constructed for Cooperative Decision Making and Operation, Journ. of Biotechnol., 52, 277-282
- Thompson, B.; Thompson, W. (1986) Finding Rules in Data, BYTE, Nov., 149-158

- Venkatasubramanian, V.; Rengaswamy, R. (1992) An Integrated Frameworks for Process Monitoring, Diagnosis and Control using Knowledge-based Systems and Neural Networks, in: Dhurjati, P. S.; Stephanopoulos, G. (eds.) On-line Fault Detection and Supervision in the Chemical Process Industries, Delaware, USA, Seite 49-54
- Weber, R.; Schimpe H. (1992) Wissensbasierte Datenanalyse mit Fuzzy Logik, Intell. Software Technol., 4,12-18
- Zöllner, B.; Mänzel S. (1996) Frühzeitige technisch-wirtschaftliche Beurteilung von Entwicklungsprojekten technischer Diagnoseexpertensysteme, atp, 38 (7), 36-44

Lebenslauf

Zur Person

Name	Russell Christian Gomersall
Geburtsdatum	10.06.1968
Geburtsort	Shornecliff, GB
Familienstand	ledig

Ausbildung

1974 – 1978	Grundschule Bad Oeynhausen
1978 – 1988	Immanuel-Kant-Gymnasium Bad Oeynhausen
Juni 1988	Abitur
Oktober 1988	Beginn des Chemiestudiums an der Universität Hannover
Oktober 1990	Diplom-Chemiker-Vorprüfung
Oktober 1993	Diplom-Chemiker-Hauptprüfung
Oktober 1993 – Mai 1994	Diplomarbeit, angefertigt im Institut für Technische Chemie Hannover
Juli 1994	Beginn der vorliegenden Dissertation