

# Intelligente leittechnische Konzepte für die Bioverfahrenstechnik: Einsatz und Perspektiven an der TH Darmstadt

Herrn Prof. Dr. rer. nat. Dipl.-Ing. Henning Tolle zum 65. Geburtstag gewidmet

Kurt Dirk Bettenhausen, Frankfurt am Main, Stefan Gehlen, Bochum und Peter Marenbach, Darmstadt



**Dr.-Ing. Kurt Dirk Bettenhausen** ist Mitarbeiter im Bereich *Corporate Research and Technology* der Hoechst AG, Frankfurt am Main. Davor war er wissenschaftlicher Mitarbeiter am Institut für Regelungstechnik der TH Darmstadt, FG Regelsystemtheorie & Robotik (Leiter: Prof. Dr. rer. nat. Dipl.-Ing. Henning Tolle). Aus dieser Zeit stammen die in diesem Beitrag vorgestellten Untersuchungen und Ergebnisse.

Adresse: Hoechst AG, Produktionstechnologien, Gebäude D 710, Technologiegruppe Leittechnik, D-65926 Frankfurt am Main, Tel.: (069) 305-83858, Fax: (069) 305-17212, E-Mail: bettenhaus@msmtemk.hoechst.com



**Dr.-Ing. Stefan Gehlen** leitet den Bereich Prozeßführung bei der Zentrum für Neuroinformatik GmbH, Bochum (ZN). Hauptarbeitsfelder: Entwicklung und Realisierung von intelligenten Automatisierungslösungen, Neuronale Netze zur Prozeßregelung, Fuzzy Control.

Adresse: Zentrum für Neuroinformatik GmbH, Am Botanischen Garten, Universitätsstraße 160, D-44801 Bochum, Tel.: (0234) 97870, Fax: (0234) 978777, E-Mail: gehlen@zn.ruhr-uni-bochum.de



**Dipl.-Ing. Peter Marenbach** ist wissenschaftlicher Mitarbeiter am Institut für Regelungstechnik der TH Darmstadt, FG Regelsystemtheorie & Robotik (Leiter: Prof. Dr. rer. nat. Dipl.-Ing. Henning Tolle). Er beschäftigt sich mit der Untersuchung und Entwicklung datenbasierter Modellbildungstechniken sowie deren Anwendung zur Führung biotechnologischer Prozesse.

Adresse: TH Darmstadt, Institut für Regelungstechnik, FG Regelsystemtheorie & Robotik, Landgraf-Georg-Straße 4, D-64283 Darmstadt, Tel.: (06151) 16-3742, Fax: (06151) 16-2507, E-Mail: peter.marenbach@rt-technik.th-darmstadt.de

*Die technologische Nutzung von Stoffwechselforgängen lebender Organismen erfordert den Einsatz von Methoden der Prozeßleittechnik, deren Leistungscharakteristika weit oberhalb derer konventioneller Ansätze zur Prozeßführung liegen. Dies ist einer der Gründe dafür, warum in der industriellen Praxis auch heute noch viele Aufgaben von Anlagenfahrern wahrgenommen werden. In der vergangenen Dekade wurden an der TH Darmstadt Methoden entwickelt, welche die bisher dem Menschen vorbehaltenen intelligenten Aufgaben der Prozeßüberwachung und -optimierung automatisieren. Diese Verfahren sind im Begriff, sich*

*im großindustriellen Maßstab zu etablieren. Der vorliegende Artikel gibt einen Überblick über die zugrunde liegenden Arbeiten sowie die bestehenden Perspektiven.*

## Intelligent methods for process control: Use and perspectives in biochemical engineering

*The technological usage of metabolims of living cells requires innovative methods of automatic process control performing better than conventional approaches are able to. This is one of the reasons why a lot of tasks during industrial fermentations today are done by human operators. During the last decade methods for automatic process control and optimisation – which are tasks that were previously only solvable by humans – have been developed. These approaches are nowadays on the way to establish in large-scale industrial applications. The article at hand gives an overview of the fundamental work and existing perspectives for the future.*

## 1 Einleitung

Die Biotechnologie wird heute als eines der Schlüsselfelder künftiger technologischer Innovation angesehen. Mit Hilfe biochemischer Prozesse kann eine Vielzahl komplexer chemischer Substanzen synthetisiert oder abgebaut werden. Der Weg der technologischen Nutzung des Stoffwechsels lebender Organismen von der Erforschung der Grundlagen in Biochemie und Molekularbiologie bis hin zur großindustriellen Produktion ist wesentlich durch die erforderliche interdisziplinäre Arbeitsweise gekennzeichnet. Hierbei reicht es nicht aus, lediglich das Verhalten einzelner Zellen zu betrachten, sondern es muß – gerade für die Übertragung der Erkenntnisse vom Labor- zum Produktionsmaßstab sowie insbesondere aus Gründen der Wirtschaftlichkeit – die rechtzeitige Mitwirkung von Verfahrenstechnik, Prozeßtechnik und Engineering sichergestellt werden.

Ein grundlegendes Problem für die automatisierte Prozeßführung ist, daß wesentliche, für den Stoffwechsel relevante Größen nicht mit vertretbarem Aufwand meßbar sind. Vielfach können Meßwerte nur durch aufwendige Laboranalysen bestimmt werden und stehen erst mit großem zeitlichen Verzug zur Verfügung. Gleichzeitig führt die experimentelle Variation der Randbedingungen des Prozesses (wie z.B. Nährmedienzusammensetzung oder physikalische Umgebungsparameter) zu deutlich unterschiedlich ausgeprägtem Prozeßverhalten, das sich zwar in den ermittelten Meßdaten widerspiegelt, aber nicht direkt zum Entwurf und zur Umsetzung von Prozeßführungsstrategien genutzt werden kann.

In der Praxis werden biotechnologische Produktionsprozesse heutzutage größtenteils von Anlagenfahrern geführt und unterliegen damit zwangsläufig den natürlichen Schwankungen menschlicher Leistungsfähigkeit. Dies kann zu Problemen bezüglich der Produktqualität und Reproduzierbarkeit der Produktion, aber auch bezüglich der Zulassung des Produktes führen. Es ergibt sich die Frage, in welchen Formen Wissen über den Prozeß automatisch gewonnen und erweitert werden kann, um somit zur Überwachung und Optimierung des Prozesses und damit auch zu einer Erhöhung der Wirtschaftlichkeit zu gelangen. Diese Fragestellung wurde in den vergangenen Jahren intensiv am Fachgebiet Regelsystemtheorie & Robotik der TH Darmstadt untersucht.

Der vorliegende Beitrag zeigt, welche Werkzeuge die Automatisierungstechnik für die Optimierung und sichere Prozeßführung biologischer Verfahren bereits heute anbieten kann. Zunächst wird auf die Modellbildung, später auf die Optimierung von Fermentationsprozessen eingegangen. Es werden mehrere Realisierungen „intelligenter“ Systeme vorgestellt, die an der TH Darmstadt entwickelt wurden oder aus den dort gewonnenen Erfahrungen hervorgegangen sind, und die Ergebnisse bei der Prozeßführung mit diesen Systemen erläutert.

## 2 Modellbildung und Wissensrepräsentation bei biologischen Prozessen

### 2.1 Grundlegende Überlegungen

Bei der Modellbildung dynamischer Prozesse in der Regelungs- und Automatisierungstechnik kann man prinzipiell folgende Ansätze unterscheiden:

1. Physikalische Überlegungen liefern mathematische Modellstrukturen, deren Parameter entweder theoretisch abgeleitet oder datengetrieben anhand von Messungen am Prozeß ermittelt werden. Dieses Vorgehen läßt sich mit Hilfe von Programmsystemen zur Prozeßidentifikation und Datenanalyse unterstützen.
2. Bei der experimentellen Prozeßidentifikation werden die Parameter einfacher Modellstrukturen daten-

getrieben so bestimmt, daß das Verhalten des geschlossenen Regelkreises zufriedenstellend ist. Bei der *adaptiven Regelung* erfolgen diese Bestimmung und der darauf aufbauende automatische Reglerentwurf on-line am Prozeß.

3. Die funktionale Abbildung von Ein- und Ausgangsmeßgrößen wird (z.B. in *neuronalen Netzen*) allgemein gespeichert, wobei durch Interpolation im betrachteten Meßbereich das Prozeßverhalten mit ausreichender Genauigkeit wiedergegeben wird.
4. Wissen über den Prozeß, das nur in qualitativer Form vorliegt, kann z.B. in Form von Wissensbasen (*Regeln, Fuzzy-Logik*) repräsentiert werden.

Für die detaillierte Prozeßmodellierung sind lediglich der erste und der dritte Ansatz relevant, wobei der theoretisch motivierte, physikalische Ansatz den Vorteil hat, anschaulich und transparent zu sein, während die interpolierende Speicherung auch komplexe Zusammenhänge automatisch ohne Strukturannahmen wiederzugeben vermag. Der vierte Ansatz ist für die Modellbildung aufgrund der gewünschten Genauigkeit oft nicht geeignet, er ist hingegen erfolgversprechend, wenn das Wissen eines Experten als Grundlage für die Prozeßführung repräsentiert werden soll.

Grundsätzlich gelangen bei technologisch genutzten Bioprozessen alle genannten Methoden zum Einsatz. In den folgenden Abschnitten werden die an der TH Darmstadt verfolgten Ansätze sowie wichtige Erweiterungen mit dem konkreten Blickwinkel auf biotechnologische Fragestellungen diskutiert. Zunächst wird jedoch ein Strukturierungsschema für biologische Prozesse angesprochen, das den notwendigen Aufwand zur Modellierung und Prozeßführung deutlich reduziert.

### 2.2 Phasenmodell zur Prozeßstrukturierung

Der zentrale Prozeß bei der biotechnologischen Produktion ist die Fermentation. Bei Batch-Fermentationen wird das Fermentationsmedium, welches die notwendigen Kohlenstoff-, Stickstoff- und Phosphorquellen enthält, zu Beginn sterilisiert und anschließend mit Zellen beimpft. Das Wachstum und die Produktion im Bioreaktor sind durch mehrere charakteristische Abschnitte gekennzeichnet:

- Zunächst erfolgt eine Adaption der Zellen an das neue Milieu, in dieser *Adaptionsphase* erfolgt kein Wachstum der Zellen.
- Nach der Adaption beginnen die Zellen, sich mit maximaler Geschwindigkeit zu vermehren (*exponentielles Wachstum*) und die Nährstoffe zu verbrauchen.
- Sind wichtige Substrate aufgebraucht oder limitiert (meist der verfügbare Sauerstoff oder die Kohlenstoffquelle), so stellt sich ein Gleichgewicht zwischen Zellwachstum und -sterben ein. In dieser *stationären Phase* ist oftmals die maximale Produktbildung zu beobachten.
- Nach der Erschöpfung der Energiereserven der Zellen beginnt die *Absterbephase*. Oft wird die Fermentation an dieser Stelle abgebrochen, wenn auch die Produktbildung zum Erliegen kommt.

Weitere Unterteilungen sind möglich, z.B. wenn sequentiell mehrere Substrate verbraucht werden. Beim Übergang von einer Phase zur nächsten stellt sich der Zellmetabolismus auf eine veränderte Umgebungssituation ein; Faktoren, die zuvor wichtig waren, können ihre Bedeutung verlieren und umgekehrt. Das Phasenmodell bietet sich – unabhängig von der gewählten Modellierungsmethode – als Strukturierungshilfe für die Prozeßmodellierung und die darauf aufbauende Regelung bzw. Optimierung an. Da in jedem Abschnitt andere signifikante Ereignisse betrachtet werden müssen, können für jede Phase vereinfachte Modelle angesetzt werden. Die Tatsache, daß durch den Ansatz auch eine auf den physiologischen Zustand der Zellen abgestimmte Modellierung und Regelung erfolgt, wird durch den von *Konstantinov* und *Yoshida* geprägten Begriff der *physiological state control* [1] auch anschaulich wiedergegeben.

### 2.3 Datengetriebene Generierung strukturierter Prozeßmodelle

An der TH Darmstadt wurde in den letzten drei Jahren ein Ansatz entwickelt, der die Vorteile der mathematisch strukturierten Verfahren mit denen datengetriebener verbindet. Mit dem Prinzip des *genetischen Programmierens* [2] wird versucht, dominante mathematisch-physikalische Zusammenhänge in einer symbolischen Notation automatisch zu formulieren und so zu verändern, daß ein Modell entsteht, welches eine möglichst gute Übereinstimmung mit gemessenen Verläufen zeigt.

Das grundlegende Konzept besteht darin, die iterative Vorgehensweise eines Prozeßingenieurs bei der Modellbildung komplexer Prozesse im Computer zu imitieren. Dazu soll soviel mathematisches Wissen wie möglich (z.B. über grundlegende Bilanzen) in Form von Teilmodellen oder Modellideen genutzt werden. Für die Verfeinerungen und Ergänzungen, die notwendig sind, um die gewünschte Modellgenauigkeit zu erreichen – die auch der Mensch nur durch gezieltes Ausprobieren findet –, werden im Rechner datengetriebene Vorschläge generiert. Bild 1 zeigt den prinzipiellen Ablauf der datengetriebenen Struktursuche und gibt an, wie dieser Prozeß vom Vorwissen des Anwenders auf der einen und vom vorhandenen Datenmaterial auf der anderen Seite beeinflusst wird. Dabei unterscheidet der Algorithmus zwischen zwei Aufgaben: Die Identifikation der Parameter, die in einer Modellstruktur enthalten sind, sowie die neuartige Suche nach einer geeigneten Modellstruktur, die auf einer symbolischen Ebene mit Hilfe des genetischen Programmierens realisiert wird.

Der Ablauf der datengetriebenen Modellstruktursuche läßt sich wie folgt beschreiben: Das evolutionäre Verfahren benötigt zunächst eine Anfangspopulation, d.h. eine Menge möglichst verschiedener Modelle, die zufällig aus den im sogenannten Modellbaukasten zusammengefaßten Übertragungsgliedern zusammengesetzt wird. Alternativ können auch Modellvorschläge vom Benutzer übernommen werden. In der Folge wird ein aus vier Schritten bestehender Zyklus solange

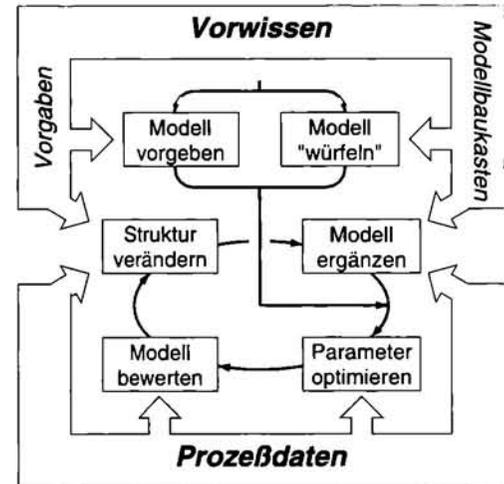


Bild 1: Funktionsschema der datengetriebenen Generierung strukturierter Prozeßmodelle.

durchlaufen, bis ein Abbruchkriterium erfüllt ist. Darin erfolgt zunächst die numerische Behandlung der Modellpopulation, bei der jedes Modell durch Einstellung seiner Parameter optimal an das gewünschte Prozeßverhalten angepaßt und anschließend bewertet wird. Bei dieser Bewertung werden neben dem verbliebenen Fehler auch die Komplexität und die Generalisierungsfähigkeit des Modells betrachtet. In Anlehnung an die natürliche Selektion der Arten wird nun eine neue Generation von Modellen generiert, wobei der Anteil der Nachkommen eines Modells um so höher ist, je besser es zuvor bewertet wurde. Schließlich werden die Modellstrukturen der Nachkommen zufällig mutiert und rekombiniert; anschließend beginnt der Zyklus von vorne.

Dieser Ansatz gibt dem Prozeßentwickler Hinweise auf die möglichen physikalischen oder auch (bio-) chemischen Zusammenhänge. Er erlaubt auch, vorhandenes Prozeßwissen und grundlegende Bilanzgleichungen in den automatisierten Modellbildungsprozeß einzubringen und liefert somit ein datengetriebenes entwickeltes Modell, das aufgrund seiner Transparenz nachvollziehbar bzw. kritisch bewertbar ist. Bisherige Untersuchungen zeigten, daß auch ein vergleichsweise „global“ suchender genetischer Algorithmus nur schwer zu einer eindeutigen Lösung kommt, wenn nur wenig den Prozeß beschreibendes Datenmaterial verfügbar ist. Dies bedeutet zum einen, daß es bei der Anwendung von Bedeutung ist, tatsächlich soviel Vorwissen wie möglich zu verwenden, und zum anderen die gelieferten Ergebnisse kritisch – auch mit Blick auf eventuell unzureichendes Datenmaterial – zu betrachten.

### 2.4 Prozeßmodellierung mit neuronalen Netzen

Künstliche neuronale Netze sind dem biologischen Nervensystem nachempfundenen Strukturen, deren Verarbeitungsleistung auf der massiv parallelen Verschaltung einfacher Elementarprozessoren (Neuronen) beruht. Ein Neuron gewichtet jedes eingehende Signal und summiert die resultierenden Werte. Das so berechnete Gesamtsignal wird über eine nichtlineare Aktivierungsfunktion auf den Ausgang des Neurons

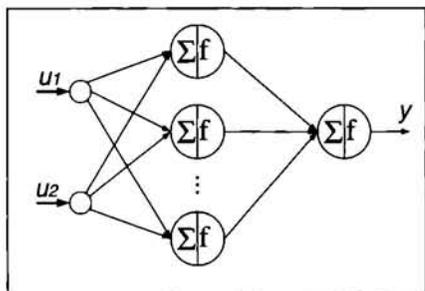


Bild 2: Grundarchitektur eines neuronalen Netzes mit drei Schichten.

geschaltet (Bild 2). Bei der Prozeßmodellierung wird mit einigen Standardtypen neuronaler Netze gearbeitet. Die wichtigsten Typen sind im folgenden kurz zusammengefaßt (eine detaillierte Übersicht findet sich z.B. in [3]).

- Das *Multilayer-Perceptron* ist ein schichtenweise organisiertes neuronales Netz. Die Signale werden von den Neuronen einer Schicht zu allen Neuronen der nächsten geschaltet. Jedes Neuron besitzt eine sigmoide Aktivierungsfunktion.
- Das Netz mit radialen Basisfunktionen (*RBF-Netz*) ist ebenfalls ein rein vorwärtsgerichtetes Netz mit drei Schichten, bei dem die Neuronen der zweiten Schicht lokale Aktivierungsfunktionen (z.B. Gauss-Glocken) besitzen. Die Neuronen der Ausgangsschicht bilden dagegen lineare Einheiten.
- Interpolierende Funktionsspeicher, wie der *CMAC-Speicher*, operieren wie das RBF-Netz mit lokal begrenzten Aktivierungsfunktionen und zeichnen sich durch besonders schnelle Adaptionsverfahren aus [4]. Sie sind besonders für ein on-line Lernen parallel zum Prozeß geeignet.

Aus mathematischer Sicht bildet ein neuronales Netz ein nichtlineares Funktionssystem, bei dem die freien Gewichtungsfaktoren über ein Optimierungsverfahren so an die vorgegebenen Prozeßmeßwerte adaptiert werden, daß eine minimale Abweichung zwischen tatsächlichen und geschätzten Werten resultiert. Ein neuronales Netz kann die eingehenden Informationen „verallgemeinern“: Für vorher nicht exakt gemessene Werte wird in der Applikationsphase der Ausgang approximiert. Im Gegensatz zu den mathematisch-physikalisch orientierten Verfahren müssen für den Einsatz eines neuronalen Netzes zur Prozeßmodellierung nur wenige Strukturvorgaben gemacht werden.

Die größte Stärke von neuronalen Netzen liegt darin, daß sie einen universellen Ansatz zur Modellbildung auch stark nichtlinearer Prozesse darstellen. Der größte Nachteil ist, daß sie *black-box*-Modelle sind, die dem Anwender zunächst keine Rückschlüsse auf die innere Struktur eines Prozesses erlauben.

## 2.5 Formulierung von Prozeßführungsstrategien in Form von Fuzzy-Regeln

Wissen über den Prozeß spiegelt sich aber nicht nur in der Analyse des Prozeßverhaltens und den daraus resultierenden Prozeßvorstellungen und -modellen wider, sondern auch in den Strategien, mit denen erfahrene Anlagenbediener den Prozeß führen. Neben

den unbewußt ablaufenden Vorgängen können verbal formulierbare Regeln vom Bediener erfragt werden. Die Berücksichtigung der von klassischen Expertensystemen bekannten Probleme bei der Wissensakquisition und der Unzulänglichkeiten der zweiwertigen Logik zur Beschreibung von Phänomenen in Übergangsbereichen legen die Verwendung der unscharfen *Fuzzy-Logik* zur rechnerinternen Regelrepräsentation nahe, deren Leistungsfähigkeit durch die Integration von Lernverfahren weiter gesteigert werden kann.

## 3 Prozeßführung und -optimierung

### 3.1 Einfache experimentelle Optimierung unter Nutzung des Phasenkonzeptes

Wie schon in Abschnitt 2.2 ausgeführt, ist der zeitliche Verlauf von Batch- und Fed-Batch-Fermentationen durch eine Abfolge von spezifischen Phasen gekennzeichnet, welche jeweils den physiologischen Zustand der Zellen widerspiegeln. Vielfach lassen sich durch experimentelle Optimierung günstige Steuerungsparameter für jeden Prozeßabschnitt festlegen, z.B. die auf die jeweilige Prozeßphase abgestimmte Fütterungsrate beim Fed-Batch-Betrieb. Betreibt man ein Klassifikationssystem am Prozeß, so können zustandsabhängig die Sollwerte der unterlagerten Regelungen modifiziert werden. Vorteilhaft ist es, das Klassifikationssystem als Fuzzy-System aufzubauen. In diesem Fall wird das „scharfe“ Schalten zwischen den physiologischen Zuständen verhindert und das Sollwertsignal kann in Abhängigkeit der Signifikanz der erkannten Abschnitte ermittelt werden. Gegenüber den in der chemischen Industrie vorherrschenden zeitbasierten Profilen wird eine robuste, zustandsabhängige Steuerung realisiert. Die Variabilität biologischer Prozesse, z.B. die Dauer der Adaption der Zellen nach Animpfung des Fermenters, kann durch den genannten Ansatz einfach berücksichtigt werden.

### 3.2 Lernende Prozeßführung und -optimierung

Die grundlegende Idee der lernenden Regelung kann man am Beispiel des in Bild 3 dargestellten lernenden Regelkreises LERNAS [4] erkennen. Im wesentlichen sind darin zwei Mechanismen gleichzeitig aktiv: Zum einen wird in einem dem Prozeß parallelgeschalteten assoziativen Speicher (neuronalen Netz) ein prädiktives Prozeßmodell erlernt. Zum anderen dient ein Optimierungsmodul dazu, gestützt auf das prädiktive Modell eine optimale Regelungsstrategie zu berechnen, die auf den Prozeß geschaltet und für spätere Anwendungen in einem lernfähigen Reglerspeicher abgelegt wird. Die Modellbildung wie auch die Optimierung können dabei sowohl on-line als auch off-line basierend auf gespeicherten Prozeßdaten erfolgen.

Da Stellgrößen im prädiktiven Modell explizit berücksichtigt werden, ist es möglich, Vorhersagen über die Auswirkungen einer Änderung der Umgebungsparameter zu machen, um so optimale Sollwerte zu ermitteln. Dies ist ein wesentlicher Vorteil gegenüber adaptiven Ansätzen, bei denen die zu optimieren-

den Umgebungsparameter nur implizit in den Modellparametern berücksichtigt werden (z.B. [5]).

### 3.3 Integration von wissensbasierter und lernender Regelung

Ein erster Schritt zur Anpassung des Grundgerüsts der lernenden Regelung auf die spezifischen Eigenschaften biochemischer Prozesse wurde in [6] vorgestellt. Das dabei verfolgte Konzept beinhaltet eine Reihe von Erweiterungen des zugrundeliegenden lernenden Ansatzes mit dem Ziel, wissensbasierte und lernende Methoden zur Prozeßführung zu integrieren (Bild 4).

Da es bei Fermentationsprozessen weniger das Problem ist, Sollwertvorgaben für Umgebungsparameter (Temperatur, pH) am Prozeß einzustellen, sondern vielmehr die Frage interessiert, welche zeit- oder zustandsabhängigen Sollwertprofile für den betrachteten Prozeß günstig sind, setzt die lernende Regelung auf einer unterlagerten konventionellen Regelungsschicht auf.

Die in Abschnitt 2.2 beschriebene Strukturierung des Prozesses ist für die wissensbasierte und lernende Prozeßführung hilfreich, weil in jeder Phase einfachere Modelle, die jeweils nur eine Untermenge aller Meßgrößen berücksichtigen müssen, gebildet werden können. Für jede Phase kann ein eigenes Optimierungsziel definiert werden, so daß z.B. zu Beginn des Prozesses das Zellwachstum und später die Stabilität des gebildeten Produktes gewichtet wird. Die Erkennung der jeweils aktuellen Phase wird durch einen regelbasierten Klassifikator realisiert.

Schließlich wurde zur Überwachung der lernenden Prozeßführung neben einer grundsätzlichen Fehlerüberwachung (Ausfall von Anlagenkomponenten, Erkennung von Kontaminationen etc.) eine zusätzliche Plausibilitätskontrolle der von der lernenden Regelung vorgeschlagenen Stelleingriffe realisiert.

Das Zusammenwirken von lernenden und wissensbasierten Komponenten ist charakteristisch für intelligente Automatisierungssysteme [7]. Bei der Generierung von Prozeßführungsstrategien will man das heuristische Wissen des Anlagenfahrers nutzen und gleichzeitig im lernenden Regelkreis optimieren. Die Gestaltung der Schnittstelle zwischen lernender und wissensbasierter Ebene ist für den Erfolg dieses Bemühens von zentraler Bedeutung. Im folgenden Abschnitt wird daher ein alternativer Ansatz zur Repräsentation der Prozeßführungsstrategie beschrieben, der die symbolische und die numerische Ebene der Informationsverarbeitung besser verbindet, indem sowohl heuristische Strategien genutzt als auch optimierte Strategien linguistisch interpretiert werden können.

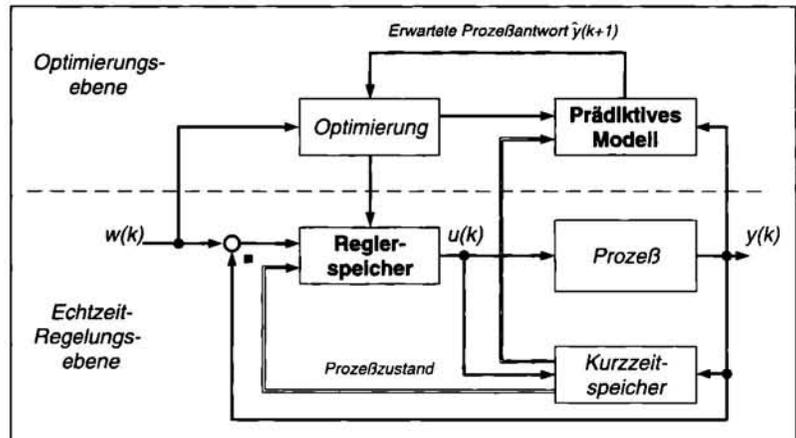


Bild 3: Blockschaltbild des lernenden Regelkreises LERNAS.

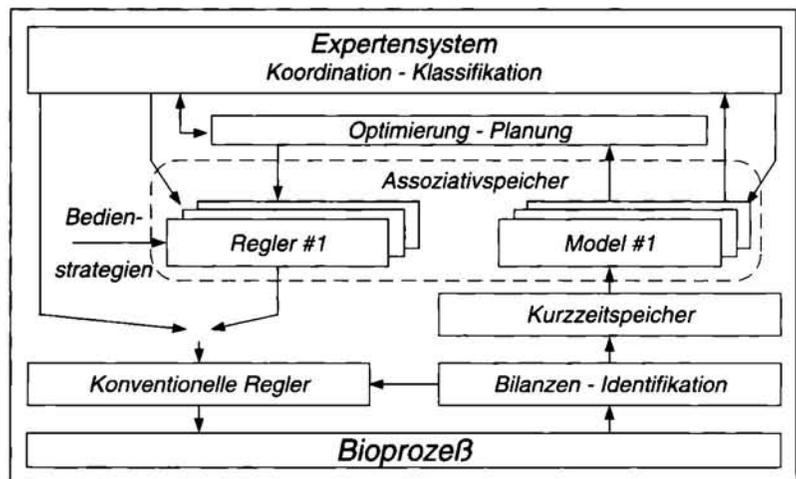


Bild 4: Architektur des Systems BioX zur integrierten wissensbasierten und lernenden Prozeßführung.

### 3.4 Transparente Generierung von Regelungsstrategien

Die Möglichkeit, auch umfangreiches Erfahrungswissen für eine automatische Prozeßführung zu nutzen, ist die treibende Kraft für die große Popularität von Fuzzy-Reglern. Sowohl Fuzzy-Systeme als auch neuronale Netze repräsentieren das in ihnen enthaltene Wissen durch die gewichtete Überlagerung nichtlinearer Basisfunktionen. Diese Ähnlichkeit, die sich auch in der internen Repräsentation wiederfindet (siehe z.B. [8]), führte zur Entwicklung lernender Fuzzy-Systeme, die oft als *Neuro-Fuzzy* bezeichnet werden.

Fuzzy-Systeme bieten eine Reihe von Möglichkeiten für die Veränderung und Optimierung ihres Ein-Ausgangs-Verhaltens. Ein Ansatz, der ein hohes Maß an Flexibilität bei gleichzeitiger Beibehaltung der linguistischen Bedeutung der Regeln bietet, ist die Einführung von Regelgewichten im *vollständig definierten Regelraum* [9].

Mit der Einführung von Regelgewichten (vgl. Bild 5 und Bild 6) ist es möglich, den Entwurf eines Fuzzy-Reglers – wie nachfolgend dargestellt – Schritt für Schritt durchzuführen und automatisch zu verbessern:

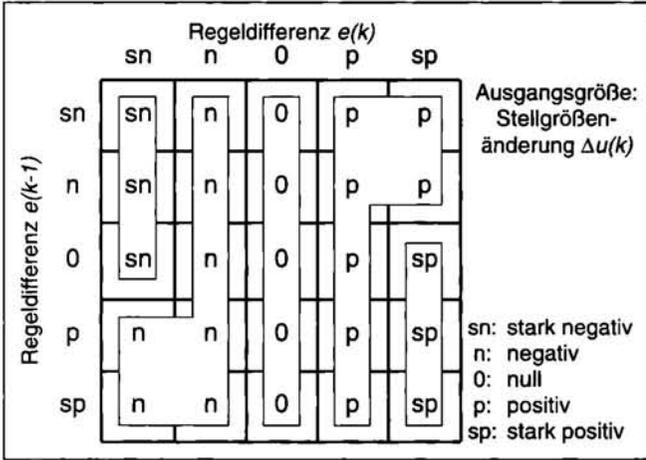


Bild 5: Regelmatrix für zwei Eingänge  $e(k-1)$  und  $e(k)$  und einen Ausgang  $\Delta u(k)$ .

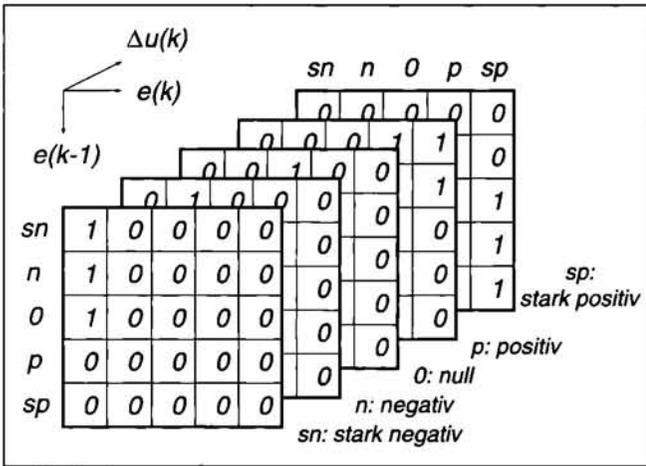


Bild 6: Darstellung im vollständig definierten Regelraum entsprechend Bild 5.

1. Erfasse vorhandenes Bedienerwissen und definiere die entsprechenden Zugehörigkeitsfunktionen sowie die Regelbasis.
2. Führe veränderbare Regelgewichte in die Regelbasis ein.

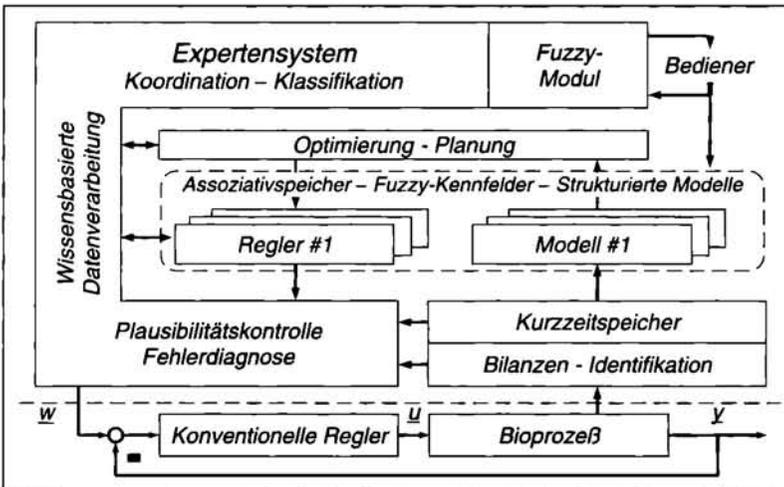


Bild 7: Struktur des Systems  $BioX^{++}$  zur Führung biotechnologischer Produktionsprozesse.

3. Optimierte die Regelbasis durch Veränderung der Regelgewichte. Dies kann on-line oder off-line geschehen.
4. Extrahiere die minimale Regelbasis und entferne die Gewichte.

Zur effizienten Optimierung der Regelgewichte in Schritt 3 wurden spezielle Algorithmen entwickelt [10]. Basierend auf dieser Vorgehensweise ermöglicht es der Neuro-Fuzzy Ansatz, die für den Bediener transparente Befolgung von linguistischen Regeln mit den aus der numerischen Optimierung gewonnenen Erkenntnissen zu verbinden.

## 4 Realisierte Systeme

### 4.1 BioX

Zur Prozeßführung biotechnologischer Verfahren wurde BioX (vgl. Bild 4) als Beispiel für ein integriertes wissensbasiertes und lernendes Pilotsystem in den Jahren 1989–1992 am Fachgebiet Regelsystemtheorie & Robotik der TH Darmstadt entwickelt [11].

Die Architektur von BioX ist zweigeteilt: Unterschieden werden das regelbasierte Kernsystem und die lernende Prozeßführung. Das Kernsystem enthält die Komponenten Wissensbasis, Wissenseditor und Erklärungskomponente und ist in der Programmiersprache LISP implementiert. Infolge der verbesserten numerischen Effizienz ist die Lernebene in der Programmiersprache C implementiert. Die wesentlichen Module in dieser Ebene beinhalten die programmtechnische Realisierung der neuronalen Kennfeldspeicher, die numerische Optimierung sowie die Prozeßanbindung. Beide Ebenen sind über ein Interface gekoppelt, wobei das Kernsystem die Kontrolle aller Komponenten der Lernebene innehat.

### 4.2 BioX<sup>++</sup>

Aufbauend auf den mit dem Vorgängersystem BioX gewonnenen Erfahrungen erfolgte eine Neuimplementierung unter Berücksichtigung objektorientierter Programmierparadigmen in der Programmiersprache C++. Damit konnte das Zusammenspiel von lernenden und wissensbasierten Verfahren verbessert werden. Der schrittweise Auf- und Ausbau des Systems ist nach Abschluß einzelner Entwicklungs- und Untersuchungsphasen in den Arbeiten [12] bis [14] beschrieben.

Bild 7 zeigt die Struktur des Systems sowie den Informationsfluß während der Verarbeitung. In  $BioX^{++}$  werden die unterschiedlichen Aufgaben blockweise abgearbeitet, wobei es die modulare, objektorientierte Implementierung gestattet, verschiedene Module gegeneinander auszutauschen. Zu den Aufgaben, die in einem Abtastschritt bewältigt werden, gehören

- Übernahme der Meßwerte und Kontrolle der Plausibilität,

- Fehlerdiagnose und Alarmbehandlung,
- Datenvorfilterung und Identifikation von Prozeßparametern sowie Auswertung von Bilanzgleichungen,
- Klassifikation des physiologischen Prozeßzustandes sowie Auswahl von phasenspezifischem Prozeßmodell und entsprechender Führungsstrategie,
- sofern zulässig, eine wissensbasierte und numerische Optimierung der Führungsstrategie auf der Basis on-line erweiterter Prozeßmodelle sowie
- abschließend die Plausibilitätskontrolle der neu erzeugten Stellgrößen, die als Sollwertvorgaben für die unterlagerten konventionellen Regelungen dienen.

### 4.3 SMOG

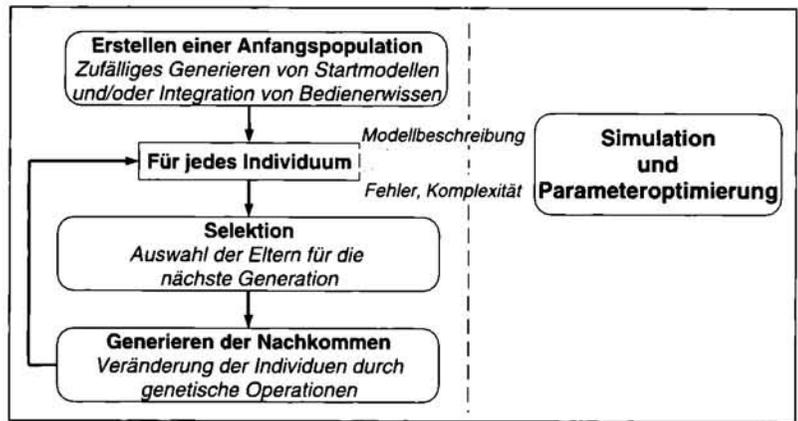
Die in Abschnitt 2.3 beschriebene datengetriebene Generierung strukturierter Prozeßmodelle wurde im Rahmen des Werkzeugs SMOG (*Structured Model Generator*) an der TH Darmstadt implementiert. Wie oben beschrieben, basiert dieser Ansatz auf einem hybriden Konzept, das die symbolische Struktursuche mit Genetischer Programmierung und die numerische Parameteroptimierung mit konventionellen Verfahren verbindet. Verdeutlicht wird die Trennung von der evolutionären Modifikation der zunächst symbolisch betrachteten Modellstruktur und der numerischen Suche nach optimalen Parametern für eine bestimmte Struktur in dem in Bild 8 dargestellten Ablaufdiagramm. Man erkennt links den typischen Ablauf eines evolutionären Optimierungsverfahrens und rechts die für jedes Modell ausgeführte Parameteradaptation. Da für die Parametereinstellung eine wiederholte Simulation des Modells not-

wendig ist, ist diese der mit Abstand rechen- und damit zeitintensivste Schritt des gesamten Ablaufs. Aus diesem Grund laufen diese Berechnungen im SMOG – unter Ausnutzung der inhärenten Parallelität evolutionärer Verfahren – effizient und ohne nennenswerten Synchronisationsaufwand verteilt auf mehreren Workstations ab.

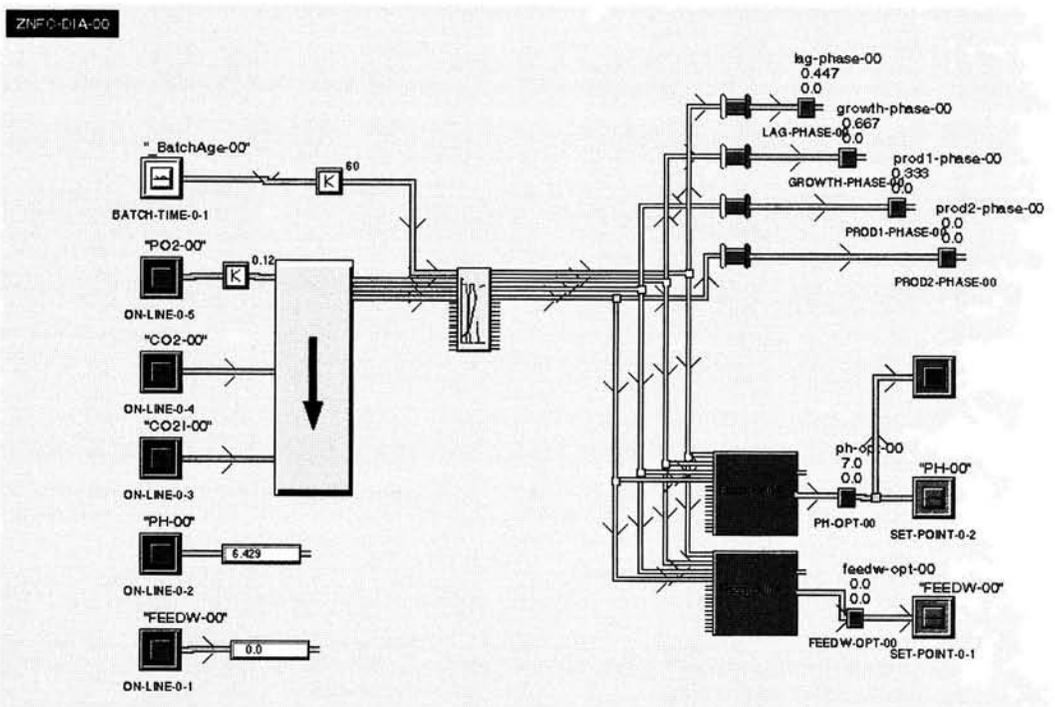
Nachdem mit SMOG zunächst Untersuchungen zur Modellierung einer Flußstauhaltung durchgeführt wurden [15], steht mittlerweile die Anwendung an Bioprozessen im Mittelpunkt. Dabei wird SMOG sowohl zur Erzeugung von sogenannten *Softwaresensoren* [16] (Modelle zur Schätzung on-line nicht-meßbarer Größen) als auch von umfangreichen Modellen des gesamten Prozesses eingesetzt.

### 4.4 ZNFC

Das Fermentationssteuerungssystem ZNFC wurde zunächst im Auftrag eines Chemiekonzerns von der



**Bild 8:** Hybrider Ansatz in SMOG: symbolische (links) und numerische Ebene (rechts).



**Bild 9:** Datenfluß-orientierte Realisierung einer Regelungsmethode mit ZNFC.

Zentrum für Neuroinformatik GmbH, Bochum, entwickelt. ZNFC setzt auf dem Echtzeit-Expertensystemwerkzeug G2/GDA der Firma Gensym auf. Dieses System zeichnet sich durch eine einfach zu bedienende graphische Benutzerführung, z.B. in Form von Daten- und Informationsflußdiagrammen (vgl. Bild 9), und eine konsequente objektorientierte Technologie aus. G2/GDA ist durch mehrere Installationen im Bereich der Bioprozeßführung bereits industriell eingeführt.

Die in G2/GDA gewählte Art der Mensch-Maschine-Interaktion wurde in ZNFC konsequent weiterverfolgt und auf die Regelung von Fermentationsprozessen zugeschnitten. ZNFC bietet für die Bioprozeßführung spezialisierte Methoden zur Fuzzy-Datenverarbeitung und zur flexiblen Einbindung neuronaler Netze. Über vorgefertigte Module wird eine transparente Datenverbindung zu bestehenden Automatisierungssystemen, z.B. zum Prozeßleitsystem MFCS/win der Firma B. Braun, Melsungen, realisiert. Im on-line Betrieb können mehrere Bioreaktoren zeitgleich geregelt werden, in einem speziellen Analysekanal lassen sich Regelungsschemata mit Daten bereits abgeschlossener Fermentationen überprüfen.

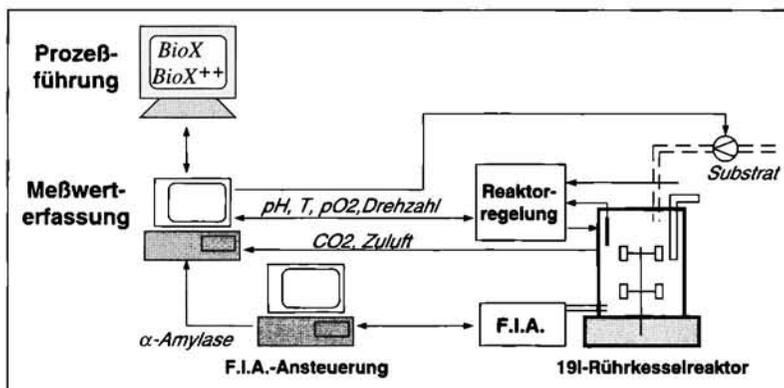
Mit ZNFC können alle in diesem Beitrag vorgestellten Regelungsmethoden zur Klassifikation, Modellbildung und Optimierung, die auf Fuzzy-Logik und neuronalen Netzen basieren, einfach realisiert werden.

## 5 Experimentelle Anwendungen

Die Leistungsfähigkeit der beschriebenen Regelungsansätze wurde über Jahre in Zusammenarbeit mit dem Institut für Biochemie der TH Darmstadt an einem industriell relevanten Prozeß, der Produktion von  $\alpha$ -Amylase mit *B. subtilis*, verifiziert.

Bild 10 zeigt schematisch den Aufbau der Pilotanlage. Der Rührkesselreaktor mit 19 l Gesamtvolumen ist mit einem Fließinjektionsanalysator (FIA) gekoppelt, mit dem die Konzentration der  $\alpha$ -Amylase on-line bestimmt werden kann. Für die Meßwerterfassung und die Steuerung des FIA-Analysators werden Personalcomputer eingesetzt. Die übergeordnete intelligente Prozeßführung wird auf einem UNIX-Arbeitsrechner realisiert, der mit dem Meßwerterfassungsrechner kommuniziert.

Bild 10: Versuchsaufbau des Beispielprozesses.



### 5.1 Optimierung der $\alpha$ -Amylaseproduktion mit BioX

Die Optimierung mit dem Grundsystem BioX konzentrierte sich zunächst darauf, für den reinen Batch-Betrieb optimale Profile für die Steuerungsgrößen Temperatur und pH zu finden [11]. Dazu wurde der Prozeß mit neuronalen Netzen vom CMAC-Typ phasenbezogen modelliert. Die neuronalen Modelle schätzen den Verlauf der Amylasekonzentration und weiterer wachstumsgekoppelter Parameter in Abhängigkeit von Temperatur und pH. In jedem Optimierungsschritt werden die Stellgrößen so variiert, daß das Modell ein günstiges Wachstum bzw. eine überdurchschnittliche Produktion vorhersagt. Die Gewichtung zwischen Wachstum und Produktion wird in Anlehnung an das Phasenmodell während des Prozesses angepaßt.

Ausgehend von Versuchen mit konstanten Einstellungen für Temperatur und pH-Wert, die dem initialen Training der neuronalen Modelle dienen, konnte in fünf Optimierungsläufen die Ausbeute der Amylasebildung um 100% gegenüber den günstigsten konstanten Bedingungen gesteigert werden. Für die Erkennung der jeweiligen Prozeßphase wurde ein Regelwerk gebildet, das auch bei Ausfall von Sensorkomponenten eine zuverlässige Klassifikation sicherstellt.

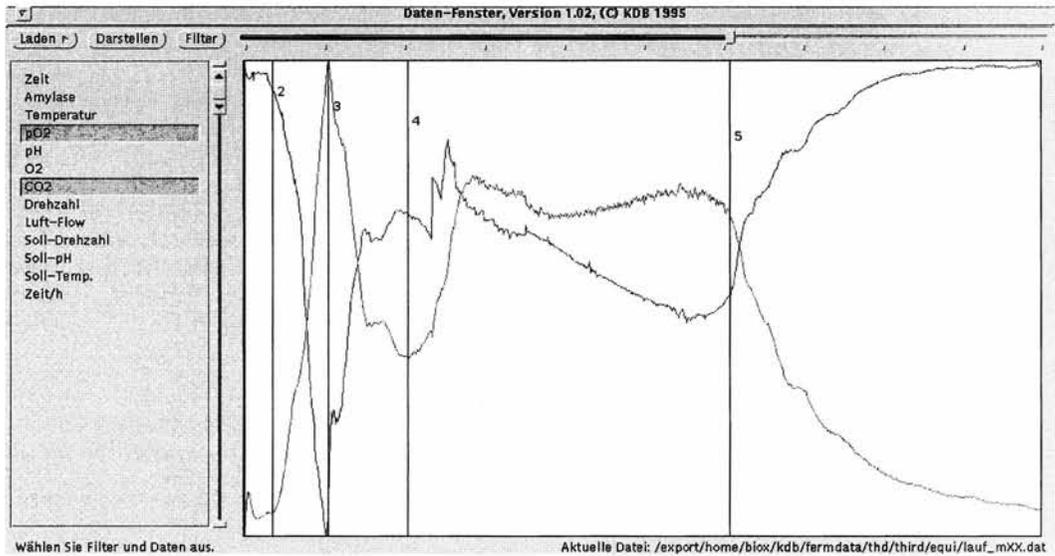
### 5.2 Automatische Regelgenerierung mit BioX<sup>++</sup>

Zur automatischen Generierung von Fuzzy-Regeln wurde zunächst eine Studie anhand des zur Verfügung stehenden Simulationsmodells des oben beschriebenen Beispielprozesses durchgeführt. Ziel war die Überprüfung der prinzipiellen Überlegungen zur Repräsentation und Optimierung der Prozeßführungsstrategien in Form von transparenten Regeln [9].

Wie bei den in Abschnitt 5.1 beschriebenen Experimenten sollten zur Optimierung der Produktausbeute Temperatur und pH in Abhängigkeit vom aktuellen Prozeßzustand dynamisch variiert werden. Zur Realisierung dieser Prozeßführungsstrategie wurde eine entsprechende Regelbasis erstellt und erfolgreich optimiert. Aufgrund des besonderen Trainingsverfahrens wurden sogar etwas bessere Ergebnisse erzielt als bei früheren Untersuchungen mit BioX. Wesentlich ist aber, daß hierbei die optimierten Prozeßführungsstrategien in einer überschaubaren Regelbasis und nicht in einem *black-box*-Speicher repräsentiert werden; somit sind sie nachvollziehbar und können im Hinblick auf Zulassungsverfahren auch offengelegt werden.

Außer zur Repräsentation der Prozeßführungsstrategien kann der Ansatz des vollständig definierten Regelraums auch angewendet werden, um die charakteristischen Prozeßphasen automatisch zu klassifizieren. Mit Hilfe des in Bild 11 gezeigten, graphisch unterstützten Werkzeuges ist es dem Bediener leicht möglich, die Phasenwechsel zu markieren. Am Pilotprozeß wurde so nach wenigen Optimierungsläufen eine stabile Klassifikation der Phasen erreicht.

**Bild 11:** Graphische Vorgabe der zu klassifizierenden Prozeßphasen anhand der gemessenen Verläufe; Adaptionphase (1), Phase verstärkten Wachstums (2), Übergangsphase (3), Phase verringerten Wachstums (4) und Absterbephase (5).



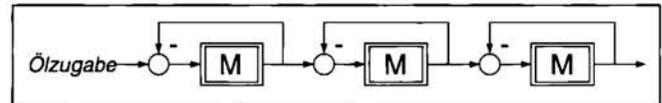
Die Umsetzung vorgegebener Strategien und deren datengetriebene Erweiterung bzw. Korrektur mit Hilfe eines lernfähigen Fuzzy-Ansatzes haben zu einer Vereinfachung der Implementierung sowie einer deutlich verringerten Anzahl notwendiger Experimente geführt, da bekanntes Vorwissen leicht integriert werden konnte [10]. Die bei veränderten Versuchsbedingungen immer wieder vorzunehmende Adaption bestehender Regelbasen zur Phasenklassifikation wird durch die automatische Modifikation anhand graphischer Vorgaben in aufgenommenen Prozeßdaten sowohl vereinfacht als auch bezüglich ihrer Stabilität verbessert.

### 5.3 Datengetriebene Generierung eines Modells zur Biomasseschätzung mit SMOG

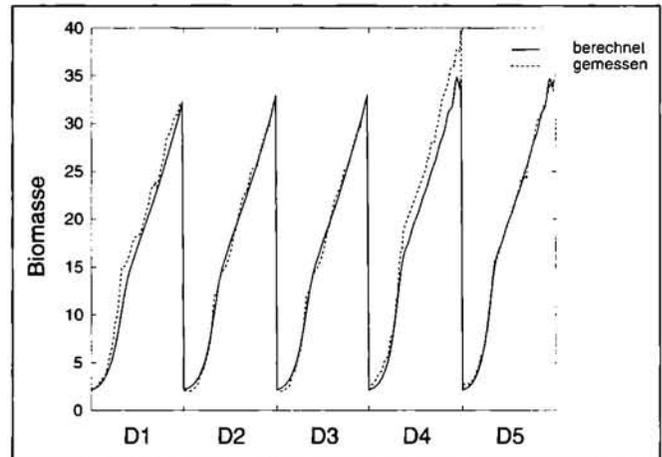
Im Rahmen einer Kooperation mit der BASF AG, Ludwigshafen, wurden auf der Grundlage industrieller Fermentationsdaten Untersuchungen zur datengetriebenen Generierung strukturierter Prozeßmodelle mit SMOG durchgeführt (die nur teilweise zur Veröffentlichung freigegebenen Ergebnisse sind in [16] nachzulesen).

Bild 12 zeigt eine nach mehreren Versuchen erzielte, biologisch plausible Modellstruktur zur Schätzung der Biomasse in Abhängigkeit von der Nährstoffzugabe im Fed-Batch-Betrieb. Sowohl die in diesem Modell verwendete Eingangsgröße als auch die Bausteine wurden vom SMOG aus einer großen Menge zur Verfügung gestellter Eingangsgrößen und Bausteine selbständig auf der Basis zweier Trainingsdatensätze ausgewählt. Der Modellbildungsprozeß wurde durch die Variation der Randbedingungen zielgerichtet beeinflusst. Bild 13 zeigt die Reproduktion der Trainingsdatensätze (*D3* und *D5*) sowie die Generalisierungsfähigkeit anhand von drei weiteren, unter variierten experimentellen Bedingungen gewonnenen Datensätzen (*D1*, *D2* und *D4*).

Zusätzliche, im Rahmen der Kooperation durchgeführte Untersuchungen haben gezeigt, daß dieses Modell im Intervall einer Variation des dynamischen



**Bild 12:** Struktur eines automatisch generierten Modells zur Schätzung der Biomasse.



**Bild 13:** Vergleich der mit Modell aus Bild 12 berechneten und der vorgegebenen Verläufe der Biomasse.

Ölzugabeprofil von  $-20\%$  bis  $+10\%$  noch sinnvolle Vorhersagen über die Entwicklung der Biomasse liefert [10].

## 6 Zusammenfassung und Ausblick

Der vorliegende Artikel beschreibt – ausgehend von den prinzipiellen Anforderungen der biologischen Prozeßführung – grundlegende Ansätze, durchgeführte Untersuchungen sowie mehrere realisierte Systeme, die in der vergangenen Dekade am Fachgebiet Regelsystemtheorie & Robotik der TH Darmstadt entwickelt, bearbeitet und implementiert wurden. Die Leistungsfähigkeit

higkeit dieser selbstorganisierenden, wissensbasierten und lernenden Verfahren konnte – sowohl bei Anwendung einzelner Methoden, als auch im Rahmen integrierender Konzepte – durch Simulationsstudien, vor allem aber experimentell an mehreren Pilotprozessen, z.B. der Produktion von  $\alpha$ -Amylase mit *B. subtilis*, nachgewiesen werden.

Vom theoretischen Standpunkt aus konzentrieren sich die laufenden Arbeiten auf die langfristige Vorhersage von Fermentationsprozessen. Hierzu wird die Methode der Genetischen Programmierung weiterentwickelt, um zu umfassenden strukturierten Modellen und einem tieferen Systemverständnis zu gelangen. Ziel ist es außerdem, entsprechende Methoden zu entwickeln, die den gesamten Prozeßablauf in die Optimierung einbeziehen.

In den letzten drei Jahren wurden – bedingt durch ein verstärktes Interesse der chemischen und pharmazeutischen Industrie – die beschriebenen Methoden auch erfolgreich in die industrielle Praxis umgesetzt. Neben den Ergebnissen, die aus einer Kooperation der TH Darmstadt mit der BASF AG entstanden sind, wurde von der Zentrum für Neuroinformatik GmbH ein Regelungssystem entwickelt, das dem Anwender in Forschung und Produktion erstmals die selbsttätige Entwicklung von Regelungsschemata zur Zustandsklassifikation und Prozeßoptimierung mit Hilfe von Fuzzy-Logik und neuronalen Netzen gestattet.

#### Literatur

- [1] *Konstatinov, K., und Yoshida, T.*: Physiological state control of fermentation processes. *Biotechnology and Bioengineering*, 33:1145–1156, 1989.
- [2] *Koza, J. R.*: Genetic Programming: On the Programming of Computers by Means of Natural Selection. The MIT Press, Cambridge, Massachusetts, 1992.
- [3] *Nowinski, G.*: Lernfähige konnektionistische Strukturen in der Automatisierungstechnik. *at – Automatisierungstechnik*, 42(1-11):A1–A40, 1994.
- [4] *Tolle, H., und Ersü, E.*: Neurocontrol, Band 172 aus Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin, 1992.
- [5] *Bastin, G., und Dochain, D.*: On-line Estimation and Adaptive Control of Bioreactors. Elsevier Science, New York, 1990.
- [6] *Gehlen, S., Tolle, H., Kreuzig, J., und Friedl, P.*: Integration of Expert Systems and Neural Networks for the Control of

Fermentation Processes. In IFAC Symposium on Modelling and Control of Biotechnological Processes, Keystone, Colorado, USA, 1992.

- [7] *Tolle, H.*: Autonomieerhöhung durch Imitation menschlicher Intelligenz. *at – Automatisierungstechnik*, 39(3/4):85-91/109-115, 1991.
- [8] *Brown, M., und Harris, C. J.*: Neurofuzzy Adaptive Modelling and Control. Prentice Hall, 1994.
- [9] *Bettenhausen, K. D., Möller, S., und Tolle, H.*: Autonomous and Transparent Generation of Control Strategies. In: Isidori, A., Bittanti, S., Mosca, E., Luca, A. D., Benedetto, M. D. und Oriolo, G. (Hrsg.), Third European Control Conference ECC '95, Band 2, Seiten 1231–1236, Rome, Italy, September 1995.
- [10] *Bettenhausen, K. D.*: Automatische Struktursuche für Regler und Strecke: Beiträge zur datengetriebenen Analyse und optimierenden Führung komplexer Prozesse mit Hilfe evolutionärer Methoden und lernfähiger Fuzzy-Systeme. Dissertation TH Darmstadt, FG Regelsystemtheorie & Robotik, 1996. VDI Fortschritt-Berichte 8 (574), VDI-Verlag.
- [11] *Gehlen, S.*: Untersuchungen zur wissensbasierten und lernender Prozeßführung in der Biotechnologie. Dissertation, TH Darmstadt, FG Regelsystemtheorie & Robotik, 1993. VDI Fortschritt-Berichte 20 (87), VDI-Verlag.
- [12] *Bettenhausen, K. D.*: Einsatz intelligenter Systeme zu Prozeßführung. In Meß-, Analysen- und Prozeßleittechnik - Anwendung der Fuzzy-Technik. *ACHEMA '94 – Internationales Treffen für Chemische Technik und Biotechnologie*, Jun 1994.
- [13] *Bettenhausen, K. D., Gehlen, S., Marenbach, P., und Tolle, H.*: *BioX<sup>++</sup>* – New Results and Conceptions Concerning the Intelligent Control of Biotechnological Processes. In: Munack A. und Schügerl, K. (Hrsg.), Proc. 6th Int. Conf. on Computer Applications in Biotechnology, Seiten 324–327, Garmisch-Partenkirchen, Germany, Mai 1995.
- [14] *Marenbach, P., Bettenhausen, K. D., und Tolle, H.*: Methods for a Transparent Development and Optimization of Biotechnological Processes. In: Gertler, J., Cruz, J. und Peshkin, M. (Hrsg.) 13th IFAC World Congress, Band N, Seiten 449–454, San Francisco, CA, 1996.
- [15] *Marenbach, P., Bettenhausen, K. D., und Cuno, B.*: Selbstorganisierende Generierung strukturierter Prozeßmodelle. *at – Automatisierungstechnik*, 43(6):277–288, 1995.
- [16] *Bettenhausen, K. D., Marenbach, P., Freyer, S., Rettenmaier H., und Nieken, U.*: Self-organizing Structured Modelling of Biotechnological Fed-batch Fermentation by Means of Genetic Programming. In: Zalzal, A (Hrsg.), Proc. Int. Conf. on Genetic Algorithms in Engineering Systems: Innovations and Applications, Band 414 aus IEE Conference Publication, Seiten 481–486, London, UK, 1995.

Manuskripteingang: 18. September 1996.

Verfügbar unter  
lediglich die vom Gesetz vorgesehenen Nutzungsrechte  
gemäß UrhG