Multiple Model Approaches to Modelling and Control

Edited by R. Murray-Smith
and T. A. Johansen

Taylor & Francis
Taylor & Francis Group

Multiple Model Approaches to Modelling and Control

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Multiple Model Approaches to Modelling and Control

EDITED BY

RODERICK MURRAY-SMITH

AND

TOR ARNE JOHANSEN

Taylor & Francis
Publishers since 1798

Published by Taylor & Francis
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN
270 Madison Ave, New York NY 10016

Transferred to Digital Printing 2010

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

ISBN 0-7484-0595-X

Library of Congress Cataloging in Publication data are available

Cover design by Amanda Barragry
Typeset in Times 10/12pt by Digital by Design

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original may be apparent.

Contents

Series Introduction

Control systems has a long and distinguished tradition stretching back to nineteenth-century dynamics and stability theory. Its establishment as a major engineering discipline in the 1950s arose, essentially, from Second World War driven work on frequency response methods by, amongst others, Nyquist, Bode and Wiener. The intervening 40 years has seen quite unparalleled developments in the underlying theory with applications ranging from the ubiquitous PID controller widely encountered in the process industries through to high-performance/fidelity controllers typical of aerospace applications. This development has been increasingly underpinned by the rapid developments in the, essentially enabling, technology of computing software and hardware.

This view of mathematically model-based systems and control as a mature discipline masks relatively new and rapid developments in the general area of robust control. Here intense research effort is being directed to the development of high-performance controllers which (at least) are robust to specified classes of plant uncertainty. One measure of this effort is the fact that, after a relatively short period of work, 'near world' tests of classes of robust controllers have been undertaken in the aerospace industry. Again, this work is supported by computing hardware and software developments, such as the toolboxes available within numerous commercially marketed controller design/simulation packages.

Recently, there has been increasing interest in the use of so-called intelligent control techniques such as fuzzy logic and neural networks. Basically, these rely on learning (in a prescribed manner) the input–output behaviour of the plant to be controlled. Already, it is clear that there is little to be gained by applying these techniques to cases where mature mathematical model-based approaches yield high-performance control. Instead, their role is (in general terms) almost certainly going to lie in areas where the processes encountered are ill-defined, complex, nonlinear, time-varying and stochastic. A detailed evaluation of their (relative) potential awaits the appearance of a rigorous supporting base (underlying theory and implementation architectures for example) the essential elements of which are beginning to appear in learned journals and conferences.

Elements of control and systems theory/engineering are increasingly finding use outside traditional numerical processing environments. One such general area in which there is increasing interest is intelligent command and control systems which are central, for example, to innovative manufacturing and the management of advanced transportation systems. Another is discrete event systems which mix numeric and logic decision making.

It was in response to these exciting new developments that the present book series of Systems and Control was conceived. It publishes high-quality research texts and reference works in the diverse areas which systems and control now includes. In addition to basic theory, experimental and/or application studies are welcome, as are expository texts where theory, verification and applications come together to provide a unifying coverage of a particular topic or topics.

The book series itself arose out of the seminal text: the 1992 centenary first English translation of Lyapunov's memoir *The General Problem of the Stability of Motion* by A. T. Fuller, and was followed by the 1994 publication of *Advances in Intelligent Control* by C. J. Harris. Since then a number of titles have been published and many more are planned. A full list is given below.

Advances in Intelligent Control, edited by C. J. Harris

Intelligent Control in Biomedicine, edited by D. A. Linkens

Advances in Flight Control, edited by M. B. Tischler

Forthcoming

Sliding Mode Control: Theory and Applications, by C. Edwards and S. K. Spurgeon

Neural Network Control of Robot Manipulators and Nonlinear Systems, by F. L. Lewis, S. Jagannathan and A. Yesildirek

Generalized Predictive Control with Applications to Medicine, by M. Mahfouf and D. A. Linkens

Control of Linear Multivariable Systems with Saturating Actuators, by Z. Lin, A. Saberi and P. Sannuti

A Unified Algebraic Approach to Linear Control Design, by R. E. Skelton, T. Iwasaki and K. Grigoriadis

Sliding Mode Control in Electro-Mechanical Systems, by V. I. Utkin, J. Guldner and J. Shi

From Process Data to PID Controller Design, by L. Wang and W. R. Cluett

E. ROGERS
J. O'REILLY

Foreword

Modelling and control of nonlinear dynamical systems is one of the most challenging areas of system theory. A great deal of recent research activity has focused on approaches such as neural networks and fuzzy logic. Much of this work on identifying nonlinear input–output models was, however, based on architectures and approaches which provided little insight into the underlying data generating process (the so called transparency problem). These methods were also not ideally suited for adaptive control applications where tracking of non-stationary plant behaviour is necessary.

For 'intelligent' modelling and control methods to make a practical contribution to real problems they should fulfil a number of requirements. They should be able to work online, with provable convergence (in learning) and provide a framework for analysing system stability. It is important that these methods are parsimonious, in that they should have good regularisation characteristics which lead to models which have low complexity and are therefore easier to interpret. Moreover, it must be feasible to implement them in real systems, and it should be easy to incorporate both mechanistic and rule-based or symbolic *a priori* knowledge into the learning schema.

In this regard basic research into local approaches to modelling and control, including classical control approaches, statistical methods, neuro- and fuzzy architectures (e.g. B-splines and Radial Basis Function networks) has shown substantial progress in recent years. In this book, these issues are addressed by authors who have made a substantial contribution towards resolving the problems. Many chapters are concerned with the use of local models and controllers which are parameterised by operating point conditions as a means of exploiting classical linear control methods for resolving global nonlinear control and modelling problems. This approach has been previously successfully exploited as an effective means of providing smooth bumpless control. The book provides a very effective and comprehensive coverage of current research in this area and provides in the opening chapter a welcome architectural viewpoint that integrates these apparently disparate approaches. The different methods include probabilistic interpretations, such as Jordan's mixture of models, fuzzy interpretations of local model methods, and Kuipers' and Åström's heterogeneous models.

Specialised topics in identification are included, e.g. the application of construction algorithms to model structure identification is analysed, active learning is used, identification methods are interpreted, and there is a specialist chapter on the normalising effects of basis functions. These effects have been observed before in neurofuzzy systems (the partition of unity requirement), but their significance to local modelling is particularly valuable. A number of chapters are concerned with control, addressing local model-based controllers such as the Takagi–Sugeno fuzzy controllers (equivalent to operating point-dependent local linear controllers), Multiple Model Adaptive Control methods, as well as local approaches in combination with more classical methods such as H_∞, and Laguerre polynomials. The stability of these methods is also investigated. Several chapters include applications which provide the reader with considerable insight into the practical implementability of these ideas.

The most valuable contribution of the book is that it brings together basically similar approaches which have appeared in different fields. Local techniques for intelligent modelling and control are shown to be a widely used approach with many practical advantages for use in analysis and design of complex systems. This provides an important contribution for researchers involved in nonlinear modelling and control, as the methods are easy to implement, have a number of useful properties, are relatively transparent, and have the ability to incorporate *a priori* knowledge with measured data from a real process.

C. J. HARRIS
Southampton University

Preface

WHY MULTIPLE MODELS?

This book presents a number of approaches which produce complex models or controllers by piecing together a number of simpler subsystems. This divide-and-conquer strategy is a long-standing and general way to cope with complexity in engineering systems, nature and human problem solving.

More complex plants, advances in information technology, and tightened economical and environmental constraints in recent years have lead to practising engineers being faced with modelling and control problems of increasing complexity. When confronted with such problems, there is a strong intuitive appeal in building systems which operate robustly over a wide range of operating conditions by decomposing them into a number of simpler linear modelling or control problems, even for nonlinear modelling or control problems. This appeal has been a factor in the development of increasingly popular 'local' and multiple-model approaches to coping with strongly nonlinear and time-varying systems.

Such local approaches are directly based on the divide-and-conquer strategy, in the sense that the core of the representation of the model or controller is a partitioning of the system's full range of operation into multiple smaller operating regimes each of which is associated with a locally valid model or controller. This can often give a simplified and transparent nonlinear model or control representation. In addition, the local approach has computational advantages, it lends itself to adaptation and learning algorithms, and allows direct incorporation of high-level and qualitative plant knowledge into the model. These advantages have proven to be very appealing for industrial applications, and the practical, intuitively appealing nature of the framework is demonstrated in chapters describing applications of local methods to problems in the process industries, biomedical applications and autonomous systems. The successful application of the ideas to demanding problems is already encouraging, but creative development of the basic framework is needed to better allow the integration of human knowledge with automated learning.

The underlying question is 'How should we partition the system – what is 'local'?'. This book presents alternative ways of bringing submodels together, which lead to varying levels of performance and insight. Some are further developed for autonomous learning of parameters from data, while others have focused on the ease with which prior knowledge can be incorporated. It is interesting to note that researchers in Control Theory, Neural

Networks, Statistics, Artificial Intelligence and Fuzzy Logic have more or less independently developed very similar modelling methods, calling them *Local Model Networks*, *Operating Regime based Models*, *Multiple Model Estimation and Adaptive Control*, *Gain Scheduled Controllers*, *Heterogeneous Control*, *Mixtures of Experts*, *Piecewise Models*, *Local Regression* techniques, or *Tagaki–Sugeno Fuzzy Models*, among other names. Each of these approaches has different merits, varying in the ease of introduction of existing knowledge, as well as the ease of model interpretation. This book attempts to outline much of the common ground between the various approaches, encouraging the transfer of ideas.

Recent progress in algorithms and analysis is presented, with constructive algorithms for automated model development and control design, as well as techniques for stability analysis, model interpretation and model validation.

OVERVIEW OF THE BOOK

Part I – Basic principles

The editors introduce the basic ideas of multiple model approaches in Chapter 1, where the existing paradigms for the application of multiple model and operating regime approaches to nonlinear modelling, identification and control are explored. The chapter also provides a survey and overview of the state of the art in terms of procedures, algorithms and tools for visualisation and interpretation.

Part II – Modelling

Part II of the book deals predominantly with modelling methods and applications, including methods for estimation and experiment design.

In Chapter 2, Babuška and Verbruggen describe the Takagi–Sugeno fuzzy model, used as an interpolating scheduler for a set of multiple linear models which are valid locally around certain operating conditions. The antecedent of the fuzzy rule provides the local region and the interpolating mechanism, while the consequent is the locally valid model. The difficult task of learning the antecedents and consequents from data is reviewed and a constructive approach incorporating fuzzy clustering is developed. The identification methods are demonstrated using experimental data from problems in biotechnology and medicine.

Gollee, Hunt, Donaldson and Jarvis show in Chapter 3 that nonlinear models based on multiple local ARX models are able to capture the nonlinear effects apparent in experiments with electrically stimulated muscles, and provide high accuracy over a wide range of input signals. This chapter is an interesting example of local methods being applied to a problem which has been studied intensively with a variety of complex mathematical modelling techniques, and comparing well. The identification methods used are those from Chapter 7.

The 'functional state' approach described by Halme, Visala and Zhang in Chapter 4 uses a discrete representation to describe the current 'functional state' of the system. This concept is close to the idea of an operating regime, and each functional state has an associated local model. A finite state automaton is used to describe the possible transitions between operating regions of dynamical processes, and multi-layer perceptron neural networks with Laguerre filters are trained to recognise transitions between states. The concept is illustrated by experiments with a two-tank problem, and a fermentation process.

In Chapter 5 Meilă and Jordan review the *Mixture of Experts* model structure and extend it to a *Markov Mixture of Experts*, where a Markov graph is used to define the transitions between multiple models in the system. This is similar in many ways to the functional state approach described in Chapter 4, but this time placed in a probabilistic framework where the transitions are described by a Markov model. The method is applied to fine motion control in robotics.

In Chapter 6 by Cohn, Ghahramani and Jordan, experiment design with the *mixture of Gaussians* model representation is studied. Local representations make it easier to produce local confidence limits, which can be used as the basis for an active learning algorithm, where the optimal search for new data can be guided by the model structure. Robotics simulations are used to illustrate the ideas. This chapter also, as with Chapter 5, gives useful insight into the probabilistic interpretation of multiple model approaches. The Expectation Maximisation algorithm is used for learning – this is also studied in Chapter 7. A further interesting aspect of the representation used in this chapter is that multiple local models are used to represent the joint input–output *density* of the data, which does not distinguish between inputs and outputs, unlike the other approaches which explicitly represent input–output or input–state mappings.

In Chapter 7, Murray-Smith and Johansen show that the commonly used global least squares method for parameter identification in multiple local models can be very sensitive and lead to ill-conditioning. They propose a cheaper locally weighted least squares identification method as a solution. The interactions between model structure and parameter identification methods are discussed – this theme reappears in several other chapters. The smoothing analysis used to illustrate the effects discussed, is also a general technique, which can be applied to other frameworks.

Chapter 8, by Shorten and Murray-Smith examines some of the side-effects of basis function normalisation – a common technique used in the weighting functions in local model and control structures to ensure that the operating range is completely covered. Normalisation also appears naturally in fuzzy and probabilistic representation of the weighting functions. Normalisation has a number of side-effects which alter the global properties of the model or controller, with respect to robustness and interpretability. These become especially important when automatic learning algorithms are used to adapt the basis functions. As well as the graphical and intuitive explanations of the side-effects, the chapter also describes some mathematical tools which can help gain a deeper understanding of the trade-offs involved.

Part III – Control

Part III of the book is dedicated to applications of multiple model methods for nonlinear control.

In Chapter 9, Kuipers and Åström present methods for developing a nonlinear controller by combining multiple heterogeneous local control laws appropriate to different operating regions. Operating regions are described using fuzzy set membership, as in Chapter 2, but the local controllers can be classical control laws with their own internal states. Qualitative simulation is suggested as a method for validation of the global behaviour of the heterogeneous controls. Some aspects of the control law can, even in the case of incomplete knowledge, be represented as a qualitative differential equation, and qualitative simulation can be used to predict the possible behaviours of the system. The methods are demonstrated on a water level controller and a highly nonlinear chemical reactor.

In Chapter 10, Sbarbaro uses operating regime based models with multiple *local Laguerre models* for identification and control. Local Laguerre models potentially have advantages over the more common ARX local model as they can cope more easily with uncertainty in time delays and model orders. The method is compared in a simulation of a chemical reactor using a model predictive control algorithm.

Chapter 11 by Schott and Bequette describes multiple model adaptive control (MMAC), which is a classical model-based control strategy. Multiple models are used and a probabilistic weighting chooses which model or combination of models best represents the current plant input/output behaviour. The authors review MMAC theory, including model bank estimation and control, and describe applications to biomedical control problems.

In the work presented by Banerjee, Arkun, Pearson and Ogunnaike in Chapter 12, the composition of multiple linear state-space models is described as a parameter-varying model. The parameters of the global model are the local model weights which are estimated on-line using a Bayesian approach similar to Chapter 11. A globally stable controller composed of multiple local linear controllers is then designed for the linear parameter varying model using H_∞ design based on Linear Matrix Inequalities. The theory is applied to a simulated chemical reactor.

Zhao, Gorez and Wertz present in Chapter 13 a method for identification and structured analysis and design of Takagi–Sugeno fuzzy models and controllers. The Takagi–Sugeno fuzzy model is based on multiple local linear state-space models that are weighted using fuzzy membership functions. An identification method based on fuzzy clustering (see also Chapter 2) is presented and experimental results from application on a glass furnace are included. The control design method guarantees stability and robustness properties. The methods are based on modern tools such as Linear Matrix Inequalities, being closely related to Chapter 12. Simulation examples are used to illustrate the methods.

We hope that this book will bring to a wider audience the progress being made in both practical and theoretical use of the multiple models philosophy, and that the workers in the field will be able to gain a deeper understanding of the relations between the different existing approaches, and tools.

RODERICK MURRAY-SMITH
Berlin

TOR ARNE JOHANSEN
Trondheim

Contributors

Karl Johan Åström
Lund Institute of Technology, Department of Automatic Control, Box 118, S-221 00 Lund, Sweden
E-mail: kja@control.lth.se
WWW address: http://control.lth.se:80/~kja/

Yaman Arkun
Georgia Institute of Technology, School of Chemical Engineering, Atlanta, GA 30332-0100, USA
E-mail: yaman.arkun@che.gatech.edu
WWW address: http://www.chemse.gatech.edu/people/ya.html

Robert Babuška
Delft University of Technology, Control Laboratory, Department of Electrical Engineering, P.O. Box 5031, 2600 GA Delft, The Netherlands
E-mail: R.Babuska@et.tudelft.nl
WWW address: http://dutera.et.tudelft.nl/people/babuska/babuska.html

Atanu Banerjee
Georgia Institute of Technology, School of Chemical Engineering, Atlanta, GA 30332-0100, USA
E-mail: atanu@cezanne.chemse.gatech.edu
WWW address: http://www.chemse.gatech.edu/grads/ab/atanu.html

B. Wayne Bequette
Rensselaer Polytechnic Institute, Howard P. Isermann Department of Chemical Engineering, Troy, NY 12180-3590, USA
E-mail: bequeb@rpi.edu
WWW address: http://www.rpi.edu/~bequeb

David Cohn
Harlequin Inc., 1010 El Camino Real, Menlo Park, CA 94025, USA
E-mail: cohn@harlequin.com
WWW address: http://www.ai.mit.edu/people/cohn/cohn.html

Nick Donaldson
University College London, Department of Medical Physics and Bioengineering, London, England
E-mail: nickd@medphys.ucl.ac.uk

Zoubin Ghahramani
University of Toronto, Dept. of Computer Science, 6 Kings College Road, Toronto, Ontario M5S 1A4, Canada
E-mail: zoubin@cs.toronto.edu
WWW address: http://www.cs.utoronto.ca/~zoubin/

Henrik Gollee
University of Glasgow, Centre for Systems and Control, Glasgow, Scotland
E-mail: henrik@eng.gla.ac.uk
WWW address: http://www.mech.gla.ac.uk/~henrik/

R. Gorez
Université Catholique de Louvain Centre for Systems Engineering and Applied Mechanics (CESAME), Bât. EULER, Avenue Georges Lemaître, 4, B-1348 Louvain-la-Neuve, Belgium
E-mail: gorez@auto.ucl.ac.be

Aarne Halme
Helsinki University of Technology, Laboratory of Automation Technology, SF-02150, Finland
E-mail: aarne.halme@hut.fi
WWW address: http://www.automation.hut.fi/

Ken Hunt
Daimler-Benz Research, Berlin, Alt-Moabit 96a, D-10559 Berlin, Germany
E-mail: hunt@dbresearch-berlin.de

J. Jarvis
University of Liverpool, Department of Human Anatomy and Cell Biology, Liverpool, England
E-mail: jcj@liverpool.ac.uk

Tor Arne Johansen
SINTEF Electronics and Cybernetics, N-7034 Trondheim, Norway
E-mail: Tor.Arne.Johansen@ecy.sintef.no
WWW address: http://www.itk.ntnu.no/SINTEF/ansatte/Johansen_Tor.Arne/

Michael Jordan
Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
E-mail: jordan@psyche.mit.edu
WWW address: http://www.ai.mit.edu/projects/cbcl/web-pis/jordan/homepage.html

Benjamin Kuipers
University of Texas at Austin, Department of Computer Sciences, Austin TX 78712, USA
E-mail: kuipers@cs.utexas.edu
WWW address: http://www.cs.utexas.edu/users/kuipers/

Marina Meilă
Dept. of Elec. Eng. and Computer Sci., Massachusetts Institute of Technology, Cambridge, MA 02139, USA,
and *Department of Automatic Control and Computers, Politehnica University of Bucharest, Romania.*
E-mail: mmp@psyche.mit.edu
WWW address: http://www.ai.mit.edu/people/mmp/mmp.html

Roderick Murray-Smith
Daimler-Benz Research, Berlin, Alt-Moabit 96a, D-10559 Berlin, Germany, from 1997, Department of Mathematical Modelling, Technical University of Denmark, Building 321, DK-2800 Lyngby, Denmark
E-mail: murray@dbresearch-berlin.de

B. Ogunnaike
E. I. DuPont deNemours & Co., Inc. Wilmington, DE 19880-0101, USA
E-mail: ogunnaike@esspth.dnet.dupont.com

R. Pearson
E. I. DuPont deNemours & Co., Inc. Wilmington, DE 19880-0101, USA
E-mail: pearsork@esspth.dnet.dupont.com

Daniel Sbarbaro-Hofer
Universidad de Concepcion, Dept. of Electrical Engineering, Casilla 53-C, Concepcion, Chile
E-mail: dsbarbar@vangogh.die.udec.cl
WWW address: http://www.die.udec.cl/~dsbarbar/

Kevin Schott
Rensselaer Polytechnic Institute, Howard P. Isermann Department of Chemical Engineering, Troy, NY 12180-3590, USA
E-mail: schotk@rpi.edu

Robert Shorten
Daimler-Benz Research, Berlin, Alt-Moabit 96a, D-10559 Berlin, Germany
E-mail: shorten@dbresearch-berlin.de

H. B. Verbruggen
Delft University of Technology, Control Laboratory, Department of Electrical Engineering, P.O. Box 5031, 2600 GA Delft, The Netherlands
E-mail: H.B.Verbruggen@et.tudelft.nl
WWW address: http://lcewww.et.tudelft.nl/~vbruggen

Arto Visala
Helsinki University of Technology, Laboratory of Automation Technology, SF-02150, Finland
E-mail: arto.visala@hut.fi
WWW address: http://www.automation.hut.fi/

V. Wertz
Université Catholique de Louvain Centre for Systems Engineering and Applied Mechanics (CESAME), Bât. EULER, Avenue Georges Lemaître, 4, B-1348 Louvain-la-Neuve, Belgium
E-mail: wertz@auto.ucl.ac.be

Xia-Chang Zhang
Helsinki University of Technology, Laboratory of Automation Technology, SF-02150, Finland
E-mail: xiachang.zhang@hut.fi
WWW address: http://www.automation.hut.fi/

J. Zhao
Université Catholique de Louvain Centre for Systems Engineering and Applied Mechanics (CESAME), Bât. EULER, Avenue Georges Lemaître, 4, B-1348 Louvain-la-Neuve, Belgium
E-mail: zhao@auto.ucl.ac.be

PART ONE

Basic Principles

CHAPTER ONE

The Operating Regime Approach to Nonlinear Modelling and Control

TOR ARNE JOHANSEN and RODERICK MURRAY-SMITH

Multiple model approaches have appeared more or less independently in several branches of science and engineering, disguised in different terminology. The purpose of this chapter is to introduce the basic ideas of multiple model approaches, focusing on their similarities rather than their differences. Existing paradigms for the application of multiple model and operating regime approaches to nonlinear modelling, identification and control are explored. A survey and overview of the state of the art in terms of procedures, algorithms and tools for visualisation and interpretation are also provided.

1.1 INTRODUCTION

Technological development is steadily increasing the complexity of process plants, vehicles and other engineered systems, and economical and environmental constraints are raising the awareness of the need for practical approaches which can be used to aid engineers to better understand and perform such complex modelling and control tasks.

These highly complex systems are characterised by a large number of components which are strongly coupled and have a wide operating range, among other factors. However, despite the increasing mathematical sophistication of research in modelling and control over the last decades, and the greater use of powerful computing facilities, many of the advanced methods have not been regularly applied to real problems under normal working conditions. This is often because of the theoretical sophistication required to understand the methods. Hence, if we want to deal with complex high-dimensional, coupled, nonlinear and non-stationary systems, tomorrow's automatic control systems must be more autonomous, robust, intelligent and user-friendly. However, dealing with complexity is obviously an inherently difficult problem, by the very definition of the word. The principle of incompatibility (Zadeh 1973) tries to make the implications of complexity more precise: *As the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance (or relevance) become almost mutually exclusive characteristics.* A consequence of this principle is obviously that models and analysis of complex systems will be less precise than for simple systems. A further consequence is that one should perhaps look for other model representations and tools that can make use of less precise system knowledge than the traditional approaches which worked well for low and medium complexity modelling and control problems in the past. This is indeed the trend in the area of intelligent control where fuzzy logic, qualitative modelling, neural networks, expert systems and probabilistic reasoning are being explored, e.g. (Åström and McAvoy 1992, Antsaklis *et al.* 1991, Åström *et al.* 1986).

In everyday life, as well as in solving engineering problems, the standard approach to complex problem solving is the divide-and-conquer strategy: *A complex problem is somehow partitioned into a number of simpler subproblems that can be solved independently, and whose individual solutions yield the solution of the original complex problem.* The key to successful problem solving with this approach is to find suitable axes along which the problem can be partitioned. In this work we will focus on one approach to the decomposition of modelling and control problems that has recently attracted significant attention, namely operating regime decomposition.

The core of the operating regime approach is to make use of a partitioning of the operating range of the system in order to solve modelling and control problems. The operating regime approach thus leads to multiple model or multiple controller approaches, where different local models or controllers are applied under different operating conditions, see Figure 1.1. The supervisor (or scheduler) will coordinate the local models or controllers. This coordination may include selection of a single one, or combining the actions or parameters of a number of local models or controllers.

The operating regimes can often be characterised by different sets of phenomena or behaviours of the system. The rationale behind this approach is basically that the development of local models (or controllers) is simpler because the interactions between the relevant phenomena in each operating regime are simpler locally than globally. For instance, if the system phenomena or behaviour change smoothly with the operating point, then a linear model (or controller) will always be sufficient locally by making each operating regime sufficiently small, even though the system may contain complex nonlinearities when viewed globally. Other motivations for operating regime approaches include:

- The model/controller structure is easy to understand and interpret, both qualitatively and quantitatively, and the approach has its roots in traditional engineering methods.
- Various types of knowledge can be incorporated and integrated within the framework, including qualitative knowledge, empiricism, measured data and available models. Operating regime based modelling is closely related to grey-box modelling.
- Reduced computational complexity compared to other nonlinear methods.

Chapter overview

In the remainder, we will discuss this approach in some detail. In section 1.2 the fundamentals are introduced and numerous simple examples are given. Next, in section 1.3, a broader perspective is taken, and the operating regime approach is seen in relation to other approaches. Details on the available algorithms, procedures and tools for operating regime based modelling and control can be found in sections 1.4 and 1.5, respectively. Further details on analysis, validation and interpretation can be found in section 1.6. Finally, some concluding remarks are given in section 1.7.

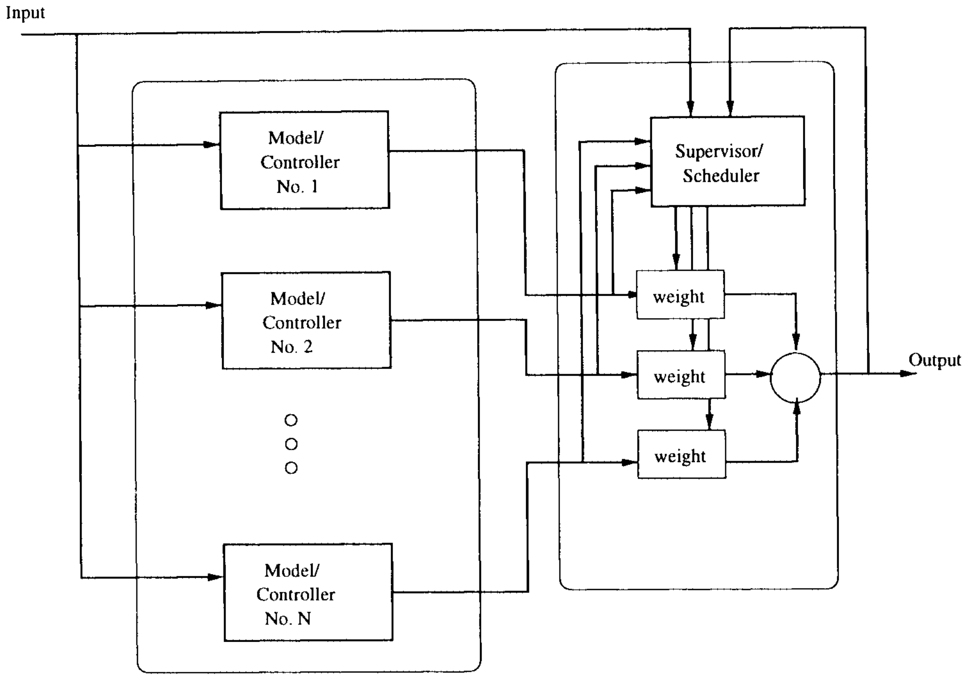

Figure 1.1 The multiple model/controller approach. This is a general visualisation of the methods described in this book, where multiple sub-models or -controllers are organised in some way by a scheduling or gating block. Each method varies in the style of scheduling done, the variables used for the scheduling, whether parameters or states are scheduled and in the structures of the individual local models or controllers.

1.2 OPERATING REGIME APPROACHES

1.2.1 The basics of the framework

Any model or controller will have a limited range of operating conditions in which it is sufficiently accurate or performs sufficiently well in order to serve its purpose. This range may be restricted by several factors, such as validity of linearisation, modelling assumptions, stability properties, or experimental conditions. A model or controller that is useful in a region less than the full range of operating conditions is called a local model or controller, as opposed to a global model or controller which is useful over the full range of operating conditions. Of course, the ultimate goal is a global model or controller. However, we will argue that in some cases it may be beneficial to achieve this goal by developing a number of local models or controllers.

The basis of the framework is a decomposition of the system's full range of operation into a number of possibly overlapping operating regimes, as illustrated in Figure 1.2. In each operating regime, a simple local model or controller is applied. These local models or controllers are then combined in some way to yield a global model or controller. Hence, model or controller development within this framework typically consists of the following tasks:

- Decompose the system's full range of operation into operating regimes. This task includes a definition of the full operating range, as well as the identification of variables that can be used to characterise the operating regimes.
- Select simple local model or controller structures within each operating regime. These structures will often be determined by the relevant system knowledge that is available under different operating conditions, as well as the intended purpose of the model or controller.
- The local model or controller structures are usually parameterised by certain variables that must be determined.
- A method for combining the local models or controllers into a global one must be applied. Numerous approaches exist, and can be characterised according to deterministic vs. stochastic assumptions, soft or hard partitions etc.

For practical problems it will not always be easy to find a natural sequence in which these tasks should be approached. Several iterations of the same tasks are usually needed before a satisfactory model or controller is found.

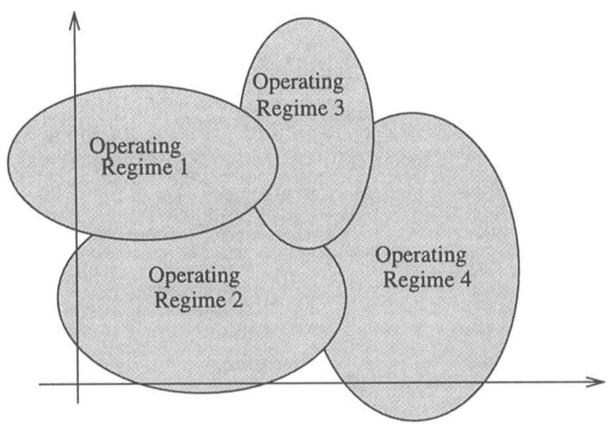

Figure 1.2 The operating range of a complex system is decomposed into a number of operating regimes. These can then be used to represent the system using associated simple subsystems. Much of this book is concerned with how to determine these regions, how to bring subsystems together, and how to interpret the resulting systems.

1.2.2 Some introductory examples

Before we proceed, this section will let us give some simple examples of how operating regime based approaches can be used to approach some nonlinear modelling and control design problems. The more fundamental aspects of partitioning approaches will then be discussed in section 1.2.3.

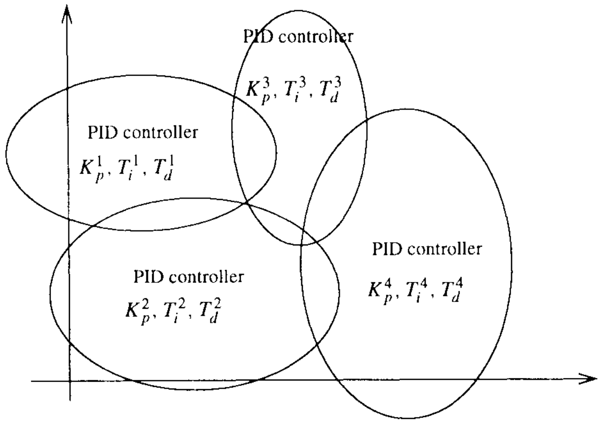

Figure 1.3 A nonlinear PID-like controller is designed by patching together local linear PID controllers.

A nonlinear PID-like controller

Suppose we want to design a nonlinear PID-like controller of the form

$$u(t) = K_p(z(t))e(t) + \frac{1}{T_i(z(t))} \int_0^t e(t)dt + T_d(z(t))\frac{d}{dt}e(t),$$

where the gain K_p, integral time T_i and derivative time T_d are functions of the system's operating point $z(t)$. In standard PID controllers these functions are all constant, and there exist numerous design procedures which guarantee that various stability, performance and robustness specifications are met, e.g. (Åström and Hägglund 1988). Of course, with the nonlinear PID-like controller above, the design problem is much more difficult. However, the operating regime based approach offers an engineering-friendly solution to this design problem. One can for instance design a number of standard linear PID controllers to meet the desired stability, performance and robustness criteria locally when the system is operating in neighbourhoods of some selected operating points $z_1, z_2, ..., z_{n_\mathcal{M}}$:

$$u(t) = K_p^i e(t) + \frac{1}{T_i^i} \int_0^t e(t)dt + T_d^i \frac{d}{dt}e(t), \qquad \text{when } z(t) \text{ is close to } z_i,$$

for $i = 1, 2, ..., n_\mathcal{M}$, see Figure 1.3. Designing such local PID-controllers is often simpler than approaching the nonlinear PID-control design problem directly, even if a nonlinear dynamic model of the system exists. In addition, we need an algorithm for combining or switching between the local PID-controllers. For instance, a weighting of the local PID-parameters as a function of the operating point $\rho_i(z(t))$:

$$u(t) = \sum_{i=1}^{n_\mathcal{M}} \left(K_p^i e(t) + \frac{1}{T_i^i} \int_0^t e(t)dt + T_d^i \frac{d}{dt}e(t) \right) \rho_i(z(t)),$$

i.e.

$$K_p(z(t)) = \sum_{i=1}^{n_\mathcal{M}} K_p^i \rho_i(z(t)),$$

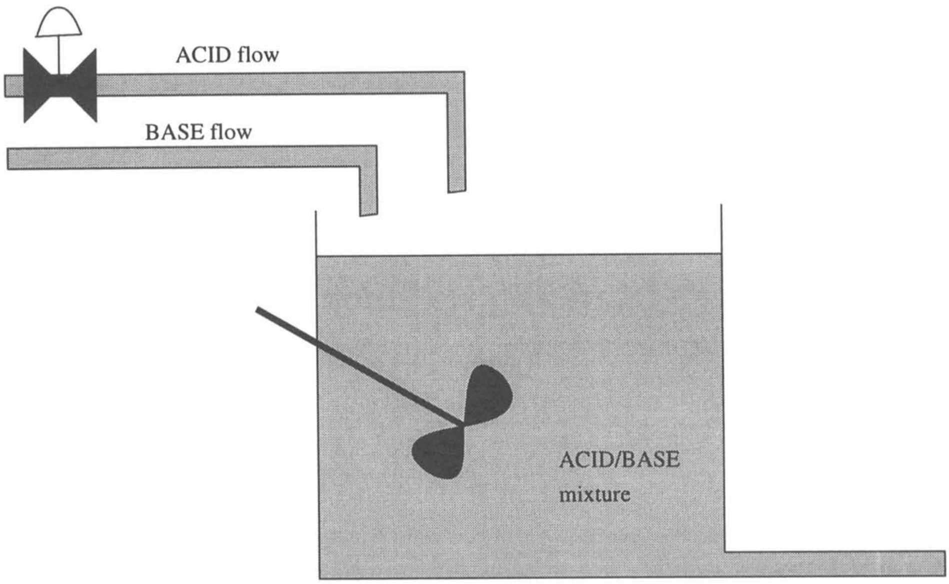

Figure 1.4 A pH neutralisation tank.

$$\frac{1}{T_i(z(t))} = \sum_{i=1}^{n_{\mathcal{M}}} \frac{1}{T_i^i} \rho_i(z(t)),$$

$$T_d(z(t)) = \sum_{i=1}^{n_{\mathcal{M}}} T_d^i \rho_i(z(t)).$$

It should be noted that even though all the locally designed PID controllers satisfy the design criteria locally, extensive analysis or simulation must usually be applied to validate and verify the global properties of the control system.

A nonlinear input–output model of a neutralisation tank

Consider, as a second example, a pH neutralisation tank where acid and base flows into a stirred tank through pipes with valves, see Figure 1.4. The mixture exits through another pipe. Suppose we select a nominal tank level, pH and flow-rates corresponding to an equilibrium point for the tank. Exciting the system about this equilibrium point and collecting experimental data, we can identify a linear transfer function model

$$\text{pH}(s) = \frac{Ke^{-\tau s}}{1 + Ts} q(s),$$

where $q(t)$ is the flow-rate through the valve, pH(t) is the pH value in the tank, and s is the complex variable in the Laplace transform. Suppose we repeat this experiment for a number of different nominal tank levels, pH and flow-rates corresponding to different equilibrium points. Then we would get different identified values of K, τ and T, because the gain would depend on the pH value, and the time-delay would depend on the flow-rate, and the time-constant would depend on both the flow-rate and the tank level. Hence, we have different linear transfer function models which are reasonable descriptions of the system only within some small operating regimes. Again, these local models can be combined using some weighting function into a globally accurate model.

Modelling the longitudinal dynamics of a car

The speed (v) of a car can be described by a dynamic model that takes into account variables such as gear position (g) and throttle angle (α):

$$v(t) = f(v(t-1), \alpha(t-1), g(t-1)).$$

The dynamics of a linearised model

$$\Delta v(t) = a_1 \Delta v(t-1) + b_1 \Delta \alpha(t-1) \tag{1.1}$$

will obviously depend on the gear position, but also on the throttle angle (due to e.g. the nonlinear engine characteristics) and the speed (due to e.g. rolling and drag forces). A model could be based on the operating regime decomposition in Figure 1.5 and local linear models of the form (1.1). A complete example can be found in (Hunt *et al.* 1996*a*).

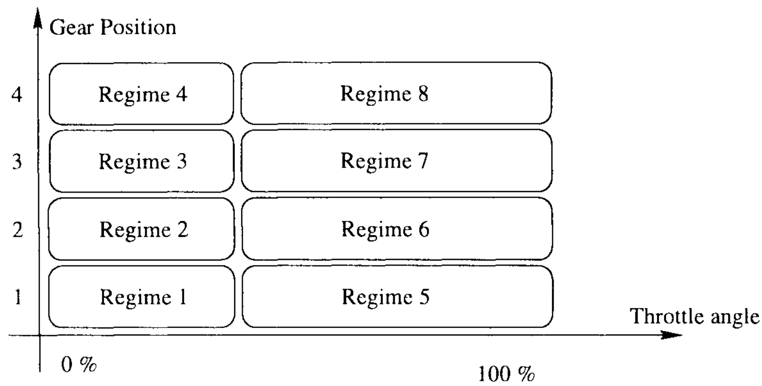

Figure 1.5 Operating regimes for a vehicle.

A nonlinear semi-mechanistic model

Consider a chemical reactor where a set of chemical reactions take place. A mechanistic model typically consists of mass and energy balances. The structure of the balance equations can typically be developed quite easily. The main problem of modelling such systems is to specify how the reaction-rates, mass-transfer coefficients, enthalpies and other model variables depend on the system's state. This typically requires a good understanding of the reaction kinetics, thermodynamics, fluid dynamics, mass- and heat-transfer phenomena present in the reactor. If such knowledge is unavailable or only qualitative knowledge about these phenomena is present, the use of operating regime based approaches can be used to formulate a semi-mechanistic model of the reactor. To illustrate this, let us consider a specific reactor, namely a semi-batch fermentation reactor where the fermentation of glucose to gluconic acid takes place, Figure 1.6.

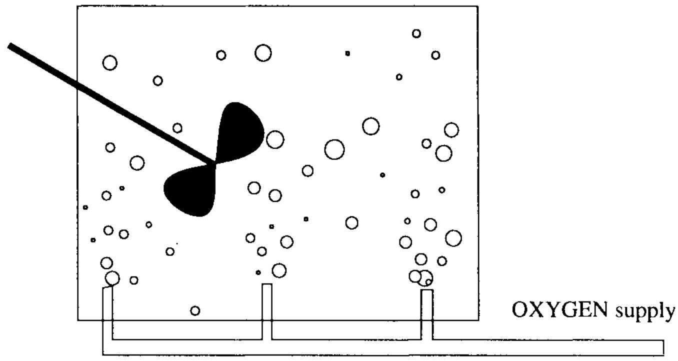

Figure 1.6 A fermentation reactor.

The main overall reaction mechanism can be described by

$$\begin{aligned} \text{Cells} + \text{Glucose} + O_2 &\rightarrow \text{More Cells} \\ \text{Glucose} + O_2 &\overset{\text{Cells}}{\rightarrow} \text{Gluconolactone} \\ \text{Gluconolactone} + H_2O &\rightarrow \text{Gluconic Acid} \end{aligned}$$

The first reaction is the reproduction of cells, consuming the substrate glucose and oxygen. The second reaction is the production of gluconolactone, again consuming glucose and oxygen. This reaction is enzyme-catalyzed by the cells, while the final product, gluconic acid, is formed by the last reaction.

Let x be the state vector that contains the compositions of cells, glucose, gluconolactone, gluconic acid, and oxygen. A mass balance can be written in the form

$$\dot{x} = r(x, T, pH) + q(x), \tag{1.2}$$

where $r(x, T, pH)$ is a vector of reaction rates, $q(x)$ is a vector of flow-rates that describes the uptake of oxygen, and T is temperature. An examination of the reaction mechanism (Johansen and Foss 1993*b*) reveals that the operation of the reactor can naturally be decomposed into four operating regimes:

1 **Initial regime** when it is the number of cells that limits the chemical reactions. Only the first reaction will be significant.
2 **Growth regime** when all reactions proceed at a high rate, limited only by the oxygen supply. All reactions will be significant.
3 **Growth termination regime**, characterised by shortage of glucose which limits the reaction rates.
4 **Termination regime** when all glucose-consuming reactions have terminated, and it is only the production of gluconic acid that is significant.

These operating regimes can for instance be characterised in terms of glucose and oxygen concentrations, see Figure 1.7. Within each operating regime the dynamics of the reactor can be described by linear differential equations with high accuracy, if the dependence on temperature and pH are neglected. Otherwise, one must decompose the regimes further on the basis of these variables, or apply nonlinear local models which take these dependences into account.

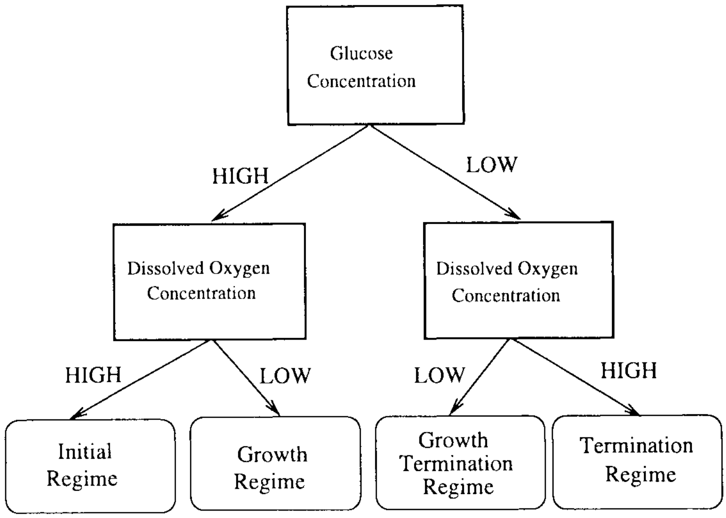

Figure 1.7 Operating regimes for batch fermentation reactor.

Similarly, the flow-rate term q in (1.2) can also be decomposed. The uptake of oxygen will depend on the oxygen concentration, stirring and other fluid dynamics phenomena that in turn may depend on the concentration of cells, since it affects the viscosity and other key properties.

Process operator heuristics

Consider a polymerisation reactor where the operator determines the feed (monomer) flow-rate (Sugeno and Yasukawa 1993). The operator's control procedures can be represented as a set of heuristic rules such as

```
IF (monomer concentration is high)
  AND (monomer concentration is increasing)
  AND (monomer feed-rate is low)
    THEN monomer flow-rate set-point is small.
```

The premise of this rule clearly determines a set of operating conditions that can be viewed as an operating regime. In (Sugeno and Yasukawa 1993), six control rules are identified on the basis of the observed behaviour of the operator.

1.2.3 Fundamental properties

The local versus global dilemma

No matter what underlying model or controller representation one chooses, e.g. state-space or input–output, lumped or distributed, discrete-time or continuous time, the nonlinear model or controller development will typically involve the specification of one or more nonlinear functions. Examples of such functions can be the dependence of the gain on certain process variables, or relations that can be linked to physical phenomena like chemical reactions. The core of the operating regime based approach is to decompose the domain

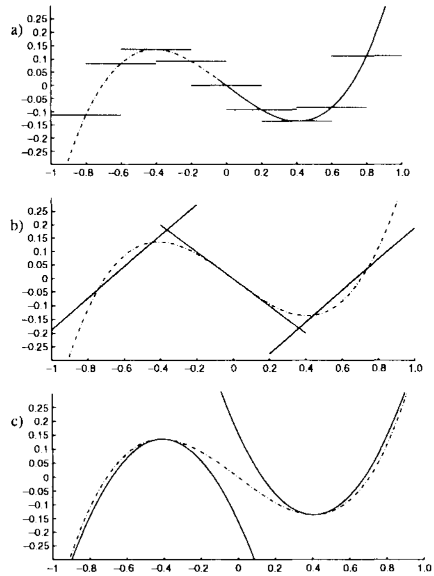

Figure 1.8 Function approximation using local models. The dashed-dotted curve is the function to be approximated, while the solid lines are local approximations. Part a) applies local constant functions, part b) applies local linear functions, while part c) applies local quadratic functions.

of functions into operating regimes and the design of simpler functions that are adequate approximations to the relationships we want to model within their respective operating regimes, see Figure 1.8. Obviously, there is a trade-off between the number and size of the operating regimes on the one hand, and the complexities of the local models on the other. For instance, at one extreme one can have only one large operating regime that covers the full range of operation. The corresponding local model must typically be complex, since it is actually the global model. In general, a decomposition into a few 'large' operating regimes will require more complex local models than a decomposition into numerous 'small' operating regimes, see Figure 1.8b) and c). On the other extreme, one can have a very fine partitioning into operating regimes based on a large number of characterising variables. In this case, the functions we are representing can be approximated by constant values locally, see Figure 1.8a).

Provided the functional relationship we want to approximate is smooth, it should be intuitively clear that approximations based on operating regimes and local models can be made arbitrarily accurate either by making the decomposition into operating regimes sufficiently fine, or by making the local models sufficiently complex. This is studied in detail by (Johansen and Foss 1993*a*), see also (Johansen 1994*c*) for some slight improvement. Explicit bounds on the approximation error as a function of the regularity (smoothness) of the underlying function, the granularity of the decomposition, and the complexity of the local models are given. Moreover, it is shown that any continuous function can be approximated to arbitrary uniform accuracy this way using polynomial local models of arbitrarily low order (like linear or constant local approximation) on a compact domain. Similar bounds and approximation results are given by (Omohundro 1987, Kosko 1994, Zeng and Singh 1994). As mentioned above, these results are quite intuitive and can almost be seen without any mathematical proof.

The curse of dimensionality

An inherent problem with all function approximation approaches that are based on partitioning of the function's domain is the curse of dimensionality: *With an increasing number of variables on which the function depends, the number of partitions required in a uniform partitioning will increase exponentially.* Consequently, a uniform partitioning is undesirable and unrealistic for anything else than low complexity problems (Friedman 1991). Fortunately, a uniform partitioning is usually not necessary. The reasons for this are diverse:

- First, the required accuracy of the model or controller may be significantly higher in some operating regions than in others. For instance, in some regions it may be necessary to have an accurate model or controller in order to optimise the control performance, while in other regions there are other criteria which are important. Fulfilment of these criteria need not always require an accurate model or controller.
- Second, in a dynamical system a large fraction of the state-space will typically be infeasible, in the sense that the system can never be in these states because they are not compatible with the physics of the system.
- Thirdly, if the local models or controllers are sufficiently complex, they will typically describe the system adequately along certain axes, while inadequately along other axes. It can be seen that it is sufficient to further decompose into operating regimes only along those axes which are inadequately described by the local model or controller (Johansen and Foss 1993*a*). This is a major advantage over simpler partitioning approaches that apply very simple local models (such as radial basis function and wavelet series expansions).
- Finally, the complexity of the system is typically not uniform. Hence, sometimes a simple local model or controller will be sufficient in a large operating regime, while in other cases a more complex local model or controller may be necessary in a smaller operating regime.

Development of a simple non-uniform decomposition of the operating range is often difficult. Prior knowledge about the system, or careful examination of large amounts of empirical data is the key to achieving this goal. Hence, the curse of dimensionality can often be 'warded off', at least to some extent.

Reducing the dimension of the scheduling variable

Consider the development of a model of the form

$$\dot{x} = f(x, u).$$

It is easy to see that such a model can always be written

$$\dot{x} = a(x, u) + A(x, u)x + B(x, u)u, \tag{1.3}$$

which is a form that emphasises the close relationship to linear models of the form

$$\dot{x} = a_i + A_i x + B_i u \tag{1.4}$$

and quasi-linear models of the form

$$\dot{x} = a(z) + A(z)x + B(z)u. \tag{1.5}$$

By comparing (1.3) and (1.5) we observe that the variable z clearly should depend on x and u. Suppose we decompose the operating range into a number of operating regimes where the system is described by linear models of the form (1.4). From (1.3) it is evident that the linearised model's parameters depend on the operating point. However, by a possible change of variables it is clear that under some conditions, one can find a characterising vector z satisfying

$$a(x, u) = a(z), \quad A(x, u) = A(z), \quad B(x, u) = B(z),$$

which contains fewer elements than the total number of elements in x and u (Johansen and Foss 1993*a*). In particular, if only an approximate model is sought, then this reduction in dimension can be significant, and the curse of dimensionality can be reduced directly. The use of prior knowledge about the system to reduce the effects of dimensionality is a recurrent theme throughout this work. How can we define modular or hierarchical structures which produce the right level of local complexity needed to model or control the real system?

Example of dimension reduction

As an example of the above considerations, consider the fermentation reactor described in section 1.2.2. A mass balance is

$$\begin{aligned}
\dot{x}_1 &= \mu(x_4, x_5)x_1, \\
\dot{x}_2 &= v(x_4)x_1 - k_1 x_2, \\
\dot{x}_3 &= k_2 x_2, \\
\dot{x}_4 &= -k_3 \mu(x_4, x_5)x_1 - k_4 v(x_4)x_1, \\
\dot{x}_5 &= k_5(x_5^* - x_5) - k_6 v(x_4)x_1 - k_7 \mu(x_4, x_5)x_1,
\end{aligned}$$

where x_1 is the cell concentration, x_2 is gluconolactone concentration, x_3 is gluconic acid concentration, x_4 is glucose concentration and x_5 is dissolved oxygen concentration. The functions (parameters) μ, v, and $k_1, \ldots, k_7$ depend on temperature and pH in addition to the states as written above. This fact follows directly from an examination of the reaction mechanism and basic principles of chemical reaction kinetics (Bailey and Ollis 1986). We clearly see that with local linear state-space models, then the operating regimes can be characterised by $z = (x_4, x_5)$ in addition to pH and temperature. Hence, the number of variables needed to characterise the operating regimes is reduced from seven to four – a significant improvement.

Such knowledge-based reduction of model complexity is fundamental to the approaches proposed in this book, and will appear in a number of guises, from the analysis of statistical distributions, hierarchical decomposition, graphical networks of behaviours and on to fuzzy logic rule bases.

1.2.4 Combining local models and controllers

Having partitioned the problem and developed a number of operating regimes and local models or controllers within each operating regime, the natural question is how to recombine the submodels – i.e. when and how to "switch" between the local models or controllers. In the discussion above we have not paid any attention to possible problems related to overlap between the operating regimes: for some operating conditions there may exist several local models that are partially relevant. Also, we have not argued whether there should be a sudden switching between the local models, or if there should be a smooth transition. The purpose of this section is to address these questions.

Hard partitions and discrete logic

The basic idea of discrete partitions is that at each operating point, exactly one local model or controller is chosen as a deterministic function of the operating point. This is often referred to as mode switching (Hilhorst *et al.* 1991, Hilhorst 1992, Söderman *et al.* 1993), or piecewise models and controllers (Opoitsev 1970, Dorofeyuk *et al.* 1970, Kasavin 1972, Rajbman *et al.* 1981, Bellman 1961, Haber *et al.* 1982, Omohundro 1987, Billings and Voon 1987, Farmer and Sidorowich 1987).

Related to the fermentation example decomposition in Figure 1.7, this means that the partition based on glucose concentration splits the set of operating points into two parts, depending on whether the glucose concentration is below or above the threshold that defines the boundary between low and high glucose concentration, see Figure 1.9.

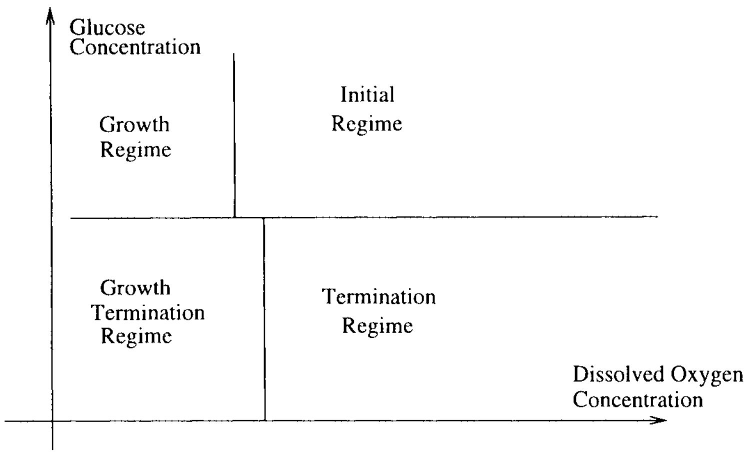

Figure 1.9 Rough decomposition of the operation of a fermentation reactor on the basis of two variables.

Convenient frameworks and representations that can be applied to describe such model or controller behaviour include decision trees (Breiman *et al.* 1984, Strömberg *et al.* 1991*b*, Sanger 1991*b*), discrete logic, expert systems, and hybrid systems (Barton and Pantelides 1994, Pettit and Wellstead 1993, Bencze and Franklin 1995, Simonyi *et al.* 1989, Konstantinov and Yoshida 1991, Konstantinov and Yoshida 1989) as well as variable structure systems (Utkin 1977, Badr *et al.* 1991). The representation in Figure 1.7 is essentially a decision tree. A discrete logic or expert system equivalent would be a number of logical statements of the form

```
IF glucose concentration IS HIGH,                    (1.6)
  AND dissolved oxygen concentration is low,
   THEN the system is operating in the Growth Regime.
```

In practice, such operating regime based models and controllers can be implemented using anything from simple programmable logic systems to sophisticated expert systems. A supervisory system based on discrete logic that lies on top of the traditional control system to handle various situations such as startup, shutdown, exceptions, safety and product changes, is the standard approach to dealing with wide operating range plants in industry.

Assume the modelling problem is a static function approximation problem, where local approximations $f_1, f_2, ..., f_{n_{\mathcal{M}}}$ are known for each operating regime. The operating regimes must form a complete partition of the operating range, and they must not overlap. The global approximation is then the piecewise approximation

$$f(u) = \sum_{i=1}^{n_{\mathcal{M}}} f_i(u)\mu_i(u),$$

where μ_i is the characteristic function for the set of points that defines the operating regime with index i.

Finite state automata

Switching between the operating regimes can also be described using a finite state automaton. A finite state automaton consists of a finite number of discrete states, each corresponding to an operating regime, or a functional state (Halme 1989, Branicky 1994, Zhang *et al.* 1994), see also Chapters 4 and 5. Transition between discrete states is described by a discrete state transition function. Suppose that within each discrete state the dynamics are described by difference equations. This leads to the hybrid model

$$q(t+1) = \mathcal{A}(q(t), x(t), u(t)), \quad (1.7)$$
$$x(t+1) = f_{q(t)}(x(t), u(t)), \quad (1.8)$$

where x is the continuous state and q is the discrete state. The deterministic function $\mathcal{A}$ represents the discrete state transition function.

Soft partitions and fuzzy logic

In some cases it may not be natural to have a sudden change between operating regimes. This may for example be the case when the operating regimes are characterised by different behaviours or mechanisms which change gradually as the operating point moves between different operating regimes. Most physical phenomena have this property. In such cases one can describe the operating regimes as overlapping sets and implement a smooth deterministic transition between them.

Consider again the decomposition described in Figure 1.9. We would like to represent that the boundary between the operating regimes characterised by low and high glucose concentration is soft. Hence, when the glucose concentration is near the overlapping boundary between low and high glucose concentration, then a blend of both local models or controllers is applied.

Frameworks that can be applied to describe such soft boundaries between operating regimes include fuzzy sets and fuzzy logic (Zadeh 1973, Takagi and Sugeno 1985) and interpolation methods (Stokbro *et al.* 1990*b*, Jones and co-workers 1991, Johansen and Foss 1993*a*). Fuzzy sets are characterised by gradual membership and are a very natural way of describing an operating regime. In fuzzy logic, the logical statement (1.6) is interpreted in terms of fuzzy sets definitions of the terms HIGH and LOW, see Figure 1.10. The theory of fuzzy logic also defines natural ways of making inference on the basis of such a rule-base. The resulting inference mechanism (Zadeh 1973, Takagi and Sugeno 1985) can be viewed as an interpolation algorithm that gives more or less weight on the local models or controllers in the different operating regimes, depending on the operating point.

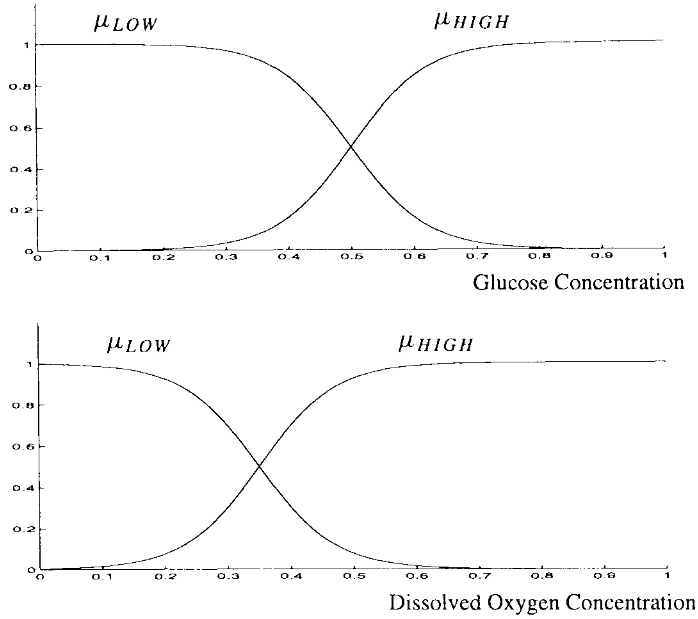

Figure 1.10 Fuzzy membership functions for representation of the terms HIGH and LOW.

Consider again the function approximation problem described above. The global approximation is now an interpolation of the local approximations:

$$f(u) = \sum_{i=1}^{n_{\mathcal{M}}} f_i(u)\rho_i(u),$$

where the smooth weighting functions $\rho_1, \rho_2, ..., \rho_{n_{\mathcal{M}}}$ provide soft transitions between the operating regimes. Typically, the weighting functions satisfy

$$\sum_{i=1}^{n_{\mathcal{M}}} \rho_i(u) = 1$$

for all u. In the case of a fuzzy logic inference (Takagi and Sugeno 1985), then

$$\rho_i(u) = \frac{\mu_i(u)}{\sum_{j=1}^{n_{\mathcal{M}}} \mu_j(u)}, \tag{1.9}$$

which clearly satisfies the equation above. The function μ_i is now the membership function for the fuzzy set that represents the operating regime with index i. The inference mechanism (1.9) is studied in (Johansen 1995).

Probabilistic approaches to partitioning

Both with the hard and soft partitioning approaches mentioned above, the characterisation of operating regimes was deterministic. An alternative is to use statistical methods to infer which operating regime is most appropriate at each time instant.[1]

The basis for mixture models (Titterington *et al.* 1985, McLachlan and Basford 1988) is that each local model or controller has an associated probability density that indicates how correct or appropriate it is (Jacobs *et al.* 1991, Jordan and Jacobs 1994, Petridis and Kehagias 1996). The basis functions can be regarded as the components of a mixture density model. For example, the function approximation problem has the following solution

$$f(u) = \sum_{i=1}^{n_{\mathcal{M}}} f_i(u) P(i|u), \tag{1.10}$$

where $P(i|u)$ is the posterior probability for the local model or controller with index i being the correct one, given the data. Depending on the probabilistic assumptions, the posterior $P(i|u)$ can be computed in a number of different ways. For example, using just priors $\rho_1, \rho_2, ..., \rho_{n_{\mathcal{M}}}$

$$P(i|u) = \frac{\rho_i(u)}{\sum_{j=1}^{n_{\mathcal{M}}} \rho_j(u)} \tag{1.11}$$

(cf. (1.9)), or priors modified by the data according to Bayes' law

$$P(i|u) = \frac{p(u|i)\rho_i(u)}{\sum_{j=1}^{n_{\mathcal{M}}} p(u|j)\rho_j(u)},$$

where $p(u|i)$ is the probability density function for the input u given that local model or controller with index i is the correct one. For dynamic modelling, an approach based on statistical pattern recognition and decision theory is described in (Skeppstedt *et al.* 1992). Moreover, there is a long tradition within the control community with multiple model estimation based on Kalman filter banks and Markov models, e.g. (Lainiotis 1976*a*, Athans *et al.* 1977, Greene and Willsky 1980, Lund *et al.* 1991, Lund *et al.* 1992, Blom and Bar-Shalom 1988) and Chapter 11. The Markov model is a probabilistic relative of the finite state automaton. Transition between the discrete states is described in terms of probabilities or probability densities. Suppose that a discrete time process is modelled by a Markov chain with discrete state $q(t)$ taking values in a finite set $\{1, ..., n_{\mathcal{M}}\}$. At each time step, the probability of transition between j and i is

$$p_{ij} = P[q(t) = i | q(t-1) = j]$$

with

$$\sum_{i=1}^{n_{\mathcal{M}}} p_{ij} = 1, \quad \text{for all } j = 1, ..., n_{\mathcal{M}}. \tag{1.12}$$

In general, the transition probabilities may be densities depending on the system state, input and some parameters. In the static function approximation case, the density may depend only on the input u and the parameters W:

$$p_{ij} = p_{ij}(u(t), W).$$

[1] The relationship between probabilistic and fuzzy methods is interesting, as in some ways, fuzzy logic is deemed *'orthogonal to probability theory as it focuses on ambiguities in describing events, rather than uncertainty of occurrence or non-occurrence'* (Pearl 1988), but it is possible to describe fuzzy logic in terms of uncertainty of which label to use. In the multiple model case a fuzzy representation of the model validity functions would amount to uncertainty about which local model was the 'correct' one for a given non-fuzzy operating point.