Conference Publication Number 79

Computers for Analysis and Control in Medical and Biological Research

Conference on

Computers for Analysis and Control in Medical and Biological Research

7—9 September 1971

Organised by
The Control and Automation Division of The Institution of Electrical Engineers

in association with
The Biological Engineering Society
The British Computer Society
The Institute of Electrical and Electronics Engineers (UK and Republic of Ireland Section)
The Institute of Mathematics and its Applications
The Institution of Electronic and Radio Engineers

Venue
Ranmoor House, University of Sheffield

Contents

The Institution of Electrical Engineers is not, as a body, responsible for the opinions expressed by individual authors or speakers

Contents

Contents

List of Authors

List of Authors

IDENTIFICATION OF METABOLIC SYSTEMS

E.R. Carson, L. Finkelstein

1. Introduction

Increasingly, interest is being shown in the measurement and identification of dynamic biological systems. Metabolic systems, involving the complex chemical exchanges taking place in living organisms, exhibit a number of phenomena which make them of special interest as subjects for joint work by physiologists and systems scientists.

It is becoming evident that the benefits of such investigations are bilateral. On the one hand, the biologist is confronted with the need to produce a systematic explanation of complex biochemical systems, leading to greater insight and clarification of assumptions made. For the systems scientist the benefits are twofold. Firstly it enables investigations to be carried out into the organisation and operation of complex, hierarchical control systems which actually function, and secondly it offers interesting problems of measurement, revealing shortcomings in existing identification methods.

By way of exemplification, in this paper it is proposed to consider the metabolism of plasma proteins, although the principles and concepts employed may be extended more generally. This investigation was carried out jointly with members of the Medical Unit of the Royal Free Hospital, London.

2. Modelling Techniques

In order to attempt any system identification, a mathematical model must first be formulated for the system. Traditionally, the description of chemical processes taking place in biological systems is formulated in terms of compartmental analysis; the compartment being defined as all of a substance in a particular form, or all the substance in a particular location, or all the substance in a particular form and location. So, for example, all radioactively labelled albumin or all albumin in the plasma or all radioactively labelled albumin in the plasma can each be considered as a compartment.

Thus in formulating a model of a metabolic system, the system is first divided into relevant and convenient compartments. The mathematical model then consists of the mass balance equations for each compartment and relations describing the rate of material transfer between compartments.

Instrument Systems Centre, Department of Systems and Automation,
The City University, London E.C.1.

The general form of the mass balance equation is as follows. If X_c is the quantity of substance in compartment C which interchanges matter with other compartments constituting its environment E then:

$$\frac{dX_c}{dt} = \sum f_{EC} - \sum f_{CE}$$

where $\sum f_{EC}$ represents the summation of the rates of mass transfer into C from all relevant compartments and $\sum f_{CE}$ the summation of the rates of mass transfer from C to other compartments of the system. The equation is shown graphically in figure 1.

Considering the dynamics of biochemical reactions, it can be shown, in general, that the rate of change of mass of a substance undergoing a chemical transformation is proportional to an integral power of its concentration.[1] Where several substances take part in a reaction, the rate of change is proportional to the product of integral powers of the concentration.

Thus for an irreversible single stage reaction

$$A + B \longrightarrow \text{Products}$$

we have

$$\frac{dA}{dt} = kC_A^{\alpha} C_B^{\beta}$$

where A = quantity of A
B = quantity of B
C_A = concentration of A
C_B = concentration of B

α is termed the order of the reaction with respect to A, β the order with respect to B and $\alpha+\beta$ the order of the reaction.

A first order reaction of the type

$$A \longrightarrow \text{Products}$$

gives with the same notation as above

$$\frac{dA}{dt} = -kC_A$$

or if the reaction takes place in a compartment of constant volume

$$\frac{dC_A}{dt} = -KC_A$$

$$\text{and } \frac{dA}{dt} = -KA$$

Representing this first order compartmental reaction in signal flow diagram form (figure 2), it can be seen that the concentration dependence of the reaction represents an inherent negative feedback loop tending to stabilise the quantity of a substance in a compartment.

Chemical reactions in biological organisms commonly form part of a chain, the products of one reaction stage forming the starting reactants of the next stage.

Plasma Protein Metabolism

Consider the metabolic system of the plasma protein albumin. As the starting point of the system consider the injection into the bloodstream of a quantity of radioactively labelled sodium carbonate. In terms of functional biology, the pathways taken by the radioactive tracer may be represented by the system of interconnected compartments shown in figure 3. Urea pathways are also included since the precursor amino acid arginine is used to synthesise both plasma proteins and urea.[2]

Mathematical formulation may thus be attempted by considering the system of interacting compartments and writing a mass balance equation for each. The tracer material introduced experimentally forms labelled compounds which behave, metabolically, in an identical fashion to the same substances, unlabelled, occuring naturally. Thus each of the postulated compartments will contain both labelled and unlabelled material. A typical mass balance equation is that obtained for the albumin synthesis compartment:-

$$\frac{d}{dt}(A_a + a_a) = f_{ra}(R_r + r_r) - f_{ad}(A_a + a_a)$$

where A = mass of unlabelled albumin
a = mass of labelled albumin
R = mass of unlabelled arginine
r = mass of labelled arginine

The subscripts, indicating the location of the substance are

a = albumin synthesis compartment
d = albumin storage compartment
r = liver arginine compartment.

Double subscripts indicate material transfer from one location to another.

In a similar fashion, mass balance equations can be obtained for all the compartments.[2,3]

3. Model Linearisation

Metabolic systems are, in behaviour, grossly non-linear. In order to simplify the problems of identification, however, it is proposed to consider a linearised small perturbation model. The reasons for this are twofold. Firstly, the dymamics and control processes of the complex system can be more readily visualised. Secondly, in practice, experimental tests only give information concerning the dynamics of the labelled substance. The labelled substance in any compartment can be considered a small perturbation and the dynamics of the labelled substance can be written in terms of linearised equations. For example the mass balance equation for albumin in the albumin synthesis compartment becomes:-

$$\frac{da_a}{dt} = \frac{\partial f_{ra}}{\partial R_r} \cdot r_r - \frac{\partial f_{ad}}{\partial A_a} \cdot a_a$$

The resulting set of linearised equations can then be expressed in signal flow diagram or transfer function form (figure 4).

4. Experimental Measurement

Mathematical models, such as that derived for plasma protein metabolism are valid only to the extent that the assumptions made concerning the systems biology are true. Thus experimental measurements are required in order to test the proposed nature of the system structure and to assign numerical values to its parameters.

Experimental tests carried out upon 'in vivo' systems are clearly limited since any perturbation which results must in no way impair the normal functioning of the biological mechanisms. Thus in metabolic systems experiments are limited to impulse testing as realised in the tracer technique. All tracers employed in this study are biosynthetic. That is to say that they are labelled substances which occur naturally in unlabelled form in the system and are thus assumed to be metabolically identical with the unlabelled substance.

In this investigation five tests were carried out:

(i) ^{14}C Carbonate Injection

Sodium Carbonate labelled with ^{14}C was injected into the bloodstream and measurement of the albumin radioactivity made in the plasma.

(ii) ^{14}C Carbonate Injection

From a similar injection, measurement of urea label concentration was made in the urea initial mixing compartment.

(iii) ^{131}I or ^{125}I Albumin Injection

Albumin labelled with ^{131}I or ^{125}I was injected into the

bloodstream and measurement of the albumin radioactivity made in the plasma.

(iv) ^{13}C Urea Injection

Urea labelled with ^{13}C was injected into the bloodstream and measurement was made of the enrichment of urea concentration by its stable isotope in the initial mixing compartment.

(v) Blood ^{14}C Bicarbonate Clearance Curve

Labelled ^{14}C Sodium Carbonate was injected into the bloodstream where measurement was made of the label concentration in the blood carbonate-bicarbonate.

Typical results for tests are shown in Figure 5. In each of these experiments, the radioactivity dosage administered was in the range 100 - 200 μcuries, and the volume of blood removed at each of the sampling instants was in the range 20 - 40 ml.

Test data available from the 'in vivo' system using the tracer technique is severely limited. The reasons for this are threefold:

(i) The number of sites of injection and sampling is severely limited on medical grounds.

(ii) The quantity of radioactively labelled substance that can be administered is restricted to prevent biological radiation damage. Resulting from this, the levels of radiation that are required to be measured may be low compared with background effects. Thus considerable error can be introduced by the subtraction of the baseline.

(iii) Since the measurements involve withdrawals of blood samples in which the radioactivity is estimated, the number of tests taken must be limited for the wellbeing of the patient.

5. Dynamic Reduction

With the foregoing constraints imposed upon experimental design, it is clearly impossible to attempt complete identification of the linearised model. Thus further reduction techniques have to be adopted.

The criteria adopted for dynamic reduction were that compartments with rapid dynamics, that is to say time constants of the order of 10^{-2}h or less, which would not show clearly in the test data could be treated as pure gains. Also, pathways wherein small quantities of substance were fed back along pathways with long time constants (10^{2}h or more) were neglected on the grounds that they would have little effect over the experimental period of ten hours.

Having estimated the various time constants[2], the system could be reduced to that shown in Figure 6.

6. Parameter Estimation

Conventional techniques for the parameter estimation of compartmental models from impulse response measurements are based on the assumption that the latter part of the response being fitted can be considered to be a single exponential (or linear if the response is plotted in logarithmic form), corresponding to the longest time constant. This terminal portion is thus identified. If this exponential component is then subtracted from the response, the terminal component which remains can again be considered to be a single exponential and may be identified in a similar fashion. This process is repeated until all the exponential components have been identified.[4] The success of this method, however, is dependent upon adequate separation of the time constant values[5], which is not generally true and is in any case not known a priori.

In this study no such assumption is made. Instead, the model is reduced to its minimum significant form and the whole of this is identified simultaneously. Digital computer based adaptive procedures were adopted in the parameter estimation process.

By way of exemplification, let us consider the transfer functions (Figure 6) resulting from injection at points (a) and (c) and measurements at points (b) and (d) respectively. These transfer functions describe the responses corresponding to testing with ^{131}I labelled albumin and ^{13}C labelled urea. These transfer functions consist of the interconnection of three first order lags, and thus the impulse responses obtained may be fitted in terms of the summation of three exponential components. In other words the modelled impulse response:

$$y_m(t) = \sum_{i=1}^{3} A_i e^{-a_i t}$$

Since the data available from the tests is contaminated by measurement noise and limited to not more than ten points on the impulse response for the reasons already outlined, and furthermore since the data is not equi-spaced, analytical exponential fitting techniques cannot successfully be applied.[6] Thus an adaptive technique was employed, based upon the system shown in Figure 7.

The reduced model was set up and its impulse response:

$$y_m(t) = \sum_{i=1}^{3} A_i e^{-a_i t}$$

was compared with data available from the experimental impulse response $y_e(t)$ at the experimental sampling times. The values of

$$(y_e(t_j) - y_m(t_j))$$

at the experimental sampling times, t_j were used in formulating an error criterion:

$$E = \sum_{j=1}^{n} (y_e(t_j) - y_m(t_j))^2$$

which could be employed in adjusting the model in order to produce a better fit. Systematic search procedures for the optimum parameters utilised the Rosenbrock hill climbing technique.[7]

Initially a least squares error criterion was used in the search procedure, with a reduction in error function indicating that a particular parameter adjustment was leading towards a best fit. The individual errors were not weighted. Typical results obtained, fitting the data in terms of the expected three exponentials are shown in Table 1.

7. Results and Analysis

The fit obtained was considerably better than the experimental error.

It was then decided to examine the uniqueness of the fit. The fitting as described in terms of three exponentials was repeated using the Conjugate Gradient hill climbing technique.[8] It was then observed that the results while yielding a similar value of mean square error gave different coefficients (Table 2.)

Another examination was carried out to see whether other models could fit the results. Two and four exponential models were chosen, results of the fitting being shown in Table 1.

Thus it can be seen that the three exponential model can explain the results but not uniquely. It can also be seen that the parameter estimates are not unique.

8. Discussion and Conclusions

Mathematical models of a complex metabolic system have been formulated on the basis of current understanding of the biology of the system. The mathematical model constitutes a good method of description of the system offering improved insight into its structure and performance.

It can be seen, however, that impulse tests on multi-compartmental models of biochemical systems cannot of themselves validate a model. The validity of the model must rely on knowledge of the unit processes and their connection.

If there is to be any prospect of determining the parameters of the system from its responses, at least the number and interconnection of the dominant components must be known. Even, however, with this knowledge of the number of compartments in a multi-compartmental model which is to be fitted to an impulse response contaminated with noise, the parameter estimates will not be unique.

To overcome the problems outlined, two parallel approaches are

being investigated.

Firstly, the modelling techniques which have been expounded have assumed that it is reasonable to describe complex chemical processes in terms of mass balance equations and simplified considerations of transport. This clearly entails gross approximation. It is proposed, therefore, to model in terms of the unit processes of biochemistry such as enzyme reactions and expanding compartments and to include a more detailed treatment of the transport processes. This should provide greater insight into the process dynamics and thereby indicate methods whereby the estimation procedures might be improved.

Model reduction and parameter estimation techniques are also being examined. Among approaches being tried is the use of various loss and risk functions which better reflect the varying confidence attached to the different portions of the response curve or which might improve the estimates of individual time constants.

It is proposed to report on this approach in due course.

Bibliography

1. Bray, H.G. and White, K. Kinetics and Thermodynamics in Biochemistry (Churchill 1966).

2. Barr, C.D., Carson, E.R., Finkelstein, L. and Jones, E.A. A Study of the Dynamics of Plasma Protein Metabolism, Proc. IFAC IV, Warsaw, 1969.

3. Carson, E.R. and Finkelstein L. The Dynamics and Control of Chemical Processes in Man, Measurement and Control 3, T157 - 167,

4. Matthews, C.M. The Theory of Tracer Experiments with ^{131}I labelled plasma proteins, Physics in Medicine and Biology 2, 36 (1957).

5. Myhill, J. Investigation of the effect of Data Error in the Analysis of Biological Tracer Data from Three Compartment Systems, J. Theoretical Biology 23, 218-231, (1968).

6. Wilson, R. A Study of the Methods of Parameter Estimation for Multicompartment Models of Physiological and Biological Systems, M. Sc. Thesis, The City University, London (1970).

7. Rosenbrock, H.H. and Storey, C. Computational Techniques for Chemical Engineers (Pergamon 1966).

8. Davis, R.H. and Roberts, P.D. Method of Conjugate Gradients applied to self-adaptive digital control systems, Proc. I.E.E. 115 562 - 571, (April 1968).

Table I

Fitting of Data by Rosenbrock Hill Climbing Technique using Least Squares Error Criterion

Data 'A'					Data 'B'				
Time (min.)	Test y_e Values	Fitted Values y_m			Time (min.)	Test y_e Values	Fitted Values y_m		
		2 Exp.	3 Exp.	4 Exp.			2 Exp.	3 Exp.	4 Exp.
13•1	1•178	1•175	1•174	1•154	5•0	0•896	0•894	0•890	0•892
41•0	0•774	0•778	0•777	0•804	21•0	0•501	0•511	0•520	0•517
84•0	0•609	0•611	0•620	0•636	43•0	0•475	0•466	0•466	0•466
125•0	0•572	0•557	0•564	0•559	61•0	0•444	0•451	0•451	0•451
192•0	0•491	0•503	0•498	0•478	109•0	0•425	0•416	0•416	0•416
235•0	0•472	0•472	0•464	0•444	154•0	0•367	0•386	0•385	0•385
362•0	0•387	0•393	0•384	0•377	273•0	0•356	0•316	0•315	0•315
470•0	0•342	0•336	0•335	0•341	383•0	0•268	0•262	0•263	0•262
594•0	0•283	0•281	0•292	0•307	423•0	0•228	0•245	0•246	0•245
					552•0	0•185	0•197	0•199	0•198
Value of error criterion		0•00044	0•00050	0•0042			•00297	•00316	•00277

Table 2

Fitting of Test Responses to 3 Exponential Component Model using Least Squares error criterion

$$y_m = Ae^{-at} + Be^{-bt} + Ce^{-ct}$$

DATA 'B'

Parameter	Rosenbrock Method	Conjugate Gradient Method
A	0·1625	0·505
a	0·00114	0·00172
B	0·3378	0·770
b	0·00198	0·1905
C	0·8181	0·299
c	0·1462	0·2260
Value of Error Criterion	0·00316	0·00283