Preferred citation: J.F. Coales. The future of control engineering. In **Papermaking Systems and their Control**, *Trans. of the IVth Fund. Res. Symp. Oxford, 1969*, (F. Bolam, ed.), pp 653–672, FRC, Manchester, 2018. DOI: 10.15376/frc.1969.2.653.

THE FUTURE OF CONTROL ENGINEERING

J. F. COALES, O.B.E., M.A., F.I.E.E., F.Inst.P., F.I.E.E.E., F.R.S.
Professor of Engineering, University of Cambridge

Introduction

UP TO a few years ago, all industrial processes were either batch processes or were operated on a steady state basis, although some (like papermills, automatic looms and steel-rolling mills) operated in steady state only for a limited period until the specification or grade of the output had to be changed. Thus, industrial processes can be divided into three classes—

1. Batch processes.
2. Continuous processes.
3. Quasi-batch processes.

Batch processes

In the early days, all processes were batch processes and this meant that they were all time-varying and therefore difficult to control accurately. They were in fact controlled by time schedules based on previous knowledge of the process, but this is in general possible only when both the desired product and the raw materials are closely specified. Any physical process can be represented by the simple box on the left of Fig. 1 in which the raw materials are represented by **v**, which includes composition, quantity and such other relevant variables as temperature, viscosity and physical dimensions, depending on the nature of the process. These raw materials are operated on by the plant that has certain characteristics such as physical dimensions, temperature and pressure, which appear as coefficients in the set of differential equations, by which they may be represented mathematically.

In addition to the characteristics of the plant, the characteristics such as rates of reaction and catalyst activity of the process must also be known and they also appear as coefficients in the set of differential equations. All these coefficients are denoted by **a**. The raw materials having been operated on by the plant and process, their composition, temperature, etc. (or in the case of

Under the chairmanship of D. Attwood

mechanical processes, physical dimensions) change with time and these quantities can be represented by **x**, which is called the state variable of the system. Certain of the components of this state variable correspond to those components of the product which appear in its specification designated **z** and the object is to control the process by the application of heat or chemical additives in the case of chemical and metallurgical processes or by moving a tool in metal-working, designated by **u**, to change the state variable so that the relevant components of **x** approximate to the desired output **z** in the shortest possible time. It should be noted that the **u** values, called the controlled variables, always imply the application of energy or material to the process.

All physical processes can be represented by a set of simultaneous differential equations of the form—

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{a}, \mathbf{u}, \mathbf{v}, \boldsymbol{\upsilon}, t) \quad . \quad . \quad . \quad . \quad (1)$$

in which $\boldsymbol{\upsilon}$ represents unwanted disturbances that cannot be measured.

If the **a**, **u** and $\boldsymbol{\upsilon}$ terms are known, then the process is completely defined by equation (*1*) and the **x** can be calculated for any t. In the case of a batch process, **v** will be the initial value of **x** at time $t = 0$ (that is, $\mathbf{x}_0$) and the **v** will be functions of time. In general, the material throughout the plant will not be homogeneous and diffusion processes will be involved, resulting in a set of partial differential equations, which makes the derivation of a mathematical model in the form of equation (*1*) extremely difficult, if not impossible. Fortunately, however, continuously variable systems of this type can frequently be broken up into sections that can be represented by lumped parameters, but this inevitably increases the number of state variables **x** and plant coefficients **a**. At the present time, it is impossible to solve large sets of partial differential equations involving time even with the largest computers available and so from now on we shall consider only sets of ordinary differential equations and assume that any distributed parameter system in which we are interested can be approximated by a lumped parameter system.

It will now be apparent that, in order to model a batch process, it is necessary to know how the state variable **x** changes with time for different values of **u** over the whole range of **x** met in practice. If the **u**, **v** and **x** values can be measured as functions of time, then it is possible to hillclimb on the coefficients **a** and so determine them. The coefficients **a** will then truly represent the plant characteristics. Unfortunately, two difficulties usually arise. Firstly, there are unwanted disturbances, some of which may arise from the ambient conditions that cannot be measured and that are represented by $\boldsymbol{\upsilon}$ in both Fig. 1 and equation (*1*). Thus, the values of **a** that can be obtained by regression analysis or hillclimbing can be no more than estimated values designated by $\hat{\mathbf{a}}$. Secondly, it is not usually possible to measure all the compoents of **x** required

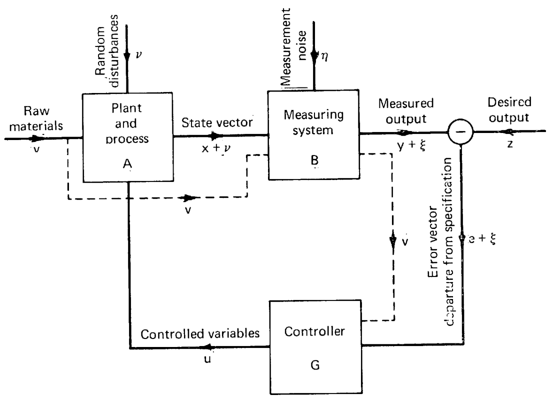

Fig. 1

to represent the state of the plant and process by a set of first-order differential equations of the form of equation (*1*) and, in any case, they cannot be obtained directly, but only by means of measuring instruments. These have to be included in the system and are represented by the box on the right of Fig. 1. This means that we must write—

$$\mathbf{y} = B(\mathbf{x}+\boldsymbol{\upsilon})+\boldsymbol{\eta} \qquad . \quad . \quad . \quad . \quad . \qquad (2)$$

where $\mathbf{y}$ is the measured values of those components of $\mathbf{x}$ that can be measured and B is called the measurement matrix. All measurements are to a greater or lesser extent noisy and so we have to add a noise vector $\boldsymbol{\eta}$ in equation (*2*). By combining equations (*1*) and (*2*), we get—

$$\dot{\mathbf{y}} = \varphi(\mathbf{y},\mathbf{a},\mathbf{u},\mathbf{v},\boldsymbol{\upsilon},\boldsymbol{\eta},t) \qquad . \quad . \quad . \quad . \quad . \qquad (3)$$

from which theoretically the estimated values $\hat{\mathbf{a}}$ of the plant parameters can be determined. This rather formidable equation can fortunately very often be simplified, because the noise vectors $\boldsymbol{\upsilon}$ and $\boldsymbol{\eta}$ are uncorrelated with the other variables. In this case, equation (*3*) can be rewritten in the form of equation (*4*)—

$$\dot{\mathbf{y}} = \varphi(\mathbf{y},\mathbf{a},\mathbf{u},\boldsymbol{\upsilon},t)+\xi \qquad . \quad . \quad . \quad . \quad . \qquad (4)$$

where ξ is the combined noise vector.

When a number of the components of $\mathbf{x}$ cannot be measured, however, equation (*4*) is not a true representation of the system and there are insufficient equations to include all the values of $\mathbf{a}$. In this case, we have to start with a structural model of the system in form of a set of higher order differential equations of the form of—

$$\mathbf{y}^n = f(\mathbf{y},\mathbf{y}^1,\ldots,\mathbf{y}^{n-1},\boldsymbol{\alpha},\mathbf{u},\mathbf{v},t)+\xi \qquad . \quad . \quad . \qquad (5)$$

in which $\mathbf{y}^1 \ldots \mathbf{y}^n$ are the first and higher derivatives of the $\mathbf{y}$ vector and the $\boldsymbol{\alpha}$ values are the plant parameters—that is, the coefficients (some representing time constants) and indices, if the system is non-linear, appearing in the equations. The $\mathbf{y}$ to $\mathbf{y}^n$ being measurable, it is now possible to determine $\hat{\boldsymbol{\alpha}}$, the estimated values of these coefficients and/or indices and so an approximate structural mathematical model is obtained. In general, this will be a set of non-linear equations, but most industrial processes are fortunately operated over a narrow range of variables and the equations can usually be piecewise linearised with a small number of linear zones. In this case, the structural model can be simplified to the normal* linear form of equation (*6*), where A' is the plant matrix and C' is the control matrix—

$$\dot{\mathbf{y}} = A'(\mathbf{y},t)+C'(\mathbf{u},t),\ \mathbf{y}_0 = \mathbf{v} \qquad (6)$$

Theoretically, it is now possible by the calculus of variations to calculate the u in terms of the error matrix, equations (*7*) and (*8*), which will bring $y(t)$ to the desired value z with minimum expenditure of energy or any other criterion J on which it may be required to optimise the system. The criterion J is frequently called a *cost function*, since it should include, besides the measure $g_1(\mathbf{e})$ of the divergences of the output product from the desired specification, the cost of applying the controls $g_2(\mathbf{u})$ and the fixed costs such as labour and capital charges $g_3(\mathbf{w})$. The form of the cost function J will be as given in equation (*9*), in which the kernel function is usually made up of the algebraic functions $g_1(\mathbf{e})$, $g_2(\mathbf{u})$ and $g_3(\mathbf{w})$. By minimising J, the $\mathbf{u}$ are obtained as functions of time as given in equation (*9*)—

$$\mathbf{u} = G(\mathbf{e},\mathbf{v},t) \qquad (7)$$

$$\mathbf{e} = \mathbf{z}-E\mathbf{y} \qquad (8)$$

$$J = \underset{T\to\infty}{\text{Limit}} \int_0^T [g_1(\mathbf{e})+g_2(\mathbf{u})+g_3(\mathbf{w})]dt \qquad (9)$$

In the case of a batch process, equation (*6*) must be integrated to give equation (*10*), which completely defines it.

$$\mathbf{y} = H(\mathbf{a},\ \mathbf{v},\ \mathbf{u},\ t), \qquad \mathbf{y}_o = \mathbf{v},\ \mathbf{y}_T = \mathbf{z}/E \qquad (10)$$

where $t = 0$ and $t = T$ are taken to be the times at the start and completion of the process, respectively.

Continuous processer

With the growth of the chemical and oil industries, the need to process large quantities of fluids led to the development of continuous processes that proved much easier to control than batch processes. The great advantage of

* The term *normal form* is used for a set of first-order differential equations

continuous processes is that, because the product is being removed and raw materials are being added continuously, the rates of reaction remain constant and so the process can be operated in steady state and is therefore time invariant, provided the plant itself does not change and there are no changes in ambient conditions to affect the process. Time can be dropped from equation (*1*) and it can be represented by a set of differential equations of the form of equation (*11*)—

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{a}, \mathbf{u}, \mathbf{v}, \boldsymbol{\upsilon}) \qquad . \quad . \quad . \quad . \quad (11)$$

Furthermore, provided the plant can be approximated by a number of sections with lumped parameters, no partial derivatives will be involved and equation (*11*) will be a set of simultaneous ordinary differential equations. This makes the process very much easier to control and it will be seen that, provided the raw materials do not change, the $\mathbf{v}$ are constant and so when the plant has been run up to the state when the relevant components of $\mathbf{x}$ are equal to the desired values $\mathbf{z}$, the control variables $\mathbf{u}$ should be kept constant, provided that $\mathbf{a}$ is also constant. Thus, provided there is no change in the plant characteristics, due to wear and tear or changes in catalyst activity or in the composition and flow of raw materials, there is no control problem while the plant is running in steady state.

Unfortunately, it is often impossible to maintain the composition and flow of raw materials exactly constant and some adjustment of the controlled variables is necessary, but the required variations are usually small. This means that we have to retain the $\mathbf{u}$ in equation (*11*). Changes in the plant characteristics due to wear and tear (for example, fouling up of pipes and valves) does occur and changes in the ambient conditions may also cause unexpected fluctuations in the operation of the plant. The unwanted fluctuations and measurement noise may be represented as before by the addition of a random variable ξ as indicated in equation (*12*).

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{a}, \mathbf{u}, \mathbf{v},) + \xi \qquad . \quad . \quad . \quad . \quad (12)$$

Thus, industrial plants have traditionally been designed for the manufacture of closely specified products from closely specified raw materials. On the basis of laboratory and pilot plant experiments, the plant has been designed and manufactured, then it has been set up and maintained to do just this. Under these conditions, if the product fails to meet the specification, no attempt is made to correct this automatically nor can the product be modified, except by modifying the plant or setting it up afresh. If there were no wear and tear of the plant and if the raw materials were always closely within the specification, there would be no need for any feedback control at all, provided the plant was unaffected by external conditions. There are always uncertainties in the operation of the plant, however; raw materials do change from time to time and these result, if they go uncorrected, in changes of the product specification and therefore of spoilt work.

It is to overcome these difficulties that feedback control is used, first of all to maintain such plant parameters as temperature, pressure and fluid flow constant and, secondly, to enable them to be changed to new values should there be changes in raw materials or in the external conditions under which the plant is operating. If the physics and chemistry of the manufacturing processes were fully understood and if the essential physical and chemical properties of the raw materials could all be measured continuously, there would be no need for feedback control, provided the new operating conditions required could be calculated quickly enough to correct for the changes in raw materials. Before the days of automation, this is exactly what the craftsmen used to do. By long experience, the turner knew that, if he got an unusually hard bit of steel, he had to turn it at a lower rate and with a smaller cut and so he adjusted his turning accordingly. There was some feedback, of course, because he had no instruments as a rule to measure the hardness of the steel and it was only by observation of the way it cut that he knew that it was harder than usual.

Similarly (apart from inherently unstable processes), in the control of industrial processes, it is due to ignorance of the characteristics of the materials and the physics of the processes that negative feedback is used to give adequate control. In most processes, it is relatively easy to discover by empirical observations over a long period the effect on the finished product of small changes of temperature and pressure on the one hand or of changes of cutting speed and depth of cut on the other. Thus, by observation of the quality of the finished product, adjustments can be made to correct for small changes in the raw materials or, indeed, of tool or plant wear and tear, even when relatively little is known about the physics and chemistry of the processes. This may be considered the first step in quality control and has resulted over the last 10–20 years in increasing demand for what are known as quality control instruments.

The first step is always to set up local control loops to keep the operating conditions constant, then to add quality control instruments, observations of which are used to alter the operating conditions over a small range to correct for changes of plant parameters and raw materials. Fortunately, these changes have usually been small and relatively slow and it is for this reason that the three-term controller has proved so successful. A change in raw materials implies, of course, a change of the $\mathbf{v}$ in equation (*11*) and changes in the plant parameters a change in the $\mathbf{a}$ of equation (*11*) so that the $\mathbf{x}$ will no longer correspond with the desired output $\mathbf{z}$. Thus, the controlled variables $\mathbf{v}$, the set points in conventional process control, have to be changed in order to achieve a new and satisfactory steady state.

In conventional control, no account is taken of the route by which the plant is brought to the new operating condition nor of the amount of spoilt work that results during the change from one steady state to another. The controllers are set up in such a way, however, that the new steady state is reached in the shortest time possible without appreciable overshoot. With the advent of on-line digital computers, the operation of the three-term controllers can be replaced by calculations carried out in the digital computer as has been done in the early applications of direct digital control. The use of a computer enables the best operating conditions (that is, the set points) for different raw material compositions and plant characteristics to be computed at regular intervals so that, when materials or plant characteristics change, the set points can be corrected automatically. In general, this is the best that can be done with direct digital control until a dynamic mathematical model of the plant and process in the form of equation (*11*) has been developed.

Direct digital control

Many continuous unit processes are now provided with data-loggers by means of which the operating conditions can be recorded and stored and the three-term controllers have been replaced by algorithms in a digital computer program. For many of these processes, very simple algebraic models of the form of equation can be developed—

$$\mathbf{x} = f(\mathbf{a}, \mathbf{u}) \qquad . \quad . \quad . \quad . \quad . \quad (13)$$

in which, because it is a continuous steady-state system, the **v** are included in the state variable **x**.

We can rewrite this—

$$\mathbf{y} = f(\mathbf{a}, \mathbf{u}, \mathbf{v}) \qquad . \quad . \quad . \quad . \quad . \quad (14)$$

where **y** includes the output quantities that are specified, if the functions f are known and, during a specified period of time, the **u**, **v** and **y** are measured and stored, the plant characteristics **a** can be calculated by hillclimbing or regression analysis. Then, assuming the **a** and **v** remain constant during the ensuing period, the **u** for optimum operation in accordance with some predetermined criterion can be calculated. The **u** are then taken to be the required set points for the ensuing period and the plant is brought to the new steady state by means of the controller algorithms. During this ensuing period, the **u**, **v** and **y** are again recorded and stored so that the **a** can be recalculated at the end of this period to take account of any changes in the plant characteristics and new values for **u**, the required set points for the next period, can be calculated and applied.

In this way, the set points are updated at the end of any period in which there is a change of the raw materials; alternatively, small changes of the desired output specification can be accommodated. The system therefore operates as a feedback control system, although only an algebraic model is used. Furthermore, the system is truly adaptive in that it takes account of changes in plant parameters and so it is comparable with a time-varying dynamic system, provided the sampling periods are short compared with rate of change of plant parameters, raw materials, etc. Further non-linearities in the plant model present no special difficulties.

This simple system has limitations, of course, the most important being that the route by which the plant goes from one state to the next is predetermined by the three-term controller algorithms and will not be optimum. Provided the required changes are small and/or infrequent, however, this is of little importance; in any case, the computer will almost certainly give better steady state operation than a human operator. Another limitation arises from the difficulty of carrying out hillclimbing proceedings with a large number of variables even with very powerful computers. At the present time, the maximum number of components of **u**, **v** and **y** that could be accommodated if the **a** are to be calculated in reasonable time is 20–30, depending on the form of the equations; with computers rapidly becoming smaller, more powerful and cheaper, this limitation should disappear in a few years.

If the **a** cannot be calculated sufficiently quickly with the available computer, they can be assumed constant over longer periods, in which case calculation of the **u** (which usually has relatively fewer components) for the new values of the **v** and/or **y** is quite simple, since no hillclimbing or regression analysis is required. This simple method of computer control has considerable advantages when unwanted random fluctuations and/or measurement noise are severe, because the mean values of the variables can be taken over each recording period. Provided the noise is of higher frequency than that of the changes in raw materials or plant parameters, it is only necessary to choose the longest recording period compatible with adequate control, an exchange relationship that is always present in noisy systems.

When the noise spectrum is such that simple averaging does not give adequate rejection, more sophisticated filtering techniques must be used. Statistical methods are particularly well developed for determining algebraic models from noisy data, so noisy data is unlikely to be a serious limitation in the application of this simple DDC to continuous processes. If a dynamic model is required (as it will be, if optimum control of the plant is wanted while going from one steady state to another), parameter estimation in the presence of noise is far more difficult. The optimisation procedure for linear filtering and prediction developed by Wiener has been extended to multi-variable systems by Kalman. If the system is linear, these methods can be applied fairly easily, using a digital computer for up to about 10 variables; for more complex systems, very little can be done at the present time.

Quasi-batch processes

Other than in oil refineries and large chemical works, there are very few continuous processes that operate in steady state producing the same product from the same raw materials year in, year out. The majority of industrial products are still manufactured by batch processes. For some industries such as the motor industry, an attempt has been made, when production quantities are large enough, to achieve steady state conditions by the introduction of transfer machines and flow production. In these cases, there are a large number of parallel operations producing intermediate products, which are then subjected to a second set of operations producing another set of intermediate products and so on until the final product is complete as indicated in Fig. 2.

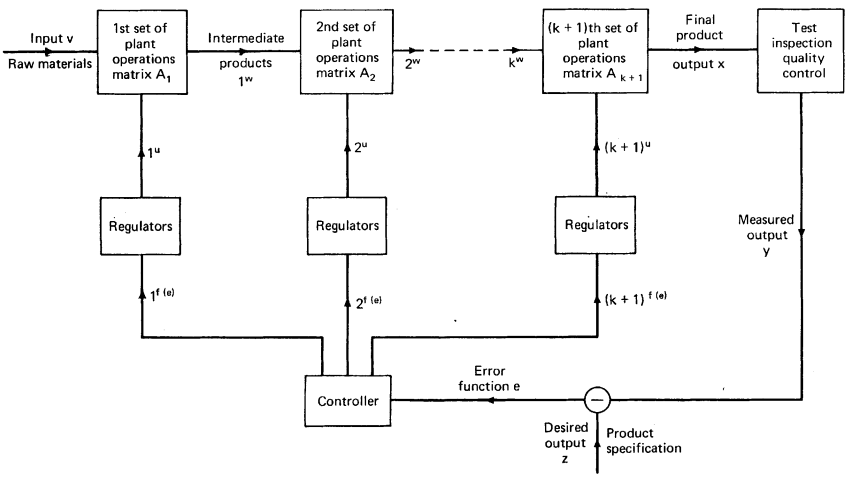

Fig. 2

This is exactly analogous to a chemical plant in which a large number of reactions go on in one unit operation (say, a distillation column), the products from which go to other unit operations for further processing.

There are other industries such as papermaking, glassmaking, steel and textiles in which the plant may operate in steady state for considerable periods while the specification is kept constant, but, because the demand for a particular grade of paper or type of cloth is limited, the specification of the output product from the same machine may have to be changed several times a day. In other cases, although the process is inherently continuous, the plant may for practical reasons have to be frequently stopped and restarted. This applies in a steel rolling mill, where the size of the slab rolled is limited by the size of the coiler. In this case, the ingot has to be rolled and cut into slabs of a certain size to match the coiler. Since the beginning and end of each slab is a gross discontinuity, it is effectively a batch process, although, for the middle 95 per cent of each slab, the mill is operating in steady state.

Similarly, a papermill, when it has been set up correctly to manufacture a particular grade, is in steady state operation and can be accurately represented by equation (*12*). When there is a change of grade, the mill must be brought to a new state in which the $\mathbf{x}$ must have different values and so different controlled variables $\mathbf{u}$ are required. In general, there is little difficulty in controlling the mill in the steady state condition, provided the raw materials can be held constant. As is well known, when a change to a new grade is required, it is difficult to achieve the new steady state in a reasonable time and without a great deal of spoiled product. It is evident that the greatest immediate savings under these conditions can be made not by improving the control of the steady state conditions, which are already relatively satisfactory, but by optimising the 'route' by which the plant is taken from one steady state to the new required steady state so as to minimise the amount of spoilt products.

This, therefore, becomes a start-up and shutdown (as distinct from a continuous control) problem and, over the period of grade change in a papermill or at the nose and tail of each slab in a steel mill, it must be considered a batch process. It is for this reason that I have called these processes quasi-batch processes. Under these conditions, we must revert to a mathematical model of the form of equation (*3*) and we shall find that, although the plant characteristics may not change with time, the $\mathbf{v}$ values are changing and so time must be retained as a dependent variable in the equations. It is also usually found that the $\mathbf{a}$ values are in the form of partial derivatives with respect to the initial conditions $\mathbf{v}$ and this is a further reason for retaining time as a dependent variable. Thus, in order to determine an optimum control strategy, the integrated form of equation (*10*) will have to be used.

Industrial processes

There are very few finished products that are produced from the raw materials by one piece of plant or unit operation. More often than not, a number of intermediate products or components are produced by a number of operations in series and parallel, then assembled together into a finished product as indicated in Fig. 2. Such is the situation in a steel works, the main processes of which are shown in Fig. 3. Even in a papermill, there are a large number of processes in series before the final product is ready for dispatch. Each of these processes can be represented by a mathematical model of the type indicated in equation (*10*), but it is found that the number of variables in each identifiable process is 30 or more. Thus, if all the models are combined to make a complete dynamic model of the whole process, the number of variables may well be several hundred. This means that there is no possibility at the present time of developing an optimum control strategy for a complete factory.

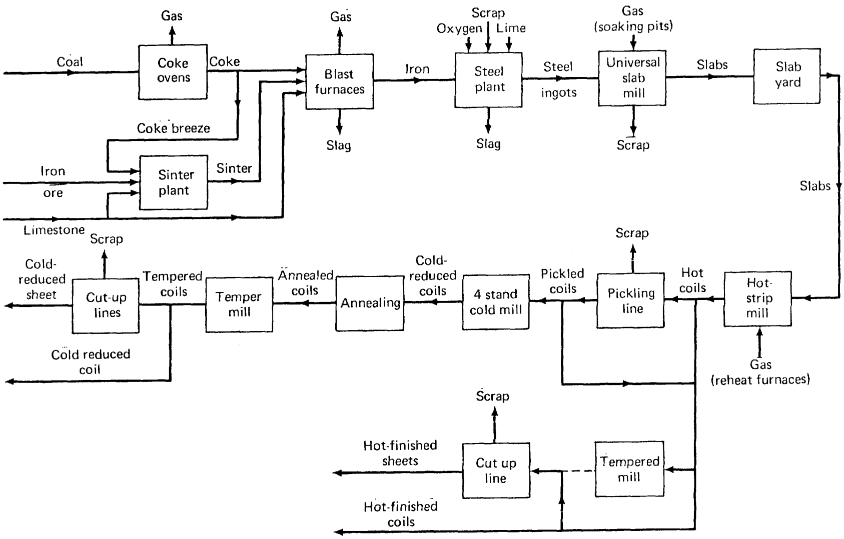

Fig. 3—Main processes of steel works

The present position

Unit operations

There is no doubt that, given time and adequate instruments to make the necessary measurements, also large enough computing facilities, a mathematical model can now be obtained for any industrial manufacturing process or unit operation. Furthermore, given a mathematical model of a process, theoretically, an optimum strategy for the control of the process can be developed by the calculus of variations (such as Pontryagin's maximum principle) for deterministic inputs. That this has been done in only very few cases is accounted for by the difficulties encountered. It has been found that, even when industrial processes are divided up into the smallest possible unit operations, these unit operations have 30 or 40 variables, the $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{x}$ of equation (*1*)—and, as mentioned above, the **x** vector includes intermediate variables that more often than not cannot be measured.

At the present time, most industrial plants are not sufficiently well instrumented for the data to be obtained and, in many cases, sufficiently accurate instruments have not yet been developed. When plants are sufficiently well instrumented, observations must be made over a considerable period and, at least, thousands of pieces of data will be required, so that some form of data recording equipment is desirable. When the data are available, a structural model of the type described above can be tested and, if it proves to be correct, the **a** of equation (*1*) can be calculated, provided a fairly large computer is available. Unfortunately, the structural model is seldom correct in every detail and modifications have to be made so that the development of an adequate model is usually both tedious and costly. I estimate that the cost of getting a model for the simplest process with about 30 variables is at least £10 000. Although the number of different processes being used in industry must run into tens of thousands, I doubt if one hundred of them have yet been modelled satisfactorily.

The development of a dynamic model is only the first step towards optimum or even improved control, because—even for deterministic systems—by use of the calculus of variations, it is possible to solve the two end-point boundary problem for a linear system only with about 10 or at most 20 variables and, if it is non-linear, it is far more difficult. The number of variables must therefore be reduced to those that are significant for control, which it must be hoped will not exceed about 10.

It is for this reason that linearisation of the model becomes important, since equation (*1*) can then be put in the matrix form for equation (*5*) and methods have been developed for reducing the plant matrix A. As mentioned above, continuous processes, because they work over a very large range, can usually be linearised, but batch processes prove to be much more difficult and approximations and changes of variables must usually be made in order to reduce the number of non-linearities. After development of the structural model, this is the next stage in the process. When the model has been linearised and reduced to a reasonable number of variables, it must again be tested, but this can be done of course with the data already collected, provided this covers the full expected range of operation of the plant. The next step is to calculate the optimum control strategy for some appropriate deterministic change in the operation of the plant such as small step changes in the controlled variables or in the composition of the raw materials according to some predetermined criterion. This strategy must then be tested either on the actual plant or on a computer using the original structural model to replace the plant. If a large hybrid computer is available, this can be done fairly quickly; if only a digital computer is available, it will have to be large and the computing time may be considerable.

Such a control strategy should ensure that within the expected range of operation the plant will be controlled in the best possible manner to take account of small variations in the raw materials and some variation in the desired specification of the product can be accommodated, but, theoretically, it will not correct for changes in the plant parameters or ambient conditions. In order to correct for these changes, some form of adaptive control will be required. Methods for doing this have been devised and used successfully in aerospace applications for which well-defined mathematical models of the systems are available, although they have not as yet been applied to industrial processes except as described in the section on DDC. Any adaptive control system almost certainly requires a model reference of some form, which, for practical reasons, should not be too large, so the development and reduction of a dynamic model is almost certainly essential if adaptive control is required.

No account has yet been taken of unwanted fluctuations in the operation of the plant or of measurement noise. Fortunately, in many processes, these random fluctuations are small compared with the accuracy to which the plant must be controlled. If they are not, deterministic control is all that can be achieved by the use of the calculus of variations, which will be certainly sub-optimum and may be far from satisfactory. If sufficient data are available, the effectiveness of the control strategy can of course be tested and sensitivity analysis can be used to predict the effect of unwanted disturbances, but more efficient control can be achieved only by the use of dynamic programming or special filtering techniques, which greatly add to the complication and the amount of computing equipment required if these methods are to be used for on-line control.

The control of complete factories

There are very few factories that produce a single product day in, day out and these are seldom, if ever, made up of few enough unit processes for the development of a dynamic model of the whole factory to be possible. As described above, individual processes, if given sufficient time, can be modelled and near-optimum control strategies can probably be developed. Even if all the processes have been modelled and can be combined, however, the number of variables will almost certainly be hundreds (if not thousands) and there will be no possibility of reducing this to the sort of number, probably fewer than 20, for which a reasonable dynamic control strategy could be developed. If the factory were producing the same product year in, year out (as does an ammonia plant), then it might be possible to develop a satisfactory control strategy to take account of changing raw materials, changing ambient conditions and wear and tear of the plant, provided a fairly large digital computer was available on-line. This is the exception, however, rather than the rule and most factories have to cater for a production programme that includes products of several different grades, colours and physical dimensions. Under these conditions, although the range of product specification may be relatively small, at least some of the processes will be in the quasi-batch category and the dynamic control of the whole factory will be out of the question. Under these conditions, the only possibility for control of the factory as a whole is either to take customers' orders and collate them or to make a future sales forecast on the basis of prediction from previous sales.

The prediction of future sales is just as much a control problem as the modelling of unit operations and there are well-established methods of doing this, but it should be pointed out that this prediction may depend on what the economists call exogenous variables arising from changing economic, climatic or demographic conditions. The prediction of these exogenous variables in turn depend on a knowledge of the behaviour of the economic, climatic and demographic systems in which the factory operates and so, as more sophisticated control of individual factories is required, the control or systems engineer is continuously forced to consider bigger and bigger systems. Assuming that a reasonable forecast of the desired output of the factory can be made, it is possible by operations research techniques (such as linear programming) to determine the best production schedule for the unit processes and this can be updated from day to day or at whatever intervals are required.

The production schedules for the individual processes having been determined, it should now be possible to use for their control dynamic strategies developed in the way described in earlier sections. If all the processes behaved exactly according to plan and the raw materials were always exactly according to specification, there would be nothing more to be done. In practice, of course, the outputs from the unit processes will not be exactly according to schedule in quantity or quality and the discrepancies must be fed back so that the future production schedule can be corrected. Changes in raw materials and faulty operation of early processes result in input changes to the later processes, which need to be optimised as far as possible to correct these discrepancies. Thus, a hierarchical system of control is required with information flowing up the line and orders flowing down. Fig. 4 shows such a system in operation at a steelworks.

Future developments

THERE are seven distinguishable areas in which active research and development is required—

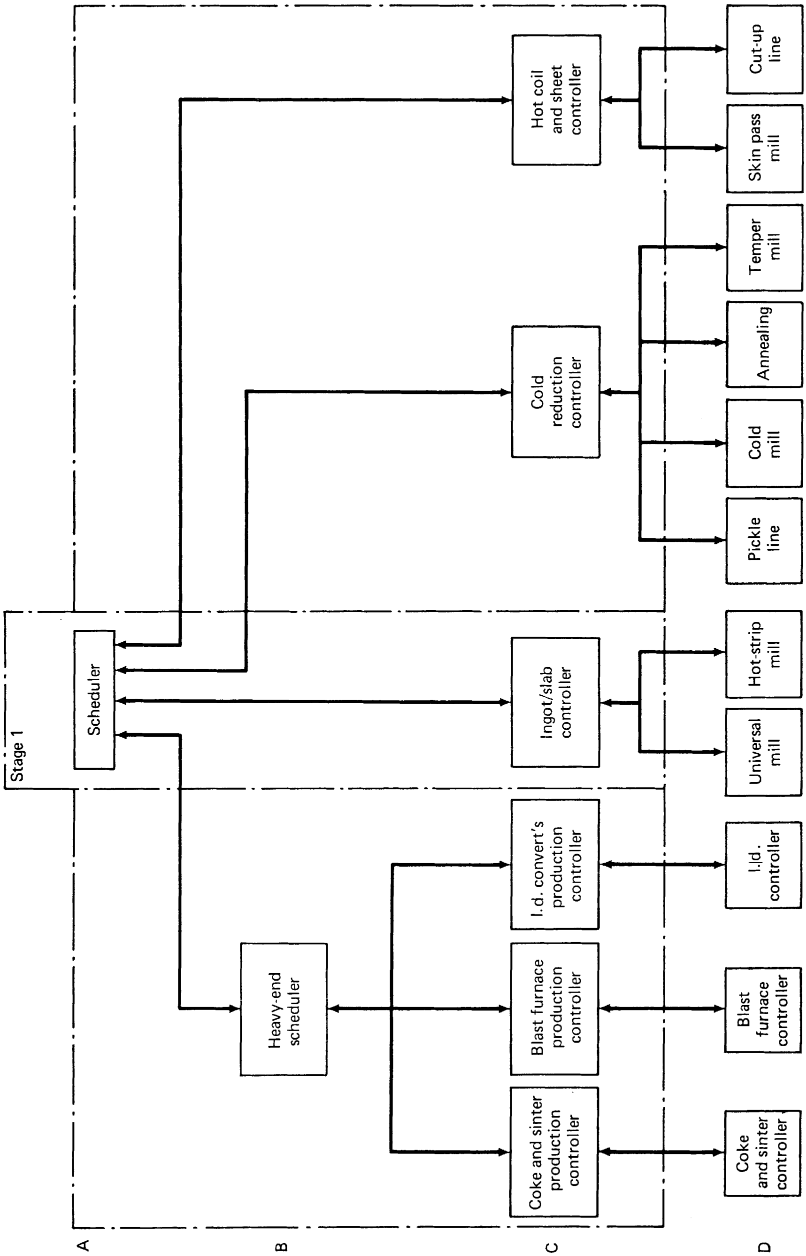

Fig. 4—Interconnections between computer systems—

Level A Plant scheduling B Plant optimisation

C Production control D Process data logging plus process control

1. Improved adaptive control of continuous processes.
2. Improved dynamic control of batch and quasi-batch processes.
3. Computer control of start-up and shutdown of complex systems.
4. The development of hierarchical computer control systems for complete factories.
5. The modelling of very large systems—for example, national economies, meteorological and demographic systems, transport and communications systems and large business and industrial groups.
6. The development of stochastic control systems.
7. Approximation of non-linearities and development of sub-optimum control strategies.

Continuous processes

So long as the range of products required from a single process is restricted and raw materials can be reasonably well controlled, it is probable that direct digital control of steady state is all that will be required. The urgent need is to improve our hillclimbing techniques for parameter estimation in systems that have a very large number of variables or cannot be linearised, so that adaptive control as described in the DDC section becomes possible.

Batch and quasi-batch processes

There can be little doubt that during the next few years the biggest economic gains will come not from improvement in steady state control of continuous processes (which is usually within 1 or 2 per cent of optimum anyway), but by increasing the throughput of batch processes and reducing the amount of spoilt product, when changing grade and at other discontinuities, in quasi-batch processes. As described in the sections on batch and quasi-batch processes, any major improvement almost certainly demands a time-varying dynamic mathematical model of the process and very little has been done as yet in this area.

Five steps are necessary in the development of optimum or near-optimum control of a batch or quasi-batch process—

1. Collection and recording of sufficient data.
2. Development and proving of an accurate structural model.
3. Simplification of this model by finding suitable approximations to allow it to be piecewise linearised and the number of variables reduced.
4. Development of a near-optimum control strategy.
5. Development and test, using the accurate model (*2* above), of the control algorithm.

At the present time, not one of these steps can be carried out for any but relatively simple systems and new and improved techniques will have to be developed.

Start-up and shutdown of complex systems

There are many very big systems such as electrical power generating plants, sugar extraction plants and chemical plants, in which for economic reasons the time taken to get into full production after a shutdown should be as short as possible, but which, for safety or other reasons, has to be carried out in accordance with a carefully worked out schedule—for example, the launch, landing on the moon and return of Apollo 11.

In an electrical generating system, the boilers and generators must be run up in such a way that components are not subjected to thermal shock and no dangerous temperature differentials arise. This may require nearly 10 000 separate commands during the start-up period and can be done in the shortest time only by means of a computer monitoring the system and issuing the commands in accordance with a predetermined program. This is a scheduling problem closely allied to optimum control, which must be considered as coming within the sphere of control or systems engineering.

In the case of a sugar factory, it takes about a week at the start of a campaign to get it into full production and a reduction of a day in this time would save many thousands of pounds. This problem should be treated as a quasi-batch process, but there are far too many variables (more than 200) for this to be possible, so the best compromise is probably to treat it as a problem of scheduling a number of unit processes as described in the next section.

Hierarchical systems

As explained in the section on factory control, a complete factory is far too complex for total dynamic control to be possible. Hierarchical systems have to be used with two or more levels of control. At present, little is known about the behaviour of multi-level systems and stability may be a problem. Relatively simple systems are in operation and research is going on with a view to a better understanding of hierarchical systems, finding principles for designing them and for comparing their performance, but very much more needs to be done.

Very large systems

As mentioned in the section on the control of complete factories, the optimum control of a whole factory inevitably depends on knowledge of a number of exogenous variables, which almost certainly have to be predicted from a knowledge of the environment in which the factory operates. A study of environmental systems such as the national economy therefore becomes necessary. Besides, prediction and control of public utilities such as electric power generation and distribution, gas distribution, transport, communications are essential to the efficient running of a community and, when these have been nationalised or are carried out on a national scale, they become very large systems indeed. The behaviour, design and control of these large systems are now being actively studied under the title of *systems engineering*, but it is a control problem, albeit a complicated and difficult one. All the techniques developed for the control of smaller systems are likely to be required and control will almost certainly be hierarchical.

Stochastic systems

As more accurate control of an industrial process is required, either to improve its efficiency or to make it conform better in a hierarchical system, the effect of measurement noise and unwanted fluctuations becomes more marked. If the system is non-linear, there comes a time when it must be treated as a stochastic system. The analysis of multi-variable linear stochastic systems is difficult enough: when they are non-linear, the mathematical techniques required for their analysis have not yet been developed.

At the present time, the best that can be done to control them efficiently is by the use of dynamic programming. Effectively, this consists of estimating the present state of the system and, by integrating forward along a number of controlled paths to the required end point and by comparison, finding the path that is optimum in some predetermined sense. Since many of the variables are noisy or subjected to random variations, only the most probable (expected) optimum path can be determined, but this is the best that can be done under the circumstances. Such a procedure requires a vast amount of computation and storage of information, so very large computers are necessary; even then, it can be done in reasonable time only for a few variables. Whereas for deterministic systems the time taken to compute the control variables goes up in proportion to the number of state variables, for dynamic programming of stochastic systems, the time required for computation goes up as the square of the number of state variables.

Rapid advances in computer technology are making very much more powerful computers both available and economically justifiable for on-line control of industrial processes; even so, dynamic programming is unlikely to be feasible for most unit processes until new methods have been developed.

Approximat on of non-linearities and development of sub-optimum control strategies

In view of the difficulties to be overcome before truly optimum control can be applied in an on-line adaptive form to any but the simplest processes, probably the most valuable results will come during the next few years from the intelligent use of approximate linear models and sub-optimum control strategies, of which predictive control is an example. When a process has been modelled, it is often possible to find new variables by transformation and combination of the measured variables; these result in approximately linear relationships, which in turn allow the model to be put into the normal form of equation (*6*). The successful linearisation of the structural model often calls for considerable ingenuity and may require different variables to be measured; if the necessary instruments are available or can be developed, it is likely to be worth the effort. Since the first step must be to develop a structural model of the process and the development of new instruments may be required, the most urgent requirement is for the development of structural models of many more unit processes.

Conclusions

DIRECT digital control can now give reasonable adaptive steady state control for continuous processes, but it is in better control of batch processes and of grade changes and unavoidable discontinuities in quasi-batch processes that the greatest gains are likely to be made during the next few years. The most urgent need is for the development of mathematical models for many more manufacturing processes so that dynamic control can be applied to batch and quasi-batch processes and adaptive DDC to continuous processes. In parallel with this, hierarchical control needs to be developed for the better control of complexes of unit processes and complete factories and the study of very large systems should continue. These is little hope of optimum dynamic control of whole factories or of any but the simplest stochastic systems in the foreseeable future, but research into the analysis of stochastic systems is of great importance.

List of symbols

$\mathbf{a}$	Plant and process characteristics: the coefficients, indices, time constants, etc. appearing in the true mathematical model
$\hat{\mathbf{a}}$	Estimated values of $\mathbf{a}$
$\boldsymbol{\alpha}$	The coefficients, time constants, indices, etc. in the true structural model of the system
$\hat{\boldsymbol{\alpha}}$	The estimated values of $\boldsymbol{\alpha}$
$\hat{A}$	The estimated plant matrix when the model is in the normal form of a set of ordinary linear first-order differential equations
$\hat{B}$	The measurement matrix
C	The control matrix
E	The error matrix
F	The function representing the physical structure of the process as a set of first-order differential equations

f	The function representing the physical structure of the process in terms of measurable quantities
g_1, g_2, g_3	The kernel functions of the performance criterion
G	The control function
J	The performance criterion that is to be minimised
υ	Unmeasurable process fluctuations
ξ	Combined noise $\xi - \mathbf{v} + \eta$
η	Measurement noise
t	Elapsed time
T	The time of completion of a batch process
$\mathbf{u}$	The vector representing the controlled variables—that is, movement of tools, flow of steam to heat exchangers, fuel, chemical additives, etc. that are used to control the process
$\mathbf{v}$	The vector representing the composition, quantity, etc. of the raw materials that are not continuously controllable nor used to control the process
$\mathbf{w}$	Fixed costs unaffected by the mode of operation of the plant, such as labour costs and capital charges, which have to be included in the cost function
$\mathbf{x}$	The state variable of the plant and process—that is, the vector representing the state of the plant and the materials in it at time t
$\dot{\mathbf{x}}$	The first derivative of $\mathbf{x}$ with respect to time
$\mathbf{y}$	The measured values of the state variable $\mathbf{x}$
$\dot{\mathbf{y}}$ or y^1	The first derivative of $\mathbf{y}$ with respect to time
$\mathbf{y}^n$	The nth derivative of $\mathbf{y}$ with respect to time, d^ny/dt^n
$\mathbf{z}$	The vector representing the desired output of the plant—that is, the specification of the product plus such other variables as the quantity required