Process Control, Past and Future

Howard Rosenbrock

July 1995

To be presented at the Nordic Process Control Workshop VI,
Mariehamn, Åland, Finland. Aug 21-22, 1995.

Introduction

You will notice that my title refers to process control, past and future, but does not mention the present. This is because my active work in the area of control ceased about twenty years ago, so that my acquaintance with modern practice is that of an observer rather than a participant. Yet this lack of direct involvement does not arise from a loss of interest. I have a deep concern for the way in which control, and technology more generally, will affect the future, and it is this which I will try to make clear.

Retrospective

To begin, however, I may perhaps in my 75th year be permitted to give some reminiscences of earlier days, which will serve to show how my present concerns arose. After the middle 1940s I worked in a number of areas, but always had a background interest in control. The subject then was very much in its infancy, though in a state of rapid development. There had been extensive, and secret, work during the war on control of guns and radar, and the wartime developments had been published[1] in 1947. A seminal conference was held at Cranfield in 1951, and there were a few specialists in the subject at British Universities, though the most active interest was in the USA.

In the period from 1951 to 1954 I worked at John Brown & Co. as assistant to the Engineer Consultant on a number of problems, including arc length control in welding, vibration problems in a water turbine, and vibration and stability of a wind turbine, and I also contributed a chapter describing feedback techniques in measurement for a revised edition of E.W.Golding's book on electrical measurements. From 1954 to 1962 I worked for Constructors John Brown (CJB) who designed and built chemical plants, oil refineries, pipelines, etc., and had an instrumentation subsidiary, and it was there that my interest in process control arose. It was also in that period that digital computers were beginning to become available. At CJB we were using the Leo computer for engineering calculations when it was the only commercially available computer in London, and subsequently we used the English Electric Deuce, IBM 650, Stantec Zebra, Ferranti Pegasus, and Manchester University's Mark One Star as these were introduced.

Out of this experience came some work which has been actively carried forward by others: on optimisation[2], on stiff differential equations[3], and on the digital simulation[4] and control[5] of distillation. Some other work fell upon more stony ground, but it seems to me that it might be worth revisiting with the greater understanding that subsequent developments have given us. My attitude to the analysis of chemical engineering systems was coloured by my training in electrical engineering, where energy considerations allow very general and powerful results to be obtained. Some of my work was therefore aimed at finding similarly general results, if they exist, in chemical engineering.

For example, many unit operations in chemical engineering are, in terms of irreversible thermodynamics and with a reservation which is not immediately apparent, open systems. Then if Onsager's relations are generalised[6] to a steady state far from equilibrium, it is easy to show that with constant boundary conditions the steady state of a stable system is not approached with the usual exponentially damped oscillations. Instead, there is only a finite number of zero crossings before the steady state is reached: the system is either unstable, or very stable. Yet it is possible to devise an open system including evaporation and fluid flow which does give damped sinusoidal oscillations. A little thought shows that some of the 'forces' (as for example gravity) are not obtained as partial derivatives of the entropy, so that distillation and many other unit operations do not fall completely within the scope of irreversible thermodynamics. Nevertheless, it would be interesting to know whether some results from that theory can be obtained for the dynamic behaviour of these chemical engineering operations.

Failing energy and entropy it is possible, for certain systems where a quantity (material, heat, etc.) is conserved, to define a function[7] which can play an analogous role. In a two-component distillation column, for example, we can define a 'disturbance function', equal to the sum of the moduli of the rates of change of composition, taken over all plates and the condenser and reboiler. Then subject to some reasonable assumptions, this function always decreases when the operating conditions are steady. Hence we can obtain a Lyapunov function and confirm stability. Moreover, in steady operating conditions the 'disturbance function' can decrease only through two causes:

(a) the compositions on two adjacent plates are changing in opposite directions

(b) the composition of one or more product streams is changing.

In case (b) there is a numerical relation between the change in the 'disturbance function' and the integral over time of the rates of change of the products.

The analogy here with energy is clear, and some general conclusions[8] can be drawn. If no product is withdrawn, and the operating conditions are steady, the column is either in a steady state, or there is at least one point where two adjacent compositions are changing in opposite directions. Again, if product is being withdrawn, and all compositions are changing in the same direction, then with steady operating conditions all compositions will continue to change in this same direction. Further results of this kind can be obtained, and one would expect that some underlying meaning could be given to the 'disturbance function', but none has been found.

One other analogy with electrical engineering can be suggested. For electrical multiports there is a well-developed theory which derives fundamental results from two defining properties:

(a) the elements (R,L,C, etc.) from which the network is constructed

(b) two matrices defining the way in which the elements are connected.

Some tentative efforts[9] were made to do something similar for more general systems, but they were not pursued. The vision of a similar theory for chemical engineering systems may be a mirage, but it is one that has strong appeal.

What has changed?

Looking back, it can seem that in earlier times there were easier problems to solve, and that, these having been solved, only more difficult problems remain. What truth there is in this thought stems, I think, from something more than just inherent difficulty. In earlier times, the engineer wrote his own computer programs, tested them and ran them. It was a slow and tedious process, but it was transparent. The source of difficulties, whether in the problem formulation or numerical methods or programming, or perhaps in the interactions between all three, was clearly evident.

Contrast this with a recent experience in preparing a mathematical paper using MS Word for Windows. The publisher could not read the equations produced by Word, and he therefore asked for a floppy disc with 'Text Only'. When the proofs were examined it was found that a paragraph which had been deleted from the paper 18 months before had reappeared. Experiment showed that it appeared in 'Text Only' and in 'MS-DOS Text' versions, but not in 'Text Only with Line Breaks'. The result was unexpected and unpredictable, and the user is distanced by the software from the ultimate result.

One consequence of this is to fragment the area of research so that, for example, difficulties such as those in solving the differential equations for a distillation column are no longer likely, as they were at an earlier time, to provoke the chemical engineer into work on numerical analysis. He may attempt to interest a numerical analyst in his difficulty, but may well fail, because numerical analysis has its own criteria for the selection of problems which are considered worthy of interest and effort. A barrier is erected in this way between new problems and the incentive to solve them.

The following example[10] shows how the distancing effect of software can have serious effects upon safety. In December 1991 an aircraft took off from Arlanda airport, Stockholm. After 30 seconds, at 1000 feet, the right-hand engine began to surge and after 64 seconds both engines were surging. At 77 seconds and 3200 feet, in cloud, both engines failed. The instruments were 'flashing lights like fireworks', the Captain piloting the aircraft was unable to obtain any guidance from them, and was anxiously considering whether he had made some error such as the one at Kegworth, where the crew shut down the good engine in mistake for one that failed. An additional Captain, who was a passenger on board, entered the cockpit and continually encouraged the pilot to ignore the alarms and glide the aircraft using the emergency instruments, in preparation for a landing. At 800 feet the aircraft broke out of cloud, the pilot selected a field, and successfully landed the aircraft. Its fuselage was broken into three parts, but no-one was killed.

Subsequent analysis showed that the first engine failed through ingesting ice. At that point Automatic Thrust Restoration (ATR) came into operation, which was designed to advance the throttles in case of engine failure during a manual noise-abatement procedure. Neither the pilots nor the airline were aware that ATR was incorporated in the control system. Neither the engine manufacturer, nor the software designer, nor the FAA had envisaged its operation in the circumstances of the accident. Its effect was to destroy the second engine.

Here again is a 'software-induced surprise', and a system which was opaque, and even misleading and distracting, to the user. The standard response is to ask for

more reliable software, more extensively tested, and for more thoroughly trained users. I would like to suggest that something much deeper is at fault, affecting the whole relationship between our science and technology on the one hand, and the skills and abilities of people on the other.

The UMIST project

I usually arrive at this point in my argument by a different route, closer to my own experience. In the middle 1970s, the coming revolution in microelectronics, communications, and computing was already clear to those 'in the trade'. Some saw it as a new Industrial Revolution, but it was more fruitful[11] to see it as part of a process which had been continuing, and increasing in speed, for two hundred years. This was the process of applying scientific knowledge to industry and commerce, the process of mechanisation and automation. This in turn rested upon and continued an earlier process of the division of labour.

Looking back from 1975, one saw a development by which human work had been fragmented, and the fragments had been, as far as possible, given to machines: a development which is associated with F.W.Taylor, one of its most vigorous defenders and publicists. Taylor called his system 'Scientific Management', and its character is defined by his statement[12], 'Under our system the workman is told minutely just what he is to do and how he is to do it; and any improvement which he makes upon the orders given to him is fatal to success'. The worker, that is to say, is treated as a programmable machine, and to make this possible the task is reduced to machine-like proportions, as in Henry Ford's production line[13]: 'The man who puts in the bolt does not put on the nut. The man who puts on the nut does not tighten it'.

Looking forward in 1975, therefore, one could see the benefits which new technology might bring, but could also see that it might very well be used to intensify and extend the reduction of human work to machine-like operations. This had been done in the past for manual work. It could be carried further in the future, and done also for intellectual work in design, medical diagnosis, some legal work, and in many more areas. Much that was said about Artificial Intelligence reinforced this disquiet.

Social Scientists had for many years condemned the Tayloristic development of technology, and had developed a 'sociotechnical' approach to the design of technical systems - one which considered simultaneously the technical and the social problems, the machines and the people who would work with them. This approach, however, had been applied mainly to small parts of larger projects, or at best to the choice among a number of existing designs. It had not been extended back into the research and development phase from which new technology arises, and in which much of the freedom of design is eliminated.

A project at UMIST[14] was therefore set up to demonstrate, on a small scale, the application of sociotechnical design principles to the early development stages of a programmable lathe controller. This was a small computing system with graphic display and interactive programming, connected to the lathe, which allowed the operator to develop the program for machining a part, and which subsequently carried out the operations that had been programmed. The aim was to return to the lathe operator responsibilities which had been removed from him at an earlier stage - for defining the way in which machining would be carried out, and for testing and modifying the program as necessary. The steering committee for the project contained a social scientist with wide industrial experience, as well as a computer scientist, a senior engineer with a trade union background, a manager, and a production engineer. The small research team was equally multidisciplinary.

A range of design principles were formulated, for example that the system should not confront the operator with behaviour that was unexpected in the light of the situation as he perceived it, and that the operator should be able to override any suggestions of the computer, and should continue to receive full computer support in doing so. The resulting lathe controller became one element in a much larger ESPRIT project.

Lessons of the project

There were a number of difficulties in this project which, while not unexpected, were more serious than had been envisaged. Engineers found great difficulty in making a joint evaluation of technical and social requirements. The first could be quantified and were familiar, the second could not be quantified on any common scale with the first. Equally, social scientists found difficulty in appreciating the technical advantages or penalties of design decisions.

More fundamentally, the scientific training of engineers leads them to see the world in terms of causal relations. The social scientists, despite their claim to scientific status, admit considerations of human aspirations and purpose which a strict interpretation of science would exclude. The outlook of the engineer, therefore, predisposes towards a rejection of human values in work, and towards a Tayloristic view of workers as equivalent to programmable machines. However much the individual engineer wishes to include human considerations he is bound to use the technical methods available to him, and these always presuppose a world in which purpose does not exist, but is only a cover for underlying causal relations. The social scientist is in the converse situation, where to put forward scientific explanations of human conduct in support of features in the system which he considers desirable will reinforce the causal basis of technology, while considerations based upon human values and aspirations will be outside the technical framework and unable to affect it significantly.

In describing science as entirely causal, I am not suggesting that it sees the world as deterministic. The essence of causality is that whatever can be predicted stems entirely from the past. We can, in principle at least, predict the complete behaviour of deterministic systems. Other systems may have random behaviour, which we cannot predict completely. Usually, however, we can predict some average properties of the system, and in the normal scientific view, these average properties are an outcome of the past. What is excluded in this outlook is purpose, the determination of future behaviour by the requirement that some objective shall be attained in the future.

A purposive science

The kind of technology at which the UMIST project was aimed has come to be termed 'human-centred'. It is, that is to say, aimed not at replacing human skill and abilities but at co-operating with them to make them more productive. A human-centred technology will be one in which machines are always subordinate to people, in the relation of tools which enable them to achieve their purposes. My conclusion from the project is that such a technology is certainly possible, but that there is a barrier which, if it is not removed, will prevent us from attaining it. This barrier is the rejection of purpose from science in favour of causality.

In a world without purpose, we can have no responsibility for our actions. They are determined by our past, by our genes, by our upbringing. There can be no ethics, no morality. To quote Einstein[15], 'For the scientist there is only "being", but no wishing, no valuing, no good, no evil; no goal'. Now of course we cannot live by such a creed, but if our technology is based upon a science of this kind, the only means it will offer us to meet our technical needs are ones which exclude the human considerations which science should respect. It will allow no place for human purpose, replacing it by a programmed response which can be demanded from a worker or built into a machine. In this way we arrive at systems which reduce human work to triviality, which take away from the human being the responsibility for work and control over the way it is done, yet return this responsibility when the technical system fails, without providing the information and support which would allow the responsibility to be accepted.

We seem to have reached an impasse. We cannot reject science, and we do not wish to reject purpose; yet the rejection of purpose is regarded as essential by science. What I suggest is that the rejection of purpose by science is chiefly due to a misunderstanding, of a kind which control theory makes very evident. There has been, since the middle of the last century, a formulation of a great part of classical mechanics in terms of a purpose. Hamilton's principle defines a cost which is the integral of the Lagrangian over a future time interval, the Lagrangian being simply the kinetic energy less the potential energy. Minimising (or more generally making stationary) this integral leads to Newton's equations or the equivalent forms of Lagrange or Hamilton. This formulation can be generalised to relativistic systems, special and general. A large part of elementary quantum mechanics can be obtained[16] by a further generalisation which introduces noise, in a particular way that makes the variables complex. Then instead of Newton's equations we obtain Schrödinger's equation.

The introduction of purpose in these ways meets with very strong objection from scientists. There is a historical basis for this objection arising from the conflict with the Church in the early seventeenth century, which placed purpose firmly in the domain of theology, but more important is a basic misunderstanding. It is suggested[17] that a purpose of this kind 'is perhaps a relic of the times when the universe was believed to be driven by a superhuman being' and that 'physicists have assigned to a

particle the knowledge which enables it to take the most convenient path of all those which would direct it to its destination'. These comments are simply wrong. Choice of a control variable to make a future cost stationary carries no theological implications. And the resulting system is, by design, non-anticipative - it needs no knowledge of the future. The particle obeys the appropriate causal laws, Newtonian, relativistic, quantum-mechanical, etc., and whatever is predictable in its behaviour is determined by its past. But the way in which the past determines the future is precisely that which is needed to fulfil the purpose: to make the future cost stationary.

My suggestion is that in the light of control theory we can describe nature causally, or alternatively we can describe it as fulfilling a purpose. The descriptions are equivalent; one implies the other and both give the same results for all scientific purposes. Yet if we acknowledge that the world can be described by means of a purpose, we need no longer reject human purpose from science. It is, of course, immensely complicated, and we are far from understanding it fully, just as we are far from being able to describe all human behaviour causally. We cannot incorporate human purpose, in any but the simplest form, in machines; and for that reason as well as for the respect that we owe to humanity, machines must always be subordinate to people.

Control in the future

This discussion may seem to have become very abstract, so let me return to the engineering practice of control in the future. The greatest problem we face, I suggest, is to combine the power of computers with the capabilities and responsibility of the human operator. There are many guides to good practice: the system must be transparent, must not mislead the user, must be reliable and very well tested, and so forth.

We have to assume that the control system which contributed to the aircraft crash that I described had been designed with all these objectives in mind. Yet it advanced the throttles when it should not have done, and it concealed this action from the pilot. It gave no assistance to the pilot when both engines had failed, but distracted him and added to the stress upon him by a display of flashing lights. The stark fact is that the control system did not care whether the passengers lived or died. It was the pilot who cared passionately that they should live, and who landed the aircraft with the most primitive assistance from his instruments. Yet 'caring', 'concern', 'purpose', are words which will never have been used in the technical development of the control system.

It seems to me that it would have been sensible for the designers to say that ultimately the safety of the aircraft depends upon the pilot. The control system can help him to avoid errors, but whatever situation arises, the controls must continue to assist the pilot to the greatest possible extent. The system must be 'human-centred'; that is to say, it must subordinate the technological function to the human requirements. The basis of our technology in a causal science makes it very difficult for us to adopt this view.

Process control systems are not the same as aircraft controls, but I make no apology for using the example of the aeroplane. It is dramatic, it makes its point strongly, and it had a happy ending. The responsibility on the designer of control systems for chemical plants, oil refineries, power stations, and similar plants is at least as great as that on the designer of aircraft controls. The lesson from one is equally applicable to the other.

References

1 H.M.James, N.B.Nicholls and R.S.Phillips, Theory of servomechanisms, 1947 (McGraw-Hill).

2 H.H.Rosenbrock, An automatic method for finding the greatest or least value of a function, Computer Journal, 1960, vol.3, pp.175-184.

3 H.H.Rosenbrock, Some general implicit processes for the numerical solution of differential equations, Computer Journal, 1963, vol.5, pp.329-330.

4 H.H.Rosenbrock, An investigation of the transient response of a distillation column, Part I: solution of the equations, Trans.Instn.Chem.Engrs.,1957, vol.35, pp.347-351.

5 H.H.Rosenbrock, The control of distillation columns, Trans.Instn.Chem. Engrs.,1962, vol.40, pp.35-43.

6 H.H.Rosenbrock, The approach to a stable steady state far from equilibrium, Proc.Phys.Soc., 1962, vol.80, pp.962-970.

7 H.H.Rosenbrock, A theorem of 'dynamic conservation' for distillation, Trans.Instn.Chem.Engrs., 1960, vol.38, pp.279-287.

8 H.H.Rosenbrock, A Lyapunov function with applications to some nonlinear physical systems, Automatica, 1962, vol.1, pp.31-53.

9 H.H.Rosenbrock, Hierarchical theory of systems, *in* J.K.Skwirzynski (Ed.), Directions in signal processing in communication and control, 1975, pp.637-645.

10 Condensed from L.Mårtensson, Requirements on work organisation, Doctoral Thesis, Royal Institute of Techno, Stockholm, 1995, pp.159-181.

11 H.H.Rosenbrock (Ed.), New technology: society, employment and skill, Council for Science and Society Report, 1981.

12 F.W.Taylor, On the art of cutting metals, American Soc.Mech.Engrs., 1906, p.55.

13 Henry Ford, in collaboration with Samuel Crowther, My life and work, 1923, p.83 (Heinemann).

14 H.H.Rosenbrock (Ed.), Designing human-centred technology, 1989 (Springer-Verlag).

15 Albert Einstein, Out of my later years, 1950, p.114 (Thames and Hudson).

16 H.H.Rosenbrock, A stochastic variational treatment of quantum mechanics, Proc.Roy.Soc.A, Aug.1995.

17 W.Yourgrau and S.Mandelstam, Variational principles in dynamics and quantum theory (third edition), 1968, pp.173-174 (Pitman).