Proceedings of Cybernetics Society Conference on Recent Topics in Cybernetics, London 20th September 1974

The role of randomness in cybernetic systems

B. R. Gaines
Department of Electrical Engineering Science
University of Essex, Colchester, Essex, U.K.

This paper demonstrates that random phenomena, although most often treated in the context of system malfunction, can play major constructive roles in the philosophy, theory and application of cybernetic systems. A control-theoretic example is given to show that a simple stochastic automaton can solve a regulator problem otherwise requiring a recursive automaton, and insoluble for finite-state automata. An identification-theoretic example is given to show that the assumption of causality in modelling a simple acausal system can lead to models which grow in size, on average, precisely at the rate of acquisition of observations. The significance and applications of these results are discussed and illustrated. Finally, it is suggested that the interpretation of stochastic automata theory in terms of system malfunction is far too restrictive in its stimulation of theoretical developments. A wider view of the role of random phenomena might aid the development of results and techniques which actually deliver what automata theory has always seemed to promise.

1 Introduction

The main role of randomness in system theory is that of subsuming residual phenomena not encompassed by the main model or theory, e.g. 'let us assume a channel noise of…' Systems are analysed into a number of causal components and a number of statistically independent noise sources (i.e. bearing no causal relationship). In application studies the noise sources invariably play the role of defects, interference, unreliability, etc., and it is the task of the system engineer to reduce the effects of such defects to acceptable levels.

This negative view of randomness is consistent with more fundamental concepts of knowledge and science. Randomness shares with such basic phenomena as space and time the property that the more one knows about it the less one's knowledge of it becomes. That is, as one axiomatizes a field, those elements which we make simplistic and primeval become inexplicable. In particular our 'scientific' explanations are historically, and psychologically, causal. Random phenomena are thought of as those at the boundary between causality and acausality where some incomplete order can be extracted from nature. This boundary is regarded classically as that of our 'knowledge', and one that should be continually expanded outwards so that the inexplicable becomes random and, eventually, causal. The psychological force behind this assumed role of randomness can be felt in Einstein's remark that, 'I can't believe that God plays dice'—the probabilistic interpretation of quantum mechanics *had* to be only a transient stage in our state of knowledge.

There are, however, results in the literature which suggest that random phenomena have a more constructive role to play in system theory: notably, Rabin's (1963) demonstration that there are regular events accepted by a two-state stochastic automaton which are only accepted by a deterministic automaton with an arbitrarily large number of states, and the practical application of this result in the stochastic computer (Gaines 1969). In relation to other problems, Hellman and Cover (1971) have shown that for deterministic automata of arbitrary size there exist

hypothesis-testing problems which they cannot solve, but for which a two-state stochastic automaton has an arbitrarily small error. And Gaines (1967, 1969) has shown that there are pattern classification that cannot be learnt by an adaptive-threshold-logic element with discrete, bounded weights, unless the weight changes are probabilistic—an effect demonstrated in a practical pattern recognizer by Clapper (1967).

All these results are characterized by the *constructive* role played by randomness. It appears not as a defect to be overcome but rather as a crucial component necessary to a system's operation. This paper is concerned with further results on the essential and constructive roles of randomness in system theory.

2 A control-theoretic result

Consider an abstract formulation of the problem of regulating a discrete dynamical system to maintain its state within a prescribed region. Represent the system as an automaton, (I, P, S, σ, π), where I is a finite input alphabet, P $\equiv$ (0,1) is a binary set of outputs, S is a set of states, σ: SxI $\rightarrow$ S is the next-state function, and π: S $\rightarrow$ P is the output function. In this formulation π is a performance function and the problem is to regulate the inputs to the automaton to cause its output to become, and remain, 1.

Suppose that there is some distinguished element $\lambda \in$ I (the 'zero' input for the autonomous system), and consider the following sets of states:-

$$W \equiv \{s : \pi(s) = 1\}$$

$$A \equiv \{s : \forall n \geq 0, \sigma(s, \lambda^n) \in W\}$$

$$B \equiv \{s : \exists n : \sigma(s, \lambda^n) \in A\}$$

W is the subset of S in which it is desired that the state should reside; A is a weak attractor within W, and B is its region of attraction (Bhatia and Szego 1967). We assume that both B and S-B are non-empty so that the autonomous system has a region of local asymptotic stability but is not asymptotically stable in the large, and consider the family of control automata whose inputs are from P and whose outputs are in I which induce global stability. For this family to be non-empty it is necessary that B be reachable (Arbib 1965) from S, that is:-

$$\forall s \in S, \exists u \in F_I : \sigma(s,u) \in B$$

where F_I is the free semigroup generated by I.

This problem is one of a class considered independently by Gold (1971) in his paper on 'Universal Goal-seekers', and by Gaines (1971) in a paper on stochastic automata. Both authors demonstrate that no finite-state deterministic automaton can act as a universal regulator for this class of problems. On the contrary it is shown that there is a finite-state dynamical system which is universally insoluble for all such regulators with less than a given number of states. Both authors give a construction for such a system which has the property that any action by the regulator other than that which it actually does would immediately solve the problem. Gaines calls such a system a *frustration automaton* and Gold terms the resulting behaviour *strongly worst*!

Having established that there is no universal finite-state deterministic regulator for this class of problems, the two papers diverge in an interesting fashion. Gold. demonstrates that there is a

univeral primitive-recursive (and hence potentially infinite-state) deterministic regulator, and Gaines demonstrates that there is a universal two-state stochastic regulator. This equivalence between recursive and stochastic solutions, together with the non-existence of finite-state deterministic solutions, demonstrates the savings in memory and. complexity resulting from the introduction of randomness. It suggests a role for random phenomena in both physical and biological systems in that they enable a simple, finite-memory structure to achieve the same control capability as a complex structure requiring unbounded memory.

A simple physical example of the principle involved is that of leveling the surface of a bowl of sand. A controller using a fine probe to push down grains on the surface could probably construct an appropriate trajectory through the state-space of the sand, but its storage requirements would be massive and I would not like to have to construct the control algorithm! Randomly tapping the bowl would be a rather more efficient means of achieving the same end. This example also illustrates a difference in informational requirements since the probe controller would require far more information as to the state of the sand (at least the locations of high spots) than the random controller (only a signal when the sand is level enough).

A psychological example of the principle in the context of instructional systems has been given by Gaines (1974) who uses it establish the existence of a universal trainer for any trainable system. The use of performance feedback only by the trainer without other information as to the structure or state of the trainee is given as an explanation of the effective operation of current instructional systems, including computer-based instruction, which do not build a model of the learner. The explanation of the observed phenomenon that many trainees are trainable is clearly outside these arguments, and requires postulates such as Pask's (1967) characterization of 'man as a system that needs to learn'.

It might be supposed that the stochastic regulator is very much less effective than the recursive one in terms of the time taken to solve the problem. However, a simple example shows that this is not so. Moore's (1956) 'combination lock' problem involves an automaton whose output is zero until the correct sequence of inputs is applied when it becomes one (the lock opens). At any stage an incorrect input causes the automaton to re-initialize. Moore shows that a sequence of length at least 2^{N-1} is necessary to guarantee to open a binary lock with N states. It may also be shown that the mean length of sequence to open the lock with a random generator is 2^N. Conway (1971) emphasizes the complexity of the deterministic solution to this problem and recommends the stochastic one. Theoretical upper bounds on the length of the deterministic solution are ridiculously high, but simulation results suggest that the actual mean length is close to that for the stochastic solution.

Another interesting example is the deadlock problem in resource allocation. Suppose a number of automata are competing for the use of a single resource for a limited period. At each instant each automaton either claims, or does not claim, the resource. If there is only one claimant it is allocated the resource, otherwise all claims are rejected. It is clear that if the automata are identical in structure and. commence in the same state then none of them will be allocated the resource (at each instant either none or all will claim), regardless of the complexity of their structure provided their behaviour is deterministic, An 'operating system' solution to the problem would be to mediate all claims through some central resource allocation system. A simpler solution that retains the independent distributed control of the individual automata is to make them claim probabilistically.

The stochastic regulator cannot be frustrated because it does not show consistent behaviour in those circumstances where it is not achieving its goal and hence cannot remain in a fruitless limit cycle. Ashbys (1960) 'homeostat' is a beautiful example of a mechanism embodying this principle, and. he suggests that limit cycles are the rule rather than the exception in complex coupled systems. Bremermann, Rogson and Salaff (1965) and Clapper (1967) have demonstrated the 'frustration' effect in deterministic, or insufficiently stochastic, learning automata. In particular, the results obtained formalize the requirement for an unconstrained random search mode in designs for universal learning automata such as Andreae's STeLLA (Gaines and Andreae 1966; Andreae and Cashin 1969).

In conclusion one may suggest that is dangerous for the observer of a biological system to assume that stochastic behaviour is 'noise' resulting from defects in the system or his imperfect observations—it may be playing a constructive and vital role. Equally the system engineer should consider that a stochastic element may provide a solution to problems otherwise requiring complex algorithms and associated storage. The pedagogue may take heart from the argument that the classic approach of presenting material in one way and, when it is not assimilated, trying another, etc., is not so 'unscientific' and inefficient as it may appear

3 An identification-theoretic result

The psychological force of the assumption that an observed system is causal, illustrated by Einstein's remark, is very strong, and Michotte's (1963) experiments indicate that causality may be a very low level percept, similar in status to shape and colour. Some observations of human problem-solving made in the course of the STeLLA project made apparent the importance of the assumption of causality in cognitive tasks. One test environment for the learning machine was a game of Nim played against a partially random player of variable optimality. To demonstrate that the task was non-trivial for people also we built a NIM automaton with 4 lamps (representing the number of matches left) and 3 keys (representing taking 1 to 3 matches) and a 'reward' light indicating that the game had been won. In practice human beings found the game virtually impossible to learn in these circumstances, whereas in the normal form NIM is trivial.

Most of the problem in learning seemed to stem from the random nature of the opponent's play. Although this makes the game easier to win (the opponent makes fatal errors) it also makes the behaviour of the automaton indeterminate. A hypothesis of randomness, or indeterminism, was *never* introduced by human players, and the longer they played the more confused they became. To investigate this further we built a simpler automaton with two keys and two lamps such that one lamp came on for a short period whenever a key was depressed. The problem was to depress the key under whatever lam was going to come on. In fact when a key was depressed either lamp came on with equal probability. Again in tests with a wide range of subjects this random element was never recognized and individuals were prepared to pit their wits against the automaton for long periods of time.

Records were kept of verbal introspection as subjects played against the automaton, and it was found that most people generated elaborate models which they felt were 'not quite correct'. One of the most interesting records is that of a behavioural scientist who played against the automaton for some 30 minutes and kept a tally of occasions correct less occasions wrong. After that period he was over 20 ahead on his count and announced he had a good strategy but felt he should explore others and seek a better one. His accumulated count then rapidly declined and

when it went negative he announced he was going back to his 'good strategy'. The count continued to decline, however, and he finally announced that the box contained a system which modelled his strategy and then structured itself in the opposing form to outwit him! The hypothesis of such a complex structure (a 'frustration' automaton in the sense of section 2) on the basis of interaction with a simple two-state stochastic automaton suggested that the assumption of causality was a dangerous one in system identification and that it would be fruitful to investigate it more formally.

Gaines (1975) analyses a formal model of a *modeller* forming a causal model of some other system based on his observations of its behaviour. The *observed behaviour* is some sequence drawn from the union of input and output alphabets, and the length of a sequence of observations is the number of output symbols in it. For any finite sequence of observations there will be an unbounded number of finite-state machines whose input/output behaviour, starting from a specified state, is identical to that observed. Out of these possible models there will be some such that no other machine has a smaller number of states, and an *optimal causal modeller* is defined as one who always chooses such a minimal state model.

Five results are established to describe the dynamic behaviour of such a modeller, in particular, the relationship between the number of states in the model and the length of the observation sequence:-

(1) The number of states in the models formed by an optimal causal modeller of a given sequence of behaviour is a monotonic non-decreasing function of the length of that sequence.

(2) The number of states in the models formed by an optimal causal modeller of a sequence of behaviour generated by a finite-state deterministic machine with M states cannot exceed M.

(3) An optimal causal modeller of an observed behaviour of length N will form a model with not more than N states.

(4) There are observed behaviours of length N such that the model formed by an optimal causal modeller has N states.

(5) The *expected* number of states in the model formed by an optimal causal modeller observing a sequence of behaviour of length N may be at least $N-\log_2 N-1$.

The first two 'results' are just a formalization of what we, and the modeller, expect to happen when it am faced by a finite-state deterministic system—the number of states in the model will increase up to final value which cannot exceed the number in the observed system. The next two results show that whilst the number of states in the model cannot exceed the length of observation it can grow as fast as the number of observations. Hence non-causal systems may give rise to behaviour requiring indefinitely complex models. However, if the system whose behaviour is observed is relatively simple, e.g. finite-state even if acausal, we might expect the modeller to behave in a similar way to when modelling a causal system. Clearly a stochastic source could generate isolated examples of the sequences noted in (4) requiring as many states in the model as the number of observations. However, it seems reasonable to suppose that such 'pathological' sequences might be generated only infrequently by a finite-state stochastic automaton.

Hence a more interesting characteristic of the observer's behaviour would be, not the maximum complexity model he might generate, but rather the expected number of states in the model (the average complexity). This might have one of three possible behaviours as a function of the

length of the observation sequence, N. The expected number of states in the model formed by an optimal causal modeller of the behaviour of a finite-state stochastic automaton might be:-

(a) asymptotic to a finite number, i.e. closely similar to the situation when the behaviour modelled is generated by a finite-state deterministic automaton.

(b) growing without limit but slower than the number of observations itself, e.g. as log N. One might hypothesize that at least the ratio:-

$$R_N = \text{expected number of states/number of observations}$$

would tend to zero with N.

(c) growing without limit at a rate similar to the maximum possible, N. This would imply that nearly all the sequences generated were 'pathological' requiring maximum-size models growing as fast as the number of observations.

The main result, (5), shows that case (c) occurs and that the ratio R_N tends not to zero but to unity. The assumption of causality when modelling acausal systems can lead to ridiculously complex models which do no more than retain al past observations. The two-lamp, two-key, automaton we used with human subjects generates sequences leading to result (5), thus explaining the complexity of the models they described. One is tempted to paraphrase Einstein and say that if one assumes God does not play dice them, when he does, one will obtain an over-complex view of the universe.

This result indicates why some of the most attractive and apparently powerful techniques in automata theory are never seen in their obvious applications. The Nerode construction (Arbib and Zeiger 1969; Goguen 1973) gives a straightforward algorithm for determining the optimal causal model (minimum-state finite automaton) corresponding to any sample of observed behaviour. Philosophically we appear to be justified in assuming that knowledge of behaviour can always be condensed into a canonical structure, and that all system-theoretic results can be applied if expressed in terms of the state-space of such structures. In practice we should be able to resolve the problems of treating general nonlinear systems, and in particular the whole of behavioural psychology, by feeding behavioural data into a canonical modeller. Such an aparoach seems to offer the possibility of a neutral behavioural science in which the structure of reality is derived without preconceptions, assumptions or distortion. In a perfectly causal, closed universe in which all information was available to us this would be so, but alas the slightest acausality (Gaines 1975) destroys completely the validity of our modelling technique.

There *are* occasions when we wish to simulate complexity in a fairly simple system, for example to enable an ELIZA-like system (Weizenbaum 1968) to pass the Turing test. An interactive conversational computer program based on simple rules is rapidly dismissed as mechanical and its rules identified unless some random selection is introduced. This is particularly relevant in tutorial systems where one wishes to establish at least parity between (automatic) teacher and student. Albritton's comment that he is more likely to apply human terms to a robot if part of its program is random is one of the more telling points in the discussion following Putnam's (1964) paper on the role of automata in psychology and philosophy. An example in the opposite direction to these is the problem created by fluctuations in the response time at a terminal to an interactive time-shared computer. The variations in response time with system loading are an acausal phenomenon to the interactive user who may attempt to ascribe them to variations in his own activity, e.g. error inputs, and become very disturbed by them (Gaines and Facey 1975).

In conclusion, whilst the theory of (deterministic) automata is immediately applicable to digital computers, its powerful and elegant results and techniques are severely limited in the normal noisy world. We need equivalent results for non-determinate and stochastic automata, and must not fall into the trap of assuming that we can obtain them by assuming that our observations can be fitted into a causal framework—they can but only at the expense of diluting reality with triviality.

4 Conclusions

It has been argued that random phenomena have a positive and. constructive role to play in the philosophy, theory and application of cybernetic systems. Control-theoretic and information-theoretic results, together with their implications, have been given which illustrate such a role. It is suggested finally that the interpretation of stochastic automata theory in terms of system, or observation, malfunction is far too restrictive in its stimulation of theoretical developments. A wider view of the role of random phenomena might aid the development of results and. techniques which actually deliver what automata theory has always seemed to promise.

5 References

Andreae, J.H., Cashin, P.M. (1969) 'A Learning Machine with Monologue', Int.J.Man-Machine Studies 1, 1-20.

Arbib, M.A. (1965) 'A Common Framework for Automata Theory and Control Theory', J SIAM Contr. Ser.A 3, 206-222.

Arbib, P.A., Zeiger, H.P. (1969) 'On the relevance of Abstract Algebra to Control Theory', Automatica 5, 589-606.

Ashby, W.R. (1960) Design for a Brain, London: Chapman and Hall.

Bhatia, E,P., Szego, G.P. (1967) Dynamical Systems: Stability Theory and Applications, Heidelberg: Springer.

Bremermann, H.J., Rogson, P.M and Salaiff, S. (1965) 'Search by Evolution' in Biophysics and Cybernetic Systems, Washington: Spartan Books.

Clapper, G.L. (1967) 'Machine Looks, Listens, Learns', Electronics October 30, 91-102.

Conway, J.H. (1971) Regular Algebra and Finite Machines, London: Chapman and Hall.

Gaines, B.R., Andreae, J.H. (1966) 'A Learning Machine in the Context of the General Control Problem', Proc. 3rd Int.Congr. IFAC, London.

Gaines, B.R. (1967) 'Techniques of identification with the Stochastic Computer', Proc. IFAC Symp. Identification, Prague.

Gaines, B,R. (1969) 'Stochastic Computing Systems' in Advances in Information Systems Science New York: Plenum Press.

Gaines, B.R. (1971) 'Memory Minimisation in Control with Stochastic Automata', Electron. Lett. 7, 710-711.

Gaines, B.R. (1974) 'Training, Stability and. Control', Instructional Science 3, 131-176.

Gaines, B,R. (1975) 'On a Danger in the Assumption of Causality', to be published.

Gaines, B.R., Facey, P.V. (1975) 'Some Experience in Interactive Systems Development and. Application', to be published.

Goguen, J.A. (1973) 'Realization is Universal'. Math.Syst.Theor. 6, 359-374.

Gold, E.H. (1971) 'Universal Goal Seekers', Inf.Contr. 18, 395-403.

Hellman, M.E., Cover, T.11. (1971) 'On Memory Saved by Randomization', Ann.Math.Stat. 42, 1075-1078.

Michotte, A. (1963) The Perception of Causality, London: Methuen.

Moore, E.F. (1956) 'Gedanken Experiments on Sequeuential Machines', in Automata Studies, Princeton University Press.

Pask, G. (1967) 'Man as a System that Needs to Learn' in Automaton Theory and Learning Systems, London: Academic Press.

Putnam, H. (1964) 'Robots: Machines or Artificially Created life', J.Philos 61, 668-691.

Rabin, M.O. (1963) 'Probabilistic Automata', Inf.Contr. 6, 230-245.

Weizenhaum, J. (1968) 'Contextual Understanding by Computers' in Recognizing Patterns, MA: MIT Press.

ONE DAY INAUGURAL CONFERENCE ON RECENT TOPICS IN CYBERNETICS.

London - 20th September, 1974.

PROGRAMME

9.00 REGISTRATION at the School of Pharmacy, University of London,

9.30 OPENING CEREMONY

The Conference will be opened by:-
Dr. F. Hartley, C.B.E., B.Sc., Ph.D., F.R.I.C.,
Dean of the School of Pharmacy and Deputy Vice Chancellor of the University of London.

Chairman: Professor J.E. Houldin, Ph.D, F.I.E.E., F.Inst.P.,
Head of Electronics Department, Chelsea College,
University of London.

MORNING SESSION

Papers relating to biological and pharmaceutical sciences

Chairman: Professor L. Saunders, Ph.D. D.Sc., F.R.I.C.,
Department of Pharmaceutical Chemistry, The School of Pharmacy, University of London.

9.45 "A Pharmacokinetic Investigation of the Distribution and Elimination of Diethylpropion and its metabolites in Man" D. Mihailova, A. Rosen, B. Testa and A.H. Beckett, Dept. of Pharmacy, Chelsea College, University of London.

10.05 "A Cybernetics Approach to Pharmacokinetics" K. Raymond, M. Sherriff and B. Warburton, Department of Pharmaceutics, The School of Pharmacy, University of London.

10.25 "Cybernetics approach to Bio-availability: Is this necessary?" S.S.Davis, Dept. of Pharmacy, University of Aston, Birmingham.

10.45 COFFEE

Papers relating to social and human sciences.

Chairman: Dr. G.B.F. Niblett, Ph.D., Barrister, Department of Law and Computing, University of Kent at Canterbury

11.00 "The Cybernetics of Design Processes", G.L. Mallin, Royal College of Art, Kensington.

11.20 "Cybernetics and the Legal System", H.A. Fatmi, Lincolns Inn, London.

11.40 "Information Retrieval of Parliamentary Legislation" G.B.F. Niblett, University of Kent and N. Price, U.K. A.E.A., Harwell.

Contd/...2.

12.00 DISCUSSION

12.30 LUNCH

AFTERNOON SESSION

Papers relating to mathematical and physical sciences

Chairman: Professor J.F. Coales, O.B.E., F.R.S., Professor of Control Science, University of Cambridge.

2.00 "Holographic Prediction", A.M. Hill, Post Office Research Department.

2.10 "A Cybernetic Approach to Pattern Recognition", D. Hayes, Mathematics Department, S.E. London Technical College.

2.20 "The Role of Randomness in Cybernetic Systems", B.R. Gaines, University of Essex.

2.30 "The Structure of Knowledge in Cybernetic Machines", G.A.S. Pask, Department of Cybernetics, Brunel University.

3.00 AFTERNOON TEA

Papers relating to engineering sciences and cybernetic art.

Chairman: Professor P.B. Fellgett, Head of Cybernetics Dept., The University of Reading.

3.15 "Hadamard Transform Coding and Filtering of Images in the Real and Finite Number Fields", E. Insam, Department of Electronics, Chelsea College, University of London.

3.35 "The Cybernetic Teaching System CASSET", P.B. Fellgate, Cybernetics Dept., University of Reading.

3.55 "High Order Predictors", F.A. Khalafalla, Dept. of Electronics, Chelsea College, University of London.

4.15 DISCUSSION

Papers relating to Cybernetic applications and robots.

Chairman: Professor M.W.Thring, Sc.D., F.I.E.E., F.I.Mech.E., Head of Mechanical Engineering Dept., Queen Mary College, University of London.

4.45 "Towards an Autonomous Robot" and "Mechanical Reflexes or Common Misconceptions Regarding Machines", M.H.E. Larcombe, Department of Computer Science, University of Warwick.

Contd/...3.

5.05 "Robots. Tëlechitic Machines and Mechanical Limbs", M.W. Thring, Mechanical Engineering Department, Queen Mary College, University of London.

5.25 "Phototaxi", J.D. Watson and P. Duncan, Ilex Sales Ltd.

5.45 CONCLUDING REMARKS

Professor G.A.S. Pask,M.A., Ph.D.,
Cybernetics Department,
Brunel University.

6.00 CONFERENCE CLOSES.

6.15 Annual General Meeting of the Cybernetic Society, London

BW/GP
18/7/74

THE CYBERNETICS SOCIETY, LONDON.

CHAIRMAN: H.A. Fatmi, BE,DIC,Ph.D., Chelsea College, University of London.

VICE CHAIRMAN: B. Warburton, MA,Ph.D., The School of Pharmacy, University of London. WC1N 1AX

SECRETARY: D. Hayes, B.Sc., Chelsea College, University of London. SW6 5PR.

Dr. B.R. Gaines,
Department of Computer Science,
University of Essex,
COLCHESTER, Essex.

1st October 1974

Dear Dr. Gaines,

"Recent Topics in Cybernetics"

On behalf of the Cybernetics Society, I must thank you for your contribution to the recent conference. The early days in the existence of any new society are crucial to its survival and we are grateful to all those who made the inaugural conference of the Cybernetics Society such a success.

It is intended to publish the proceedings of the conference, and those who attended will be informed of the progress to this end.

Yours sincerely,

Brian Warburton.

B. WARBURTON