Witten, I. H. and Gaines, B. R. (1976). The role of randomness in system theory. In Progress in Cybernetics and Systems Research III Trappl, R., Klir, G. J. and Ricciardi, L., eds., pp. 11-18. Hemisphere, Washington.

The role of randomness in system theory

I.H. WITTEN and B.R. GAINES
Department of Electrical Engineering Science, Colchester, Essex, UK

1. Introduction

The main role of the concept of randomness in system theory has always been that of subsuming residual phenomena not encompassed by the main model or theory, eg. 'we will assume that this signal is degraded by a channel noise of...'. In application studies the noise sources invariably play the role of defects, such as interference, unreliability, or imprecise measurement, and a major objective in system engineering is to reduce the effects of these defects to tolerable levels. This negative view of randomness is consistent with more fundamental concepts of knowledge and science where causality is taken as a basic postulate [1, pp.316-324] and randomness is introduced as a last resort, temporary 'explanation' when the assumption of causality fails.

In recent years, however, new foundations have been established for the concept of randomness that relate it very closely to that of *computational complexity* [2-7]. This relationship raises some interesting new system-theoretic questions about random behavior, notably: what form of world model will be obtained if we assume that the world is deterministic in circumstances where its behavior is really randomly generated?; can we use the 'complex' behavior of a random source in circumstances where we need complex behavior and would otherwise have to use substantial memory and/or computing power to generate it?

This chapter presents recent results that answer these two questions. It is shown, on the one hand, that a deterministic modeller of a world that exhibits the slightest degree of random behavior will obtain a model that becomes indefinitely complex with the number of observations. On the other hand, theoretical and practical examples are given of situations in which the complexity of random behavior may be used to solve problems that would otherwise be expensive in storage and computing power. A variety of examples and consequences of both these results are given.

The remainder of this introduction is a brief survey of the negative role of randomness in science followed by an outline of more constructive recent results in system theory.

A. Negative aspects of randomness in science

In his wide-ranging review of the emergence of the concept of probability Hacking [8] suggests and illustrates a peculiar symbiosis between causality and randomness in which the very stress on the former as a basis for scientific explanation itself necessitates the introduction of the latter to account for the gap between observation and theory. He remarks (p.2) that, "Europe began to understand concepts of randomness, probability, chance and expectation precisely at that point in its history when theological views of divine fore-knowledge were being reinforced

by the amazing success of mechanistic models". Our scientific models 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. As Hume[9] put it, "it is commonly allowed by philosophers, that what the vulgar call chance is nothing but a secret and concealed cause".

Recently Suppes [10] has been vehemently critical of this negative view of the role of randomness in science. He argues that, "the old theology has been replaced by a new theology of philosophy and science that is equally fallacious and mistaken in character", and that tenets of the new theology are, for example, "every event has a sufficient determinant cause", and "knowledge must be grounded in certainty". Suppes argues that, "the fundamental laws of natural phenomena are essentially probabilistic rather than deterministic in character". He goes on to draw a historic analogy between a refusal to accept probabilistic laws and Descartes' refusal to accept the concept of action at a distance which influenced Newton's own presentation of his laws of gravitation.

However, the negative view of randomness coupled with a positive one of causality appears to be more than an expression of scientific philosophy and to be rooted deeply in human psychology. For example, Michotte's [11] experiments in which certain moving visual patterns are immediately perceived as physical, cause-effect relationships between objects, illustrate that the assumption of causality may be inbuilt strongly at a comparatively low level of brain functioning. Gaines [12 and see section III] has reported the complex causal models generated by subjects performing extremely simple cognitive tasks which have a random component. These innate human factors rather than a reasoned scientific philosophy, appear to dominate the attitudes of many eminent scientists in their rejection of randomness as a legitimate part of scientific explanation. Charles Darwin [13] states, "I cannot think the world as we see it is the result of chance", and the conflict that he saw remains an issue in biology today [14,15,16]. Freud's [17] psychopathology of forgetting, mistakes and dreams originates from his refusal to accept explanations of these phenomena as chance occurrences and his resultant development of causal models for them. Finally, perhaps the most widely known example of this attitude towards randomness is that of Einstein in the controversy over the status of probability in quantum mechanics. In a letter to Born he remarks, "You believe in God playing dice and I in perfect laws" [18].

B. Positive aspects of randomness in system engineering

Even if the assumption of causality was not deeply embedded in the human nervous system, the attitudes of Darwin, Freud and Einstein would be justified by the very success of that assumption in the development of modern science. [1 p.316]. Hacking's argument above that the concept of randomness is complementary to, and subservient to, the determinism that led to modern physics, "this specific mode of determinism is essential to the formation of concepts of chance and probability" [8 p.3], also empasises the negative status of randomness. Thus it is not surprising that only in recent years have constructive applications of random phenomena begun to be investigated. To ask the question, 'can we make use of random phenomena', is itself an inversion of customary modes of thought. To look for practical advantages in so doing would seem foolish, and to expect there to be situations in which random phenomena are essential to achieve goals would seem akin to madness.

However, this decade has seen the growth of substantial research into, and practical application of, the constructive aspects of random phenomena. A key result for this work was Rabin's [19] 1963 paper on stochastic acceptors in which he demonstrated that there are regular events recognized by a two-state stochastic automaton that are only recognized by a deterministic automaton with an arbitrarily large number of states. This result is the theoretical basis for such engineering structures as the stochastic computer [20,21], and its commercial applications, eg. signal averaging [2], digital-analog conversion [23], and electronic [24] and fluidic [25] instruments. It is also the basis for Hellman and Cover's [25] result 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 arbitrarily small error.

A comparable series of results have been developed for stochastic generators. Gaines [20,27] has shown that there are pattern classifications that cannot be learnt by an adaptive-threshold-logic element (ATLE) with discrete, bounded weights, unless the weight changes are probabilistic. Independently Clapper [28] discovered this phnomenon in a practical pattern

recognizer. Gaines [29] generalized the result to show that there was a class of control problems insoluble for any deterministic automaton with a bounded set of states but universally soluble by a two-state stochastic automaton. Independently Gold [30] showed that this class of problems was soluble by a simple recursive automaton, thus establishing a correspondence in power between finite-state stochastic and infinite-state recursive controllers. It may be argued that this correspondence underlies the major role of chance in genetic development as emphasized by Monod [51] and illustrated in the models described by von Foerster [31] and Block [32].

The constructive role of randomness in many applications of electronic systems has been surveyed recently by Gupta [33]. In the remainder of this chapter I shall concentrate on some system-theoretic results that illustrate both the negative aspects of randomness discussed in the previous section and the positive aspects outlined above.

2. Some identification-theoretic results

The psychological force of the assumption that an observed system is causal is very strong as illustrated in section 1A. Some observations of human problem-solving made in the course of the STeLLA [34] project made apparent the importance of the assumption of deterministic 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 four lamps (representing the number of matches left) and three keys (representing taking one to three 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 its 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 behavior of the automaton indeterminate, and 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 light 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 behavioral scientist who played against the automaton for some thirty minutes and kept a tally of occasions correct less occasions wrong. After that period he was over twenty 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 that 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 (3)) 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 [12] analyzes a formal model of a *modeller* forming a causal model of some other system based on his observations of its behavior. The *observed behavior* 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 behavior, 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 behavior 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 behavior is a montonic 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 behavior generated by a finite-state deter-

ministic machine with M states cannot exceed M;

3. An optimal causal modeller of an observed behavior of length N will form a model with not more than N states;
4. There are observed behaviors 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 behavior of length N may be at least $N-\log_2 N-2$.

The first two 'results' are just a formalization of what we, and the modeller, expect to happen when it is 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 behavior requiring indefinitely complex models. However, if the system whose behavior is observed is relatively simple, eg. 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 behavior would be, not the maximum state model he might generate, but rather the expected number of states in the model (the average complexity). This might have one of three possible behaviors 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 behavior of a finite-state stochastic automaton might be:

a. Asymptotic to a finite number, ie. closely similar to the situation when the behavior modelled is generated by a finite-state deterministic automaton;
b. Growing without limit but slower than the number of observations itself, eg. as log N. One might hypothesize that at least the ratio:

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

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-sized 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 all 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 when he does, one will obtain an over-complex view of the universe—perhaps just what has happened in psychopathology.

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 [35,36] gives a straightforward algorithm for determining the optimal causal model (minimum-state finite automaton) corresponding to any sample of observed behavior. Thus, epistemologically, we appear to be justified in assuming that knowledge of behavior can always be condensed into a canonical structure, and that general system-theoretic results are applicable if expressed in terms of the state-space of such structures. In practice, we should be able to resolve the problems of treating general non-linear systems, and in particular the whole of behavioral psychology, by feeding behavioral data into a canonical modeller. Such an approach seems to offer the possibility of a neutral behavioral 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 [12] destroys the validity of our modelling technique.

The failure of deterministic modelling with random sources has prompted a wealth of research on modelling techniques for discrete stochastic systems; Fu and Booth [37] have recently surveyed this work. However, until recently, there has been no formulation of the problem comparable to that for deterministic systems on which the elegant results of Nerode, Arbib and Goguen are based. What is an 'optimal model' of a stochastic system based on a finite sample of that system's behavior—is it a complex model exactly reproducing the behavior (in

which case we ultimately reach the absurd deterministic models discussed above)—or is it a simple model that approximates the behavior (and then how simple for what level of approximation)?

Gaines [6,7] has recently given a precise formulation of the general identification problem that makes these possible trade-offs quite explicit in that the result is not an 'optimal model' but an *admissible set* of models ordered by *complexity* in one dimension and *poorness-of-fit* in the other. Models in the admissible set are such that no other model has both a less poor fit and a lesser complexity. For deterministic modelling the admissible set has only one member (up to an isomorphism) whereas for stochastic modelling it has a spectrum of members with clearly defined drops in 'poorness-of-fit' as real structure in the source is picked out by more complex models [7]. The formulation is applicable to general non-deterministic sources and, in the stochastic case, establishes a close link between the three formulations of probability theory [38] in terms of: relative frequency; subjective probability; and computational complexity. It demonstrates that the system-theoretic tools required for Suppes [10] 'probabilistic metaphsyics' can be made available.

There *is* a positive application of the results of the section in that there are occasions when we wish to simulate complexity in a fairly simple system, for example to enable an ELIZA-like system [39] 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 Putman's [40] paper on the role of automata in psychology and philosophy. A contrary example is the problem created by fluctuations in the response time at a terminal to an interactive time-shared computer. The variations with system loading are an acausal phenomenom to the interactive user who may attempt to ascribe them to variations in his own activity, eg. error inputs, and become very disturbed by them [41].

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.

3. Some control-theoretic results

The results of section 2 illustrate some of the dangers of randomness if we assume that it is not present. However, as indicated in section 1*B*, these are also constructive aspects of randomness—advantages to be gained by its deliberate introduction in system design. For example, consider the following 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 a finite input alphabet, $P \equiv \{0,1\}$ is a binary set of output, S is a set of states, $\sigma : S \times I \rightarrow S$ is the next-state function, and $\pi : S \rightarrow P$ 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:

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

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

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

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 [42]. 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 from S, that is:-

$$\forall\, s \in S,\ \ u \in I^* : \exists \sigma(s,u) \in B$$

where I^* is the free semigroup generated by I.

This problem is one of a class considered independently by Gold [30] in his paper on 'Universal Goal-seekers', and by Gaines [29] in a paper on stochastic automata. Both authors demonstrate that no finite-state deterministic automaton can act as a un-

iversal 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 behavior *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 universal 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 one would have difficulty in constructing 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 [43] who uses it to 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 trainees 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 [44] 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 [45] '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 [46] 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 [47].

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 behavior is deterministic. In a human situation typically that of two people stepping aside in the same direction to avoid one another, oscillations may occur but the deadlock is ultimately broken because their strategies are *not* identical. They do not have to communicate with one another or give way to a greater power as in the 'operating systems' solutions to such problems, eg. semaphores and mediation of claims through a central resource allocator. A simple solution to the general problem, that retains the independent, non-communicating, distributed control by individual automata is to make them claim resources probabilistically. In terms of Dijkstra's famous 'five philosophers' problem [48], one notes that each will get a fair share of the spaghetti if, on finding they have only one fork, they toss it in the air and, if it falls the right side up, put it back on the table for a little while so that another philosopher (whose toss has led him to keep his fork) may grab it!

The stochastic regulator cannot be frustrated because it does not show consistent behavior in those circumstances where it is not achieving its goal and hence cannot remain in a fruitless limit cycle. Ashby's [49] '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 [50] and Clapper [28] 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 [34].

In conclusion one may suggest that it is dangerous for the observer of a biological system to assume that stochastic behavior 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!

4. Conclusions

This chapter has been a wide-ranging, light-hearted and speculative survey of the role of randomness in science and system theory. It has shown both the destructive aspects of randomness if we reject it in modelling and the constructive aspects that made it an important factor in system design and function. If this paper has given some greater character, personality and life to the concept of randomness, and reduced the drabness of its role as a signal-destroying parasite called 'noise', then it will have achieved its objective. We need far richer results on random and general non-deterministic systems if system theory is to have wide application and automata theory is to deliver what it has always seemed to promise.

Acknowledgements

We are grateful to John Andreae and Judea Pearl for discussions that have contributed to the substance of this chapter.

5. References

1. NAGEL, E. *The Structure of Science*, London: Routledge and Kegan Paul (1961).
2. MARTIN-LOF, P. 'The definition of random sequences', *Information and Control* 9, pp.602-619 (1966).
3. KOLMOGOROV, A. 'Logical basis for information theory and probability theory', *IEEE Trans. Inf. Theor.* Vol.IT-14, pp.662-664 (1968).
4. WILLIS, D. 'Computational complexity and probability constructions', *JACM* 17, pp.241-259 (1970).
5. CHAITIN, G. 'A theory of program size formally identical to information theory', *JACM* 22, pp.329-340 (1975).
6. GAINES, B.R. 'Approximate identification of automata', *Electronics Letters* 11, pp.444-445 (1975).
7. GAINES, B.R. 'Behavior/structure transformations under uncertainty', EES-MMS-AUT-75, Department of Electrical Engineering, University of Essex, Colchester, UK (November 1975).
8. HACKING, I. *The Emergence of Probability*, Cambridge University Press (1975).
9. HUME, D. *A Treatise of Human Nature*, London (1739).
10. SUPPES, P. *Probabilistic Metaphysics*, Filosofiska Studier, Sweden: Uppsala University (1974).
11. MICHOTTE, A. *The perception of Causality*, London: Methuen (1963).
12. GAINES, B.R. 'On the complexity of causal models', IEEE Trans. Systems, Man and Cybernetics, January 1976, Vol.SMC-6.
13. DARWIN, F. *Life and Letters of Charles Darwin*, 1888, Vol.11, p.378, London: Collins (1972).
15. LEWIS, J. (ed), *Between Chance and Necessity*, London: Garnstone Press (1974).
16. AYALA, F.J., and DOBZHANSKY, T. (eds), *Studies in the Philosophy of Biology*, London: MacMillan (1974).
17. FREUD, S. *Psychopathology of Everyday Life*, London: Ernest Benn (1914).
18. SCHILPP, P. (ed), *Albert Einstein, Philosopher-Scientist*, Illinois: Evanston, p.176 (1949).
19. RABIN, M.O. 'Probabilistic automata', *Information and Control* 6, pp.230-245 (1963).
20. GAINES, B.R. 'Stochastic computing systems', *Advances in Information Systems Science* (ed. J.T. Tou), Vol.2, pp.37-172 (1969).
21. POPPELBAUM, W.J. 'Statistical processors', Department of Computer Science, University of Illinois, Urbana, Illinois, USA, (May 1974).

22. SCHUMANN, W.R. 'Method and apparatus for averaging a series of transients', United States Patent 3182181, May 5, 1965.
23. HIRSCH, J.J., CHEVALIER, G., and OLIVA, I. 'New types of digital to analog and analog to digital convertors using MOS/LSI technology', Proc. EPN Seminar, Paris (April 1972).
24. TUMFART, S. 'New instruments use probabilistic principles', *Electronics* 48, pp.86-91 (July 1975).
25. MASSEN, R. 'Stochastic fluidic computing systems', Proc. 5th Cranfield Fuidics Conference, 1972, pp.G3 45-56.
26. HELLMAN, M.E., and COVER, T.M. 'On memory saved by randomization', *Annals of Mathematical Statistics*, 1971, 42, pp.1075-1078.
27. GAINES, B.R. 'Techniques of identification with the stochastic computer', Proc. IFAC Symposium on Identification, Prague, June 1967.
28. CLAPPER, G.L. 'Machine looks, listens, learns', *Electronics*, 30 October 1967, pp.91-102.
29. GAINES, B.R. 'Memory minimization in control with stochastic automata', *Electronics Letters*, 1971, 7, pp.710-711.
30. GOLD, E.M. 'Universal goal seekers', *Information and Control*, 1971, 18, pp.395-403.
31. FOERSTER, H. von,'On self-organizing systems and their environments', *Self Organizing Systems* (eds. M.C. Yovits and S. Cameron) London: Pergamon Press, pp.31-50 (1960).
32. BLOCK, H.D. 'Simulation of statistically composite systems', *Prospects for Simulation and Simulators of Dynamic Systems'*, (eds. G. Shapiro and M. Rogers), New York: Spartan Books, pp.23-68 (1967).
33. GUPTA, M.S. 'Applications of electrical noise', *Proc. IEEE,* 1975, 63, pp.996-1010.
34. GAINES, B.R., and ANDREAE, J.H. 'A learning machine in the context of the general control problem'. Proc. 3rd International Congress IFAC, London 1966.
35. ARBIB, M.A. and ZEIGER, H.P. 'On the relevance of abstract algebra to control theory' *Automatica*, 1969, 5, pp.589-606.
36. GOGUEN, J.A. 'Realization is universal'. *Mathematical System Theory*, 1973, 6, pp.359-374.
37. FU, K.S. and BOOTH, T.L., 'Grammatical inference: introduction and survey—part II' *IEEE Transactions on Systems, Man and Cybernetics*, 1975, Vol.SMC-5, pp.409-423.
38. FINE, T.L. *Theories of Probability*, New York: Academic Press (1973).
39. WEIZENBAUM, J. 'ELIZA—a computer program for the study of natural language communication between man and machine.' *Comm. ACM, 1966,* 9 , pp.36-45.
40. PUTNAM, M. 'Robots: machines or artificially created life?' *Journal of Philosophy*, 1964 61, pp.668-691.
41. GAINES, B.R. and FACEY, P.V. 'Some experience in interactive system development and application' *Proc. IEEE* 1975, 63, pp.894-911.
42. BHATIA, N.P. and SZEGO, G.P. *Dynamical systems: stability theory and applications.* Heidelberg: Springer-Verlag (1967).
43. GAINES, B.R. 'Training, stability and control'. *Instructional Science*, 1974, 3, pp.151-176.
44. PASK, G. 'Man as a system that needs to learn' *Automaton Theory and Learning Systems* (ed. D.J. Stewart), London: Academic Press, pp.137-208 (1967).
45. MOORE, E.F. 'Gedanken experiments on sequential machines'. *Automata Studies* (eds. C.E. Shannon and J. McCarthy), Princeton University Press, pp.129-153 (1956).
46. CONWAY, J.H. *Regular Algebra and Finite Machines,'* London: Chapman and Hall, (1971).
47. WITTEN, I.H. 'Learning to control sequential and non-sequential environments', EES-MMS-CON-75, Department of Electrical Engineering, University of Essex, Colchester, UK, 1975.
48. DIJKSTRA, E.W. 'Hierarchical ordering of sequential processes', *Acta Informatica*, 1971, 1, pp.115-138.
49. ASHBY,W.R. *Design for a Brain*, London: Chapman and Hall, (1960).
50. BREMERMANN, H.J. ROGSON, M., and SALAFF, S. 'Search by Evolution', *Biophysics and Cybernetic Systems* (eds. M. Maxfield, A. Callahan, and L.J. Fogel), Washington: Spartan Books, pp.157-167 (1965).
51. MONOD, J., *Chance and Necessity*, Collins, London (1972).