Origins of Stochastic Computing

Brian R. Gaines

Abstract In the early 1960s research groups at the University of Illinois, USA, and Standard Telecommunication Laboratories (STL), UK, each independently conceived of a constructive use of random noise to implement analog computers in which the probability of a pulse in a digital pulse stream represented a continuous variable. The USA group initially termed this a *noise computer* but shortly adopted the UK terminology of *stochastic computer*. The target application of the USA group was visual pattern recognition, and that of the UK group was learning machines, and both developed trial hardware implementations. However, as they investigated applications they both came to recognize that the technology of their era did not support stochastic computing systems that could compete with available computational technologies, and they moved on to develop other computing architectures, some of which derived from the stochastic computing concepts. Both groups published expositions of stochastic computing which provided a comprehensive account of the technology, the architecture of its functional modules, its potential applications and its then current limitations. These have become highly cited in recent years as new technologies and issues have made stochastic computing a competitive technology for a number of significant applications. This chapter provides a historical a historical analysis of the motivations of the pioneers and how they arrived at the notion of stochastic computing.

Introduction

The possibility of emulating analog computers using digital hardware by representing a continuous number as the probability of the presence of a digital pulse in a train of pulses was conceived independently by Sergio Ribeiro and Brian Gaines in the

B. R. Gaines (✉)
University of Victoria, Victoria, BC, Canada

University of Calgary, Calgary, AB, Canada
e-mail: gaines@uvic.ca; gaines@ucalgary.ca

W. J. Gross and V. C. Gaudet (eds.), *Stochastic Computing: Techniques and Applications*, https://doi.org/10.1007/978-3-030-03730-7_2

early 1960s. Ribeiro was a graduate student of Ted Poppelbaum in the Information Engineering Laboratory (IEL) at the University of Illinois, Champaign, Illinois, USA, and Gaines was a graduate student of Richard Gregory in the Department of Experimental Psychology, Cambridge University, UK and also a consultant to John Andreae's learning machines group at Standard Telecommunications Laboratory (STL), UK.

The US and UK groups both implemented digitally-based analog computers using probabilistic pulse trains, the IEL group initially terming this a *noise computer* but shortly adopting the terminology of the STL group, *stochastic computer*, which became the common designation in later research. As both groups evaluated applications of stochastic computing, for IEL primarily image processing and for STL navigational aids and radar tracking, it became apparent that the stochastic computer based on the digital circuitry then available was not competitive with alternative techniques. They began to develop other computer architectures to address those applications such as *burst* and *bundle* processing [58], and *phase computers* [37] and *microprogrammed* computers [21], respectively.

Both groups published extensively on stochastic computing in the late 1960s [24, 26, 30, 61, 68] which stimulated research in other research groups world-wide and many of those publications continue to be widely cited in the current renaissance of stochastic computing as they provide tutorial material on the fundamentals and the commonly adopted terminology for stochastic computer components, representations and applications. They also contain critical commentaries on the strengths and weaknesses of stochastic computing which are still applicable today.

Ted Poppelbaum

When I was asked to contribute a foreword to this collection of articles on the current state of the art in stochastic computing and its applications, my first reaction was sorrow that Ted Poppelbaum was no longer available to co-author it with me. Ted died in 1993 at the age of 68 and did not live to see the massive resurgence of stochastic computing research in the past decade.

Wolfgang (Ted) Johan Poppelbaum was born in Germany in 1924 and studied Physics and Mathematics at the University of Lausanne from 1944 to 1953. In 1954 he joined the Solid State Research Group under Bardeen at the University of Illinois and researched an electrolytic analog of a junction transistor. In 1955 he joined the faculty of the Department of Computer Science and became a member of the Digital Computer Laboratory developing the circuits for the ILLIAC II and later computers. In 1960 he received a patent for his invention of the transistor *flip-flop* storage module [59]. In 1972 he became Director of the Information Engineering Laboratory and received a major Navy contract to support his research on statistical computers and their applications. He retired in 1989.

Ted had many and varied projects in his laboratory. His 1973 report [57] on the achievements and plans of the Information Engineering Laboratory summarizes some 45 distinct projects during the post-Illiac II phase from 1964 to 1973. They are grouped under the categories: Storage/Hybrid Techniques; Stochastic and Bundle Processing; Displays and Electro-Optics; Communication/Coding; World Models and Pattern Recognition; Electronic Prostheses.

Fig. 1 Paramatrix: online digital/analog processing of picture information, Information Engineering Laboratory, University of Illinois, 1965

Ted and I became aware of our common interests in stochastic computing in 1967 as we both commenced publishing about the research and he invited me to present a paper on stochastic computing [18] at the IEEE Convention in March 1968 in New York where he was organizing a session on *New Ideas in Information Processing*. I also visited his laboratory, saw the many systems he had developed including the *Paramatrix* image processor (Fig. 1) which was one of his target applications for stochastic computing, and met John Esch who had built the *RASCEL* stochastic computing system.

Ted and I found we had a common background in solid state electronics and computer innovation, and discussed them at length as if we had been colleagues for many years. I met with him again and introduced him to John Andreae and David Hill at the IFIP conference in August 1968 in Edinburgh (Fig. 2). We kept in touch intermittently and planned a joint book on stochastic computing but I had moved on to other projects and introduced him to Phil Mars at Robert Gordon Institute in Aberdeen who was actively pursuing stochastic computing research. They co-published *Stochastic and Deterministic Averaging Processors* in 1981 [47].

Fig. 2 Three pioneers of computational intelligence: from left to right, John Andreae (learning machines), David Hill (speech recognition), Ted Poppelbaum (stochastic computing in image processing), IFIP Congress, August 1968, Edinburgh

Ted published several additional major articles that placed stochastic computing in the context of other computing technologies, notably his surveys in *Advances in Computers* in 1969 on *what next in computer technology?* [60], in 1976 on *statistical processors* [58] and in 1987 on *unary processing* [62]. His 1972 textbook on *Computer Hardware Theory* [56] that was widely used in engineering courses includes a chapter on *analog, hybrid and stochastic circuits*.

Sergio Ribeiro, Cushin Afuso and John Esch

Sergio Telles Ribeiro was born in Brazil in 933, received an Engineering degree there in 1957 and taught electronics at the Institute of Technology and Aeronautics. In 1960 he received a fellowship from the Brazilian Government to study in the USA and entered the University of Illinois, receiving his masters in 1961 and his doctorate in 1963. His doctoral topic was a *phase plane theory of transistor bistable* *circuits* [67] reflecting Ted's continuing interest in the circuit he had invented and its dynamic behavior that determined the speed and reliability of its operation.

After his doctorate Ribeiro continued as a research assistant working with Ujhelyi on the electronic deflection [65] and intensity modulation [79] of laser beams, and in 1964 they joined Carson Laboratories to pursue the industrial applications of that research. In July 1966 he submitted a paper to the IEEE Transaction on Computers on *random pulse machines* [68] that has become one of the classics in the stochastic computing literature.

It appears that Ribeiro's research on study of the architecture and potential of random pulse machines was theoretical. He notes in footnote 2 that "In the spring of 1963 while working with Dr. W.J. Poppelbaum at the University of Illinois the author suggested that a research program be undertaken to investigate theoretical and practical aspects of random-pulse systems." He thanks Dr. Carson for his support of the writing of the paper without implying that it is a project at Carson Laboratories.

Ribeiro had left Ted's laboratory before I visited and I never met him and have not been able to trace any publications by him after a Carson Laboratories 1966 patent for a display device based on his research with Ujhelyi [80]. There is no specific information about how Ribeiro came to be interested in random pulse computing. However, there is some strong circumstantial evidence that indicates how the notion may have occurred to him.

In 1964 Ribeiro [66] published a correspondence item in the IEEE Computer Transactions critiquing Schmid's [71] 1963 paper on a providing analog-type computation with digital elements. He corrects some errors in Schmid's discussion, suggests improvements in his implementation and then, whilst discussing the utility of pulse rate computers, suggests that studies of artificial neurons show that the implementation could be simple. Ribeiro cites three papers on electronic neuron models [7, 48, 50] from the *Bionics* session at 1961 National Electronics held in Chicago, about an hour away from Champaign, suggesting he may have attended that meeting, and a 1963 paper [44] from the IEEE Transactions of Biomedical Electronics suggesting he continued to follow the related literature.

However, none of the cited papers mention the notion that neurons had stochastic behavior which was common in the neurological literature going back to at least to Lashley in 1942 [42, p. 311]. In 1962 Jenik [40, 41] showed that the rates of the non-coherent pulse trains of two neurons afferent to a third were multiplied in its efferent train. Ribeiro might have become aware of such analyses or he might have considered the optoelectronic approximate multiplier described in one of the neuron model papers [7] and realized that if the pulse streams were independent random events then the output of an AND-gate would be the product of their generating probabilities.

In his 1967 paper Ribeiro mentions neurons in his abstract and index terms, commences the introduction with a presentation of Von Neumann's book on *The Computer and the Brain* [81], discusses the neural analogy extensively throughout, and has a *Bionics* subsection in his references with 12 citations. However, he does not specifically attribute the source of his introduction of the notion of random pulses into Schmid's architecture to any specific material that he cites.

In 1964 Ted initiated a research program to study the computational potential of random-pulse systems by making it the topic of Afuso's doctoral research in 1964 and that of Esch in 1967. Cushin Afuso was born in 1933 in Japan, studied for his masters at the University of Illinois in 1959–1960, and returned for his doctorate in 1964–1968. He states that his 1968 dissertation, *Analog computation with random pulse sequences* [1] is "is a feasibility study of a stochastic computing system" taking the operations of an analog computer as his target and showing how multipliers, dividers, adders and subtractors may be implemented.

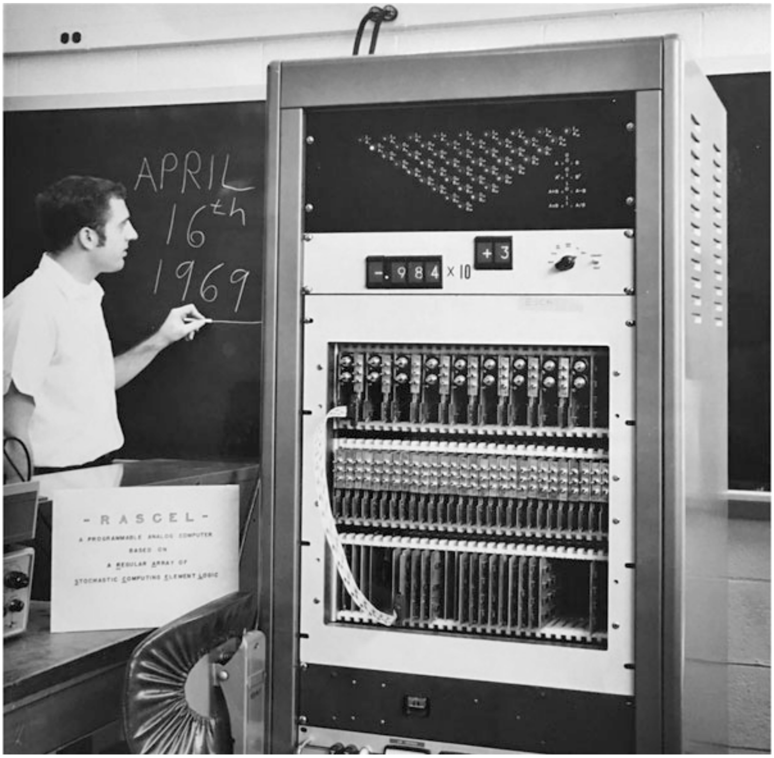

Fig. 3 John Esch presenting his RASCEL stochastic computer, Information Engineering Laboratory, University of Illinois, 1969

John W. Esch was born in the USA in 1942, studied for his masters at the University of Illinois in 1965–1967, and for his doctorate in 1967–1969. He states in his 1969 dissertation, *RASCEL - A programmable analog computer based on a regular array of stochastic computing element logic* [11] (Fig. 3) that "in February of 1967 this author joined Afuso and worked with him to extend the SRPS system to a sign-magnitude number representation and to develop a programmable arithmetic computing element." He went on to a career in the IT industry becoming a major contributor to knowledge representation technologies and literature and developing *conceptual graphs* innovations, implementations and applications [12, 13].

Neither Afuso nor Esch mention the neural analogy in their theses and their research was focused on developing a stochastic computer that emulated the functionality of an analog computer. It seems that the Schmid's version of the digital differential analyzer and research on artificial neurons jointly provided the original inspiration for Ribeiro's ideas but the later implementation became an engineering project with an initial focus on circuits to generate random pulse trains and a later one on how to emulate an analog computer.

Yiu Kwan Wo undertook graduate research with Ted commencing in 1967 receiving his masters in 1970 and his doctorate in 1973 for a thesis entitled *APE machine: A novel stochastic computer based on a set of autonomous processing elements* [85]. However, his work does not extend the stochastic computing aspect of Afuso and Esch's research but instead focuses on radio-frequency transmission of data between the modules supporting inter-operation and reconfiguration without physical interconnection, an intriguing possibility in its own right.

Brian Gaines

It should be easier to describe my own research and the influences on it, and I do have some detailed recollections, but, after five decades, much has been forgotten and I have had to go back to files of notes, documentation, reports papers, memoranda and correspondence from the 1960s that I have dragged around the world for over 50 years but not previously opened—there were many surprises.

I was born in the UK in 1940, studied at Trinity College, Cambridge, from 1959 to 1967 for my bachelors in mathematics and theoretical physics and doctorate in psychology. Electronics became my primary hobby when I was 12 after my father banned analytical chemistry when hydrogen sulphide permeated the house. The availability of low-cost government surplus electronics components and systems after the war made it feasible to create a professional laboratory at home and I built my first oscilloscope from the components of a rocket test set when I was 14.

My school library had several of Norbert Wiener's books and I became fascinated by his notion of cybernetics as the common science of people and machines and his portrayal of what is was to be a mathematician. The headmaster taught a small group of students philosophy in the library and I audited his lectures becoming very interested in Kant and the issues of human understanding of the world and of the nature of scientific knowledge. I found Ashby's writings on cybernetics and admired the way that he solved very general problems using algebraic techniques and I also found Carnap's logical structure of the world and Wittgenstein's tractatus provided formal approaches to the issues that Wiener and Kant had raised.

I was on the science side at school and obtained a state scholarship in mathematics to attend University in 1958 and applied to Trinity College, Cambridge but they made it a condition of acceptance that I also qualify in Latin and delay entry until 1959. I went to the Latin teacher's home for an hour on Saturdays for 3 months to prepare for the examination, and spent the interim year working as a laboratory assistant at ITT's semiconductor research laboratory[1] working on phosphorus and boron diffusion, epitaxial growth of silicon and the fabrication of gallium arsenide tunnel diodes. I also designed and built a nanoamp measuring instrument to determine the greatly reduced leakage current in transistors as we experimented with the planar process and was surprised to find it still in routine use at the end of a 74n integrated circuit family production line when I visited STC at Footscray again some 5 years later.

When I went up to Cambridge I planned to continue my activities in electronics and took with me many samples of the transistor and tunnel diodes that I had fabricated. At the *Societies Fair* in my first term I asked the chairman of the Wireless Society, Steve Salter, whether he knew anyone who might be interested in using them as I hope to find a home in some electronics laboratory. Steve was Richard Gregory's instrumentation engineer and introduced me to Richard who agreed that I could act as Steve's electronics assistant. Richard's primary research was how the brain reconstructed depth information from the disparate images of the separated eyes. I built an oscilloscope with two cathode ray tubes and prisms that allowed the eyes to be stimulated separately. This enabled me to display the 3D projection of a rotatable 4D cube and I studied how the projection was perceived as the subject manipulated it.

In 1961 saw an advertisement in *Nature* for staff for a new *learning machines* project at STL,[2] ITT Europe's primary research laboratories, about an hour away from Cambridge. I applied to John Andreae, the Project Leader, to be employed there in vacations and became his part-time assistant in mathematics, electronics and computing. In particular, I worked on the interpretation of neural net simulations and on the theoretical foundations of the STeLLA[3] learning robot [5] which John was simulating on the KDF9 and his electronics assistant, Peter Joyce, had built in the laboratory (Fig. 4).

Richard and John's laboratories were my focus of attention during my Cambridge years. Trinity only required me to attend a 1 h tutorial with a college Fellow once a week and work on questions from past examination papers, and eventually take the part II mathematics tripos exam to qualify for a degree. Lectures were offered by the university and open to all students but not compulsory or assessed. I largely went to philosophy topics that interested me and lectures by renowned visitors such as Murray Gell-Mann and Richard Feynman in cutting-edge research areas where it was fascinating to meet those who were creating new theories of the nature of matter.

[1] Standard Telecommunications Company (STC), Footscray, Kent.

[2] Standard Telecommunication Laboratories (STL), Harlow, Essex.

[3] Standard Telecommunication Laboratories Learning Automaton (STeLLA).

In June 1962 I took the part II tripos examination in mathematics, and asked the state scholarship committee if I could have a further year of funding to study for the mathematics tripos part III. However, Richard was a consultant to the Ministry of Defence and had been offered funding for a graduate student to undertake a study of the *adaptive training* of perceptual-motor skills. He offered me the opportunity but Oliver Zangwill, the head of the department, said I needed a psychology degree to do so. My scholarship sponsors approved this variation, and Richard asked Alan Watson, the eminent behavioral psychologist, to be my tutor. My positivist leanings suited him well and he set me to write essays for him on a very wide range of topics in psychology, debating my extremely behavioristic mathematical interpretations.

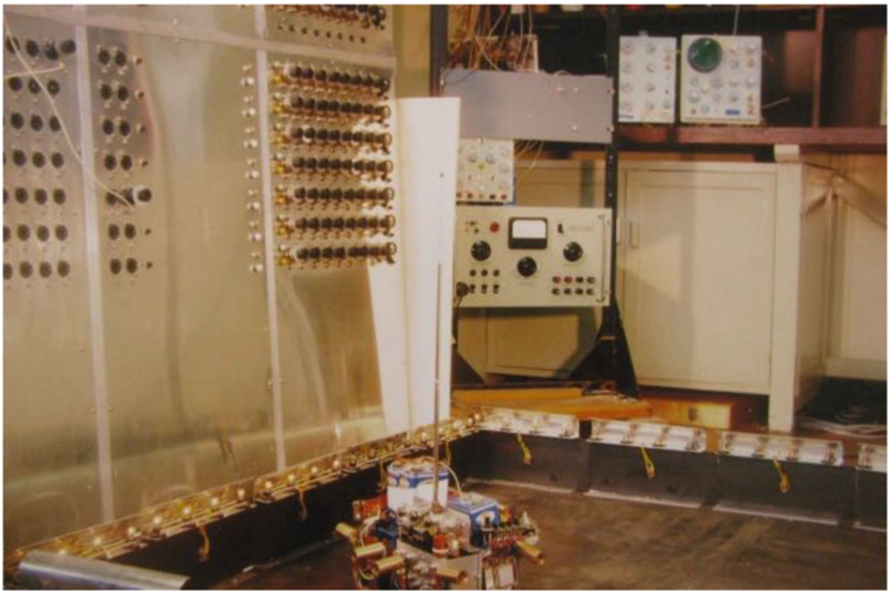

Fig. 4 Andreae's STeLLA, Standard Telecommunications Laboratories, Harlow, Essex, UK, 1963: in the foreground, robot with its sensors in its environment; in the background, racks of electronic equipment and potentiometers for adjusting weights implementing the learning algorithms

In June 1963 I took the part II tripos examination in psychology and became Richard's graduate student funded through the contract. *Adaptive training* is a technique to generate a learning progression for a skill by automatically adjusting the task difficulty based on the trainee's performance, thus turning a simulator into a teaching machine [29]. Common sense suggests it should be effective but nearly all studies to date had negative outcomes. I analyzed the situation Ashby-style assuming that a skill was constituted as a number of dependent sub-skills ordered such that the probability of learning one was low if prior ones had not been learned and showed that in such a situation adaptive training should be very effective even if one has no knowledge of the sub-skill structure or trainee's state in terms of it. I examined previous laboratory studies and felt that the tasks investigated had been insufficiently challenging. The task of interest to the sponsor was classified but the training literature suggested that a tandem-propeller submarine involving managing position by controlling the rate of change of acceleration was extremely difficult and I decided to simulate that.

Fig. 5 Brian Gaines working with the analog computer and stereoscopic oscilloscope that he built, Department of Experimental Psychology, Cambridge, 1964

I built a 10-amplifier analog computer using commercially available operational amplifiers (Fig. 5) but needed an analog multiplier to adjust the parameters that varied the difficulty of the task which was not within the available budget. I designed what I termed a *chopping multiplier* where one signal was intermittently switched by a comparator whose inputs were the other signal and a uniformly distributed waveform. The latter could have been random but the most convenient to generate was a sawtooth. The output was smoothed by a resistor-capacitor feedback loop to generate an estimated product of the two signals. This reduced the bandwidth but it was still more than adequate, and in practice I found it useful to leave some ripple as a disturbance to create a challenging control task. The electronic multiplier tended to drift and for the final study I designed a motorized potentiometer multiplier where the bandwidth for one signal was high but for the other was low but more than adequate for the performance-adaptive adjustment. The experience of designing multipliers sensitized me to the issues of providing them as analog computer modules.

The notion of stochastic computing came from three distinct influences. From Wiener and Ashby I came to see random processes as a source of variety and innovation that became structured through the various filtering processes discussed by Kant, Carnap and Wittgenstein. From my experiences in constructing analog computer multipliers the simplicity of the multiplication of probabilities of the conjunction of uncorrelated events seemed to have engineering potential. From Richard I developed an interest in the neurological basis of depth perception and proposed that the representation of visual intensity by neuronal discharges could be used to extract depth information by spatial correlation through a simple conjunctive processes if the discharges were essentially asynchronous and hence uncorrelated.

In addition, my studies of adaptive training had three components: a theoretical one to show that a very general framework for what is was for any system to *learn a skill* showed that adaptive training accelerated learning; an empirical one of training humans; and an empirical one of training learning machines undertaking the same task as the humans. For the last I used a digital version of Rosenblatt's [69] *perceptron* which did not have the same convergence properties as an analog version. I had noticed this when analyzing Novikoff's [53] proof of perceptron convergence as one of steepest descent. I had previously deconstructed Pontryagin's [55] *maximum principle*[4] to understand why there was no discrete version even though there were several erroneous attempts in the literature to derive one. It seemed to me that a discrete perceptron would have similar problems because it could only approximate steepest descent and might enter a non-convergent limit cycle. I hypothesized that random variation in the training sequence might overcome this as might random variation in the weight changes and showed empirically that the limit cycles did prevent convergence and theoretically that randomness in the training sequence or weight changes could overcome this [30]. However, even though I envisioned a discrete perceptron with random weight changes I did not at that time extend the notion to more general applications. I also did not implement at that time a stochastic perceptron but I found the issues of training one that had problems of convergence very useful to my analysis of the dynamics of training both people and learning machines [17].

All these notions came together when I visited STL in May 1965 and found that Peter Joyce had designed a digital module that John Andreae termed an *ADDIE*[5] that enabled the reinforcing weight adjustments for the STeLLA learning robot to be made automatically rather than by manually adjusting potentiometers. The weight update equation was in the form of a running average, $w' = \alpha w + (1 - \alpha)x$, and Peter had approximated this with a complex of integrated circuits. I noted that the component count could be greatly reduced and the approximation improved if the variables were represented as the generating probability of a digital pulse train, and sketched out circuit diagrams for ADDIE's with various resolutions using 74n series

[4] I became interested in Pontryagin's work because one of my experiments in Richard's laboratory was to replicate the results in a memorandum by Bartlett where he had investigated reaction times in a tapping task with variations in target difference. His results were consistent with the hypothesis that people made Pontryagin-type *bang-bang* movements using maximum acceleration following by maximum deceleration, and I was later able to demonstrate this in my control task[20].

[5] Adaptive Digital Data Integrating Element (ADDIE).

integrated circuit up-down counters, flip-flop arrays with logic gates to generate a pseudo-random number sequence and adders acting as binary number comparators.

John approved further investigation and during the next week I developed the statistical theory for the behaviour of an ADDIE, Peter breadboarded the circuit, and we were able to confirm that theory and practice conformed and provided a module providing the functionality required in the STeLLA architecture. I realized the ADDIE emulated an operational amplifier with a negative feedback loop in my analog computer, that a logic gate acted in the same way as my *chopping multiplier* and that the [0, 1] range of probabilities could be used to emulate $[-1, +1]$, $[0, \infty]$ and $[-\infty, +\infty]$ ranges through appropriate transformations.

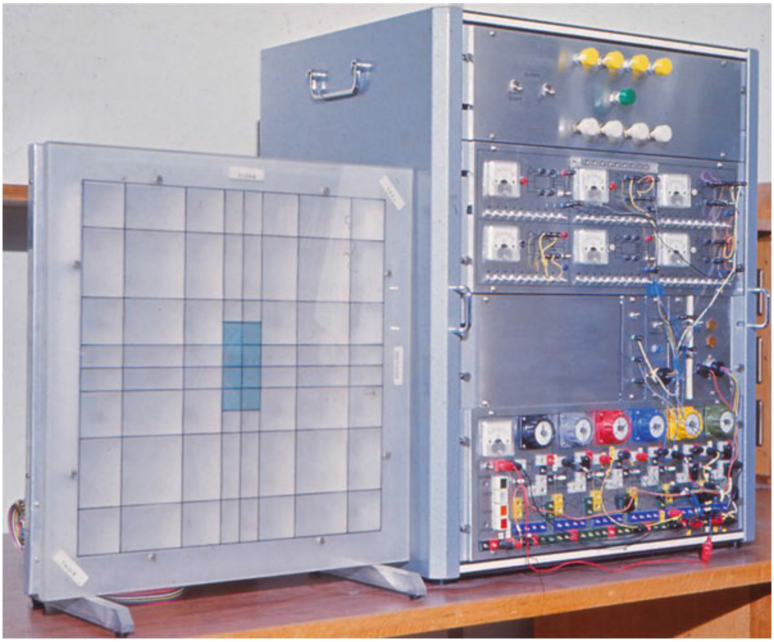

Fig. 6 On right, stochastic analog computer designed by Brian Gaines and built by Peter Joyce; on left, visual display of STeLLA learning controller's trajectory to the specified region of state space, Standard Telecommunication Laboratories, 1965

This led me to propose that the ADDIE be generalized to a complete *stochastic analog computer* with objective of providing all the computational support required by the STeLLA architecture. When it came to a later presentation to other ITT researchers John suggested that the term *analog* could be confusing as the computer was a digital one emulating analog functionality and we abbreviated the name to *stochastic computer*. Peter rapidly constructed a six-integrator stochastic computer with both analog and digital input-output capabilities (Fig. 6) and we were able to test its performance in a variety of applications, not only those relevant to learning machines such as the ADDIE, a stochastic perceptron, Bayesian prediction, and so on, but also the simulation of dynamic systems, digital control, solution of partial differential equations, and so on. STL's primary output was ITT patents and I worked with our resident agent who had trained as a mathematician and understood the principles of the computer to file in March 1966 a comprehensive patent that had 54 claims covering the computer, representations and applications [27].

Once the patent was filed ITT approved the publication of details of the stochastic computer. The first public disclosure was at the IFAC Congress in June 1966 where I presented a paper with John on *A Learning Machine in the Context of the General Control Problem* [35] which updated his at the 1963 Congress [5]. The stochastic computer was the focus of my discussion contribution reporting progress since the paper was submitted which was transcribed in the published proceedings [35].

The first full paper I wrote on stochastic computing was requested late in 1966 by Roger Meetham at the National Physical Laboratory (NPL) who had heard a non-disclosure presentation that John and I gave to some NPL researchers and requested an article for the *Encyclopaedia of Linguistics, Information and Control* that he was editing. The paper was written early in 1966 and approved for submission by ITT in April but the book did not appear until 1969 [25].

The IFAC discussion [35], encyclopaedia entry [25], an internal presentation in December 1965 [16] and the patent [27] together provide a good account of how we perceived the stochastic computer at the time of its invention and before we were aware of a similar invention at the University of Illinois.

The magazine, *Electronics*, had published a short news item in December 1966 noting that "at the University of Illinois, a group of computer researchers has designed a series of analog computer circuits that depend on noise and therefore needn't be protected from it" and providing circuit examples [10]. I asked the editor if they would like an article on the similar research at STL and he took my draft, redrew all my diagrams as hand-drawn sketches to make them appear doodles from a research notebook, and retitled it as *Stochastic computer thrives on noise* [24].

I submitted a paper to the analog computing session Spring Joint Computer Conference in Atlantic City [26] as part of my first trip to the USA where I visited IBM, Bell Laboratories and Xerox research laboratories, under a research-liaison agreement between the major electronics companies. The doyens of analog and hybrid computers, Granino Korn and Walter Karplus also presented and, in leading the discussion on my paper, Walter remarked that he had never expected to see further radical innovations in analog computing.

At the conference exhibition I met Gene Clapper from IBM who was exhibiting his research on character recognition and speech recognition based on digital perceptrons [9]. He remarked that he had been surprised to find character recognition less accurate but ascribed it to a lower variety in the training sequences, and we discussed the role of noise in aiding the convergence of perceptrons. I also presented a paper [30] at the IFAC conference on system identification in Prague which focused on the modelling applications of the stochastic computer such as gradient techniques, the digital perceptron and Bayesian predictor.

My December 1965 presentation at STL aroused interest is several other research groups, notably those developing navigational aids and air traffic control systems and the written version was also distributed to other ITT companies in Europe. It aroused great interest, not only as a new computing technique but also because computing research had been discouraged in ITT after early commercial losses but several ITT manufacturing companies had become involved in both computer manufacture and application through government funding in their own countries. There was a need for computer expertise to be developed in the research laboratories.

I visited ITT companies in France, Germany, Sweden and Norway and found some potential applications of stochastic computers but was also able to contribute knowledge of general-purpose computers from my own experience and also from our usage of large-scale integration logic circuits in the 74n TTL family where usage had previously been discouraged because of early experience of unreliability. We were asked to investigate the feasibility of digital versions of navigational aids for the Decca and Omega systems, and for radar automatic-tracking systems, and I was also tasked with developing plans for a computer research division that I might lead when my university research was complete in 1967.

It rapidly became apparent that stochastic computing in itself would be too slow to be competitive in the proposed applications but I realized that the low-cost multiplication with AND-gates only required uncorrelated pulse streams, not necessarily random ones, and developed a digital version of my *chopper multiplier* where one pulse stream was uniformly distributed and the other clumped so that the two were uncorrelated. It was again possible to develop a complete analog computing framework and I termed it a *phase computer* [37] since the processing was now cyclic and the results were accumulated over each phase. Peter Joyce quickly produced a general-purpose prototype structured similarly to an analog computer (Fig. 7) and we were able to test this in a variety of applications such as a Decca navigator, a radar tracker and a digital PID control system.

In June 1966 John Andreae relocated to New Zealand to take up a chair in electrical engineering and I took over as part-time project leader but kept closely in touch with him by correspondence which has provided a useful record of what was happening at STL in late 1966 and early 1967. We were awarded a British government contract from the Ministry of Technology Advanced Computer Project for *Pulse Rate Modular Computing* and Ray Shemer joined the project to develop the phase computer and investigate its applications, also registering as an external doctoral student with the University of London [72, 73].

I was invited to be the keynote speaker at the IFAC conference on *Pulse Rate and Pulse Number Signals in Automatic Control* in Budapest in 1968 and presented a paper on *varieties of computer — their applications and inter-relationships* [34] which analyzed the relations between general-purpose and special-purpose computer architectures and applications. From the other presentations and discussions I also came to understand how what we had developed fit into a much larger sphere of applications of pulse rate computing. There was extensive discussion of the paper largely focusing on the relative merits of general-purpose and special-

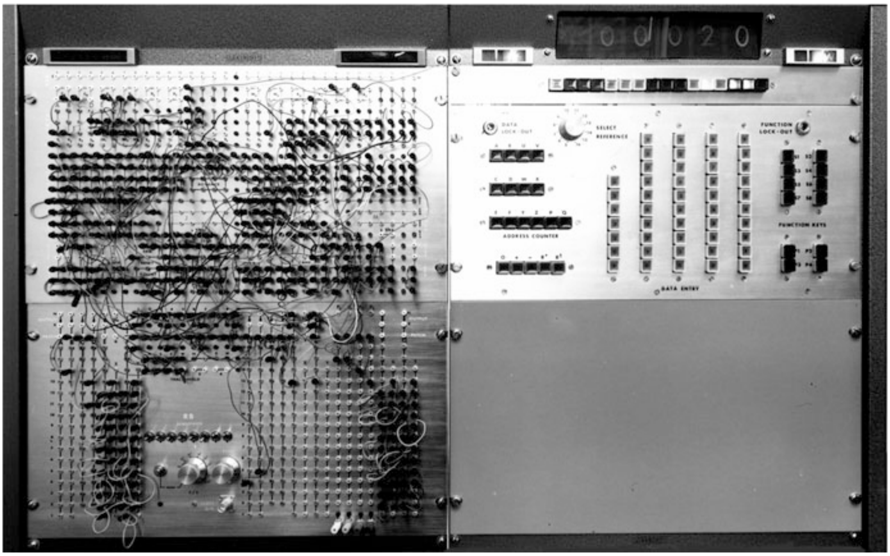

Fig. 7 Phase computer designed by Brian Gaines and built by Peter Joyce, Standard Telecommunication Laboratories, 1966

purpose computers and I still remember the joking, but perceptive, comment of one discussant that buying a general-purpose machine was safer because if it turns out to be unsuitable you can always find another use for it.

I was asked to summarize the conference by the editor of *Automatica* and Ray and I wrote an overview that concluded "Whatever the state of the special-purpose/general-purpose controversy, it is clear that the advent of low-cost integrated circuits has opened up a tremendous range of possibilities for new developments in control hardware; the re-development of DDA-like incremental computing techniques is one of the more attractive possibilities which is likely to lead to practical applications" [38]. I was also asked by the editor of the IEEE Computer Transactions to write a commentary on Ribeiro's 1967 paper and concluded that "the main obstacle to the practical application of the stochastic computer is, at present, the generation of the random variables required in a reliable and economical manner. It may well be that we should look to truly random physical processes, such as photon-photon interactions, to provide the hardware foundation for stochastic computing systems" [23].

In May 1966 ITT decided to tender for the Boeing 747 simulators required in the UK on the basis of the simulation capabilities of LMT, their French company, but needed to establish a British company to manage the tendering process and support the products if they were successful. I was told I was to be appointed chief engineer of the new company rather than head of the new advanced developments and computing division. I had already arranged for that division to have research links with the Electrical Engineering Science department at the newly formed University of Essex, and when I expressed my dismay at no longer having that opportunity to Barrie Chaplin the incoming Head of Department he offered me a position as one of the four initial faculty with a contract that encouraged me to develop industry links. I accepted, remaining a consultant to STL and the phase computer development for some years but becomingly increasingly involved with other activities.

After my 1967 presentation at the Spring Joint Computer Conference, Julius Tou asked me for a survey of stochastic computing for *Advances in Information Technology* and I extended my encyclopaedia article to what was essentially a 136 page *brain dump* of all that I knew about stochastic computer. It was published in 1969 [28] and came to establish much of the stochastic computing terminology in the literature.[6] I intended that to be my final paper on stochastic computing but was asked in 1976 to provide an overview of progress for an IEE colloquium on *parallel digital computing methods* [32]. In 1987 I was asked to introduce the topic at the first IEEE International Conference on *Neural Networks* [33], a paper that attracted a large audience and substantial discussion despite my noting that the research reported was 22 years old and has become cited as a primary source in many papers on random pulse neural networks.

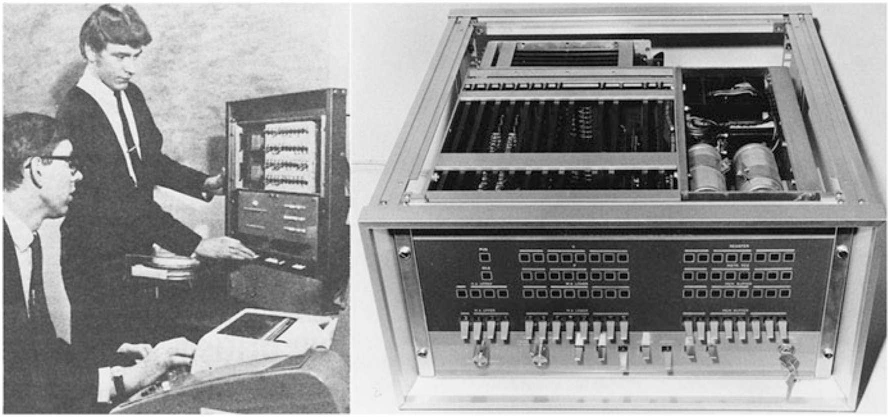

Fig. 8 Minic 8-bit microprogrammed microcomputer designed by Brian Gaines and built by Tony De'ath, Essex University, 1968; on left. the university prototype; on right, the commercial version

At Essex University my research became focused on the design of general-purpose and hybrid computers and associated operating systems, compilers and applications, human factors of computing and artificial intelligence. I received a contract from the RAF to continue my research on adaptive training in flight

[6]Earl Hunt was one of the first to cite this chapter (in the context of von Neumann's book [81]) in his 1971 paper on "what kind of computer is man?" and comes to the conclusion that man is a stochastic computer. Earl unfortunately died in 2016 just before the advent of stochastic deep learning neural networks [45] and the assessment of how the behaviour of deep networks emulated human visual perception [54] that begins to validate his conjecture.

simulators, and decided to add hybrid capabilities to my analog computer to support more flexible training strategies. I had a PDP9 in my laboratory but needed a small dedicated computer, could not find a suitable one and decided to build my own which would give me full control over both hardware and software. I designed an 8-bit microprogrammable computer using the 74n series integrated circuits supporting a main memory from 1 to 64 Kbytes and a drum backing store, and my technician, Tony De'ath, built a prototype (Fig. 8). The experience of designing a conventional computer was very similar to that of designing the stochastic and phase computers, even though the architectures appear so different—all such digital computers are finite-state automata and the design process is akin to writing programs in the same language but satisfying somewhat differing requirements.

When I was evaluating a 64 Kbyte Sperry drum as a permanent storage device for Minic, the sales manager, Dave Seale, and his engineer, Tony Maine, said they would be interested in forming a company to manufacture the computer. I agreed and we formed Microcomputer Systems with Dave as CEO, Tony as chief engineer and myself as part-time technical director. We discussed funding with several potential investors but eventually went with the George Kent, an industrial instrumentation company who stated in their press release that "the Minic's concept is far in advance of anything in the UK or the United States" [78, 82]. That may appear an exaggeration given DEC's PDP7, 8, 9 and 10 architectures, but Minic seems to be the first microcomputer for sale at below £2000 and also the first with customer-programmable microprogramming. It was a precursor of the personal computers yet to come based on the Intel 8080, Zilog Z80, and so on, single-chip microcomputers.

The machine tool company, Alfred Herbert, needed an industrial computer for numerically-controlled machines and became our biggest customer. I had provided for 256 microcode instructions but only used 128 for the computer itself, and we were able to encode the stepping machine geometry in the remaining microcode using the phase computer techniques in a way that minimized the interface needed between computer and machine tool [74]. Alfred Herbert was able to sell its *Batchmatic* machine tool controller for £3500, half the price of their major competitor.

My graduate student, Peter Facey, wrote an emulator for MINIC on our central PDP10 computer [14], and I programmed the MINIC operating system, assembler, loader and a general-purpose system and application language, MINSYS [15, 19, 22], entirely in the emulator so that they were available before the first machines were manufactured. I also programmed some of the initial applications such as an 8-bed intensive care monitoring system for University College Hospital that operated on a 1 Kbyte MINIC with a 64K drum and output charts of blood pressure, temperature, and so on, on a Tektronix storage oscilloscope.

We received a British government contract from the Ministry of Technology Advanced Computer Project for the development of a more powerful version of MINIC, codenamed MINIC-S, and intended for industrial applications requiring high-performance computing and IO, security of operation and high reliability. The original MINIC was used to provide the IO processor, or processors, and the

MINIC-S had a descriptor and capability architecture targeted on the constructs required by high-level language compilers and the hardware assurance that running programs could not unintentionally interfere with one another [36, 39, 84]. A prototype was built and an operating system, assembler, linking loader, FORTRAN compiler and part of an ALGOL-68 compiler were written. However, in 1974 George Kent was acquired by Brown Boveri who decided to use DEC computers rather than manufacturer their own and MINIC manufacture and MINIC-S development were cancelled.

These computers may seem remote from the issues of stochastic computing but for me they illustrate the continuity between special-purpose and general-purpose computers. The MINIC modules were the same as those in the stochastic and phase computers but the microprogram used them to emulate a von Neumann architecture stored-program computer. If we had continued the development of the stochastic computer then its modules would have been similarly controlled to provide a programmable version of the analog computer patchboard. The principles of program control apply to all computer architectures, even neural networks. There is always a need to be able to configure general-purpose modules for special-purpose tasks.

I became aware of an analogous phenomenon in humans through discussions with the Russian psychologist, Alexander Luria, during his visits to the Psychology Department at Cambridge. I had been very impressed by his investigations on the effect of linguistic behaviour on the performance of perceptual-motor skills [46], and investigated verbal instruction as a priming technique for both my human and learning machine subjects [31]. One role of language is to provide program control, or at least behaviour priming, in both human and artificial neural networks.

The Invention of Stochastic Computing

Ted and I took for granted the independent simultaneous invention of stochastic computing at the University of Illinois and STL and never discussed it or tried to ascertain who was 'first.' We became aware of earlier pulse rate computers and of statistical linearization techniques in polarity coincidence correlators [86] and saw noise/stochastic computing as an extension of such techniques.

Multiple discovery and invention [51] is a common and well-studied phenomenon across many disciplines [83] and the usual explanation is that those involved were stimulated by the availability of the same, or similar, information. I have tried to ascertain that common inspiration for Ribeiro's and my research, and have suggested that it is the overlapping neural analogy in Ribeiro's considering artificial neurons as modules of pulse rate computers, and my considering the multiplicative processes implementing correlation in the interaction of the pulse streams of natural neurons.

In addition, the history of stochastic computing also exhibits another phenomenon of multiple discovery/invention where later researchers are unaware of previous work. One of my colleagues at STL, David Hill, found in a patent search in the early 1970s that an invention filed by William G. Thistle in 1962 entitled *Integrating Apparatus* [77] that carried out computations using random pulse trains.

Thistle was an electronics engineer conducting research for the Canadian Ministry of Defence at the Canadian Armament Research and Development Establishment in Québec. David contacted him for further information and received both the patent and an internal report entitled *A novel special purpose computer* [76]. He sent me copies at the time and I recollect reading the patent and noting it was related to stochastic computing but have only now read the report in detail whilst writing this paper.

The abstract of Thistle's report states: "A type of computer is described for the real time evaluation of integrals of the form $I = \int ydx$, where x and y are functions of time. It is believed to be novel in it use of quasi-random processes, operating on pulse trains, to achieve the desired result. The method may be extended to cases where y is a function of several variables dependent on time. Accuracies comparable to analog methods are obtainable without the drift problems usually associated with analog methods."

Thistle describes techniques for addition, subtraction, multiplication, division and integration using random pulse trains, provides circuit diagrams, and described an application to a simple navigation system. His computer encompasses the basic architecture of the stochastic computers developed at the University of Illinois and STL and would constitute *prior art* from a patent perspective.

His report was not widely circulated. The distribution list shows that only 3 copies were issued (to the superintendents of systems and of electronics, and the to chief superintendent) and 25 were lodged in the documents library. Thistle has three other patents (for power supplies and a gas discharge matrix display), and seems to have no other publications although there will likely be other internal reports. It is probable that much of his work was associated with classified systems.

A google scholar search on his name returns two of his patents, one of which is the US version of his *Integrating Apparatus* retitled *Computer for evaluating integrals using a statistical computing process*. His patent is not cited in other patents as prior art, and it seems unlikely that, even today, a content-based automated search would be able to link his text to the stochastic or pulse rate computing literature. As far as I know, Thistle's research is completely unrecognized and has had no influence, and there is no indication of how he came to invent a stochastic computer, but it deserves recognition in the history of computing as the earliest documented invention of a fully-functional stochastic computer.

Thistle's invention is also relevant to another question frequently asked about discoveries and inventions, what would have happened if neither the Illinois or STL teams had developed stochastic computers, would others have done so? The answer is clearly *yes*—it had already happened but no one knew. There was also research in neurology where it became known empirically, possibly as early as the 1950s, that the coincidence of neurons firing could result in a common afferent neuron firing and that this might be the basis of motion detection [64]. This led to an empirical analysis of the jitter in neural firing that was shown to be sufficient for the afferent neuron to be acting as a pulse frequency multiplier [75]. Thus, *stochastic bit-stream neural networks* [8] were conceived from biological studies uninfluenced by stochastic computing (even though the similarity to the stochastic computer is often noted in that literature, e.g. [43]).

Conclusions

In the three decades after Ted and I completed our research in stochastic computing research continued elsewhere but at a low intensity. We received papers to referee, were asked to be thesis examiners, and were aware that there was continuing activity by colleagues across the world, such as Phil Mars in the UK, Sadamu Ohteru in Japan, Robert Massen in Germany (who in 1977 wrote the first book on *stochastic computer technology* [49]) and others, but no major growth in interest. However, in the recent past there has been a significant growth in research as illustrated in Fig. 9 which shows the citations to my 1969 survey (a more robust estimator based on a basket of several commonly cited articles shows a similar pattern). This book provides a much-needed overview of this burgeoning literature through tutorials and overviews by some of those who make have major contributions to its growth.

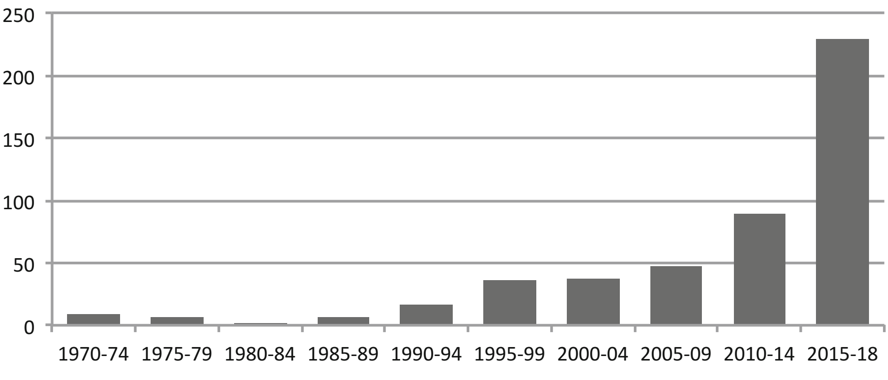

Fig. 9 Counts of the citations of *Stochastic computing systems* [28]

In the conclusions of my 1969 survey I noted three aspects of stochastic computing that seem to me to remain relevant to current research:

- "The stochastic computer has as yet had no major practical impact on data processing systems. Equally, the analogy between nervous processes and stochastic computing has not been carried to a stage where the stochastic computer can be justified solely as a model of the central nervous system. Indeed, present justification for interest in stochastic computing is of a scientific nature—it extends the range of known data-processing systems."

- "If true stochastic processes are utilized, e.g., from light photons or radioactive sources, then it would seem better to establish computing elements working directly with the original sources; photon/photon interactions brought about by certain electron systems would seem to offer a great potential for *natural* stochastic computation."
- "The design of learning machines and pattern recognizers which take full advantages of the properties peculiar to stochastic computers offers the greatest promise for the future development and exploitation of the systems surveyed."

To these, in the light of the ensuing developments in Ted and my laboratories, I would add:

- The essence of the stochastic computer is that, by representing a number as the frequency of occurrence of a binary pulse stream, simple modules can be used to perform significant computations, but, as shown by the phase computer, the pulse streams do *not* necessarily have to be truly random, only uncorrelated.
- The maintenance of this lack of correlation such that the output of a computation may be used as the input to another is a major issue in stochastic computer applications (which makes the apparent simplicity of stochastic computing misleading). Whilst a general solution might be feasible (e.g. using technologies with intrinsic stochastic behaviour), the requirement may also be addressed by application-specific techniques to manage correlation (e.g. [2]) that take advantage of the structure of the computations required.
- Hybrid architectures that combine the modular, parallel processing of the parallel computer and programmability of the general-purpose computer, particularly to control module interconnection will become increasingly significant.
- There are computational problems where randomness plays a significant role in computing the optimal or the least complex solution, and these merit special attention as actually requiring a stochastic implementation.

I will not attempt to present the current state of the art in stochastic computing as it relates to these issues. Very perceptive surveys by those deeply involved in current research are available [3] and the contributions to this book provide in-depth studies of the state of the art. Ted would be interested to see his interest in image processing addressed in the research that in many way triggered the resurgence of interest stochastic computing when it was shown to be as effective, error-tolerant, and requiring less energy usage than competing image processing technologies [4, 63]. My early interests are addressed by the applications of stochastic neural networks to deep learning [45] and other neuromorphic applications [6], by the wide range of alternative technologies being investigated for stochastic computing [3, 6, 70], and by the more computationally efficient deterministic variants which parallel our transition to the deterministic phase computer and are now accepted as variants of stochastic computing [52]. There are also many significant theoretical and application innovations that we never envisioned—interesting ideas take on a life of their own, nurtured by the community of research.

Acknowledgements I am grateful to John Esch for his help in verifying my commentary on the research at the University of Illinois and for providing the photograph of his RASCEL stochastic computer. I am grateful to David Hill for prompting my memory of certain dates and events and for providing the material by Thistle documenting his early stochastic computer. I would also like to thank the editors of this volume for providing the opportunity to contribute this account of the origins of stochastic computing knowing that there are very few of us still alive to do so. I hope it will be of interest to the stochastic computing research community of this era, and wish them well.

Note Early stochastic computing material is available at http://cpsc.ucalgary.ca/~gaines/reports.

References

1. Chusin Afuso. “Analog computation with random pulse sequences”. PhD thesis. University of Illinois, 1968.
2. A. Alaghi and J. P. Hayes. “Exploiting correlation in stochastic circuit design”. *IEEE 31st International Conference on Computer Design (ICCD)*. 2013, pp. 39–46.
3. A. Alaghi, W. Qian, and J. P. Hayes. “The promise and challenge of stochastic computing”. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 37.8 (2018), pp. 1515–1531.
4. Armin Alaghi, Cheng Li, and John P. Hayes. “Stochastic circuits for real-time image-processing applications”. *Proceedings of the 50th Annual Design Automation Conference*. Austin, Texas: ACM, 2013, pp. 1–6.
5. John H. Andreae. “STeLLA: A scheme for a learning machine”. *Proceedings 2nd IFAC Congress: Automation & Remote Control*. Ed. by V. Broida. London: Butterworths, 1963, pp. 497–502.
6. A. Basu et al. “Low-power, adaptive neuromorphic systems: Recent progress and future directions”. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 8.1 (2018), pp. 6–27.
7. T. E. Bray. “An optoelectronic-magnetic neuron component”. *Proceedings National Electronics Conference*. 1961, pp. 302–310.
8. Peter S. Burge et al. “Stochastic bit-stream neural networks”. *Pulsed Neural Networks*. Ed. by Wolfgang Maass and Christopher M. Bishop. Cambridge, MA: MIT Press, 1999, pp. 337–352.
9. Gene L. Clapper. “Machine looks, listens, learns”. *Electronics* (Oct. 30, 1967).
10. Electronics. “Dropping the guard”. *Electronics* (Dec. 12, 1966), p. 48.
11. John W. Esch. “RASCEL - A Programmable Analog Computer Based on a Regular Array of Stochastic Computing Element Logic”. PhD thesis. University of Illinois, 1969.
12. John W. Esch. “System and method for frame and unit-like symbolic access to knowledge represented by conceptual structures”. Pat. 4964063. Unisys Corporation. Sept. 15, 1988.
13. John W. Esch and Robert Levinson. “An implementation model for contexts and negation in conceptual graphs”. Conceptual Structures: Applications, Implementation and Theory. Springer Berlin Heidelberg, 1995.
14. Peter V. Facey and Brian R. Gaines. “Real-time system design under an emulator embedded in a high-level language”. *Proceedings DATAFAIR 73*. London: British Computer Society, 1973, pp. 285–291.
15. Brian R. Gaines. “A mixed-code approach to commercial microcomputer applications”. *Conference on Microprocessors in Automation and Communications*. London: IERE, 1978, pp. 291–301.

16. Brian R. Gaines. *A Stochastic Computer: Some Notes on the Application of Digital Circuits to the Operations of Arithmetic and Differential Calculus by Means of a Probabilistic Representation of Quantities*. Tech. rep. Standard Telecommunication Laboratories, Dec. 9, 1965.
17. Brian R. Gaines. "Adaptive control theory: the structural and behavioural properties of adaptive controllers". *Encyclopaedia of Linguistics, Information & Control*. Ed. by A.R. Meetham and R.A. Hudson. London: Pergamon Press, 1969, pp. 1–9.
18. Brian R. Gaines. "Foundations of stochastic computing systems". *Digest of IEEE International Convention*. New York: IEEE, 1968, p. 33.
19. Brian R. Gaines. "Interpretive kernels for microcomputer software". *Proceedings Symposium Microprocessors at Work*. London: Society of Electronic & Radio Technicians, 1976, 56–69.
20. Brian R. Gaines. "Linear and nonlinear models of the human controller". *International Journal of Man-Machine Studies* 1.4 (1969), pp. 333–360.
21. Brian R. Gaines. *MINIC I Manual*. Colchester, UK: Department of Electrical Engineering Science, 1969.
22. Brian R. Gaines. *MINSYS Manual*. Colchester, Essex: Department Electrical Engineering Science, 1974.
23. Brian R. Gaines. "R68-18 Random Pulse Machines". *IEEE Transactions on Computers* C-17.4 (1968), pp. 410–410.
24. Brian R. Gaines. "Stochastic computer thrives on noise". *Electronics* (July 10, 1967), pp. 72–79.
25. Brian R. Gaines. "Stochastic computers". *Encyclopaedia of Linguistics, Information & Control*. Ed. by A.R. Meetham and R.A. Hudson. London: Pergamon Press, 1969, pp. 66–76.
26. Brian R. Gaines. "Stochastic computing". *Spring Joint Computer Conference*. Vol. 30. Atlantic City: AFIPS, 1967, pp. 149–156.
27. Brian R. Gaines. "Stochastic Computing Arrangement". British pat. 184652. Standard Telephones & Cables Ltd. Mar. 7, 1966.
28. Brian R. Gaines. "Stochastic computing systems". *Advances in Information Systems Science, 2*. Ed. by J. Tou. New York: Plenum Press, 1969, pp. 37–172.
29. Brian R. Gaines. "Teaching machines for perceptual-motor skills". *Aspects of Educational Technology*. Ed. by D. Unwin and J. Leedham. London: Methuen, 1967, pp. 337–358.
30. Brian R. Gaines. "Techniques of identification with the stochastic computer". *Proceedings IFAC Symposium on The Problems of Identification in Automatic Control Systems*. 1967, pp. 1–10.
31. Brian R. Gaines. "Training the human adaptive controller". *Proceedings Institution Electrical Engineers* 115.8 (1968), pp. 1183–1189.
32. Brian R. Gaines. "Trends in stochastic computing". *Colloquium on Parallel Digital Computing Methods—DDA's and Stochastic Computing*. Vol. 30. London: IEE, 1976, pp. 1–2.
33. Brian R. Gaines. "Uncertainty as a foundation of computational power in neural networks". *Proceedings of IEEE First International Conference on Neural Networks. Vol.3*. Ed. by M. Caudhill and C. Butler. 1987, pp. 51–57.
34. Brian R. Gaines. "Varieties of computer —— their applications and inter-relationships". *Proceedings of IFAC Symposium on Pulse Rate and Pulse Number Signals in Automatic Control,* Budapest: IFAC, 1968, pp. 1–16.
35. Brian R. Gaines and John H. Andreae. "A learning machine in the context of the general control problem". *Proceedings of the 3rd Congress of the International Federation for Automatic Control*. London: Butterworths, 1966, 342–348 (discussion, session 14, p.93).
36. Brian R. Gaines, M. Haynes, and D. Hill. "Integration of protection and procedures in a high-level minicomputer". *Proceedings IEE 1974 Computer Systems and Technology Conference*. London: IEE, 1974.
37. Brian R. Gaines and Peter L. Joyce. "Phase computers". *Proceedings of 5th Congress of International Association for Analog Computation*. 1967, pp. 48–57.
38. Brian R. Gaines and Ray A. Shemer. "Fitting Control Mathematics to Control Hardware: An Aspect of the 1968 IFAC Pulse-Symposium". *Automatica* 5 (1969), pp. 37–40.

39. Brian R. Gaines et al. “Design objectives for a descriptor-organized minicomputer”. *European Computing Congress Conference Proceedings, EUROCOMP 74*. London: Online, 1974, pp. 29–45.
40. F. Jenik. “Electronic neuron models as an aid to neurophysiological research”. *Ergebnisse der Biologie* 25 (1962), pp. 206–245.
41. F. Jenik. “Pulse processing by neuron models”. *Neural Theory and Modeling: Proceedings of the 1962 Ojai Symposium.* Ed. by Richard F. Reiss. Stanford, CA: Stanford University Press, 1964, pp. 190–212.
42. K. S. Lashley. “The problem of cerebral organization in vision”. *Visual Mechanisms*. Ed. by H. Kluüver. Oxford: Cattell, 1942.
43. Robert Legenstein and Wolfgang Maass. “Ensembles of spiking neurons with noise support optimal probabilistic inference in a dynamically changing environment”. *PLoS Computational Biology* 10.10 (2014), e1003859 1–27.
44. E. R. Lewis. “The locus concept and its application to neural analogs”. *IEEE Transactions on Bio-medical Electronics* 10.4 (1963), pp. 130–137.
45. Y. Liu et al. “An energy-efficient online-learning stochastic computational deep belief network”. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* (2018), pp. 1–1.
46. A.R. Luria. *The Role of Speech in the Regulation of Normal and Abnormal Behavior*. Oxford: Pergamon Press, 1961.
47. P. Mars and W. J. Poppelbaum. *Stochastic and Deterministic Averaging Processors*. Stevenage: IEE/Peregrinus, 1981.
48. T. B. Martin. “Analog signed processing by neural networks”. *Proceedings National Electronics Conference*. 1961, pp. 317–321.
49. Robert Massen. *Stochastische Rechentechnik: Eine Einfuührung in die Informationsverarbeitung mit zurfaälligen Pulsfolgen*. Munich: Hanser, 1977.
50. E. P. McGrogan. “Improved transistor neural models”. *Proceedings National Electronics Conference*. 1961, pp. 302–310.
51. Robert King Merton. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press, 1973.
52. M. Hassan Najafi, David J. Lilja, and Marc Riedel. “Deterministic methods for stochastic computing using low-discrepancy sequences”. *IEEE/ACM International Conference On Computer-Aided Design (ICCAD ’18)*. New York: ACM, 2018.
53. Albert B. Novikoff. *On Convergence Proofs for Perceptrons*. Tech. rep. SRI Project 3605. Menlo Park, CA: Stanford Research Institute, 1963.
54. Joshua C. Peterson, Joshua T. Abbott, and Thomas L. Griffiths. “Evaluating (and improving) the correspondence between deep neural networks and human representations”. *Cognitive Science* (2018), 42.8, 2648–2699.
55. L. S. Pontryagin et al. *The Mathematical Theory of Optimal Processes*. Oxford, England: Pergamon Press, 1962.
56. W. J. Poppelbaum. *Computer Hardware Theory*. New York, Macmillan, 1972.
57. W. J. Poppelbaum. *Record of Achievements and Plans of the Information Engineering Laboratory*. Champaign, Urbana, IL: Department of Computer Science, University of Illinois, 1973.
58. W. J. Poppelbaum. “Statistical processors”. *Advances in Computers*. Ed. by Rubinoff Morris and C. Yovits Marshall. Vol. Volume 14. Elsevier, 1976, pp. 187–230.
59. W. J. Poppelbaum. “Transistor Flip-Flop Circuit”. Pat. 2933621. University of Illinois Foundation. Aug. 2, 1956.
60. W. J. Poppelbaum. “What next in computer technology?” *Advances in Computers*. Ed. by L. Alt Franz and Rubinoff Morris. Vol. Volume 9. Elsevier, 1969, pp. 1–21.
61. W. J. Poppelbaum, C. Afuso, and J.W. Esch. “Stochastic computing elements and systems”. *Fall Joint Computer Conference*. Vol. 31. New York: Books, Inc, 1967, pp. 635–644.
62. W. J. Poppelbaum et al. “Unary Processing”. *Advances in Computers*. Ed. by C. Yovits Marshall. Vol. Volume 26. Elsevier, 1987, pp. 47–92.

63. W. Qian et al. “An architecture for fault-tolerant computation with stochastic logic”. *IEEE Transactions on Computers* 60.1 (2011), pp. 93–105.
64. Werner Reichardt. “Evaluation of optical motion information by movement detectors”. *Journal of Comparative Physiology A* 161 (1987), pp. 533–547.
65. S. T. Ribeiro and G. K. Ujhelyi. “Electro-Optical Modulation of Radiation Pattern Using Curved Electrodes”. U.S. pat. 3433554. Secretary of the Navy. May 1, 1964.
66. Sergio Telles Ribeiro. “Comments on Pulsed-Data Hybrid Computers”. *IEEE Transactions on Electronic Computers* EC-13.5 (1964), pp. 640–642.
67. Sergio Telles Ribeiro. “Phase Plane Theory of Transistor Bistable Circuits”. PhD thesis. University of Illinois, 1963.
68. Sergio Telles Ribeiro. “Random pulse machines”. *IEEE Trans. Electronic Computers* EC-16.6 (1967), pp. 261–276.
69. Frank Rosenblatt. “The perceptron: A probabilistic model for information storage and organization in the brain”. *Psychological Review* 65.6 (1958), pp. 386–408.
70. Sayed Ahmad Salehi et al. “Computing mathematical functions using DNA via fractional coding”. *Scientific Reports* 8.1 (2018), p. 8312.
71. Hermann Schmid. “An operational hybrid computing system provides analog-type computation with digital elements”. *IEEE Transactions on Electronic Computers* EC-12.6 (1963), pp. 715–732.
72. Ray A. Shemer. “A hybrid-mode modular computing system”. *Proceedings of IFAC Symposium on Pulse Rate and Pulse Number Signals in Automatic Control*. Budapest, 1968.
73. Ray A. Shemer. “A Hybrid-Mode Modular Computing System”. PhD thesis. 1970.
74. I. W. Smith, D. A. Hearn, and P. Williamson. “Software development for Batchmatic computer numerical control system”. *Proceedings of the Fourteenth International Machine Tool Design and Research Conference*. Ed. by F. Koenigsberger and S. A. Tobias. London: Macmillan, 1974, pp. 381–389.
75. Mandyam V. Srinivasan and Gary D. Bernard. “A proposed mechanism for multiplication of neural signals”. *Biological Cybernetics* 21.4 (1976), pp. 227–236.
76. William G. Thistle. *A Novel Special Purpose Computer*. Tech. rep. CADRE Technical Note 1460. Valcartier, Queébec: Canadian Armament Research and Development Establishment, 1962.
77. William G. Thistle. “Integrating Apparatus”. Canadian pat. 721406. Ministry of National Defence. Nov. 30, 1962.
78. Financial Times “George Kent backs a micro-computer venture”. *Financial Times* (Feb. 5, 1970), p. 11.
79. G. K. Ujhelyi and S. T. Ribeiro. “An electro-optical light intensity modulator”. *Proceedings of the IEEE* 52.7 (1964), pp. 845–845.
80. Gabor K. Ujhelyi, Sergio T. Ribeiro, and Andras M. Bardos. “Data Display Device”. U.S. pat. 3508821. Carson Laboratories. Aug. 11, 1966.
81. John Von Neumann. *The Computer and the Brain*. New Haven, Yale University Press, 1958.
82. Computer Weekly. “MINIC system is bought by George Kent”. *Computer Weekly* (Feb. 12, 1970).
83. Wikipedia. *List of multiple discoveries*. 2018. https://en.wikipedia.org/wiki/List_of_multiple_discoveries.
84. F.K. Williamson et al. “A high-level minicomputer”. *Information Processing 74*. Amsterdam: Noth-Holland, 1974, pp. 44–48.
85. Yiu Kwan Wo. “APE machine: A novel stochastic computer based on a set of autonomous processing elements”. PhD thesis. 1973.
86. S. Wolff, J. Thomas, and T. Williams. “The polarity-coincidence correlator: A nonparametric detection device”. *IRE Transactions on Information Theory* 8.1 (1962), pp. 5–9.