Quantum Neurodynamics
or
"Where is the State Vector"

Robert L. Dawes
Martingale Research Corporation
100 Allentown Parkway, Suite 211
Allen, Texas 75002
(214) 422-4570

1. Introduction.

This paper deals more with the computation of physics than with the physics of computation. That, in itself, does not make these ideas unique among those presented at this Workshop: Fredkin's dinner discourse on the universe as a computer is in the same category, for example. However, the focus of this paper is the description of a computational architecture that is capable of autonomous modeling at least part of the universe, observing and comprehending it, predicting its behavior, and -- through motors and transducers -- altering its evolution. To achieve this capability, the architecture computes physics in the abstract, and uses methods from neurocomputing to build associative maps between its internal abstract dynamics and the observed world. It then uses these maps to transform observations of its environment into predictions about its behavior.

The algorithms that we discuss below are not mere abstractions. They have been implemented in both serial and parallel form and they have been demonstrated to work on simple tasks comparable to those performed by the human visual system in tracking multiple moving objects in the visual field simultaneously. They have been developed over the past five years with almost $2 M in funding from Army, Navy, Air Force, DARPA, NSF and NASA sponsors, whose support we gratefully acknowledge.

2. Computation and cognition of the physical world.

Modern systems theory examines the formulation and solution of three fundamental problems in the design of machines and algorithms that "intelligently" observe, predict, and interact with the objects and systems of the physical world. These problems are known as the system identification problem, the stochastic filtering problem, and the adaptive control problem.

The stochastic filtering problem concerns the construction of an optimum "filter" for the transformation of the ensemble of all past observations on some system into an estimate of its state at some past, present or future time. My daughter solves this problem in real time on the soccer field every time she traps a pass. The solution makes use of some form of dynamical model obtained from "system identification", and it comprehends the statistics of the variations (noise) that have occurred in all prior observations. *Every solution to a stochastic filtering problem, according to R.S. Bucy [5], involves the computation of a time-varying*

0-8186-3420-0/93 $03.00

probability density function on the state space of the observed system. This is important to remember.

The principal computational tool for stochastic filtering is the Kalman filter. Unfortunately, the Kalman filter is not suitable for "intelligent" computing, because its restrictions (linear systems, Gaussian observation noise) are too restrictive, its "givens" (a good dynamical model with Gaussian residuals) are seldom given, and the algorithm is too computationally intensive for systems of a practical complexity [21]. Nonetheless, the abstract mathematical theory of stochastic filtering is powerful and comprehensive [3, 5, 13, 15-18, 20, 24], and anyone who deigns to build a machine that can trap a soccer ball or catch flies, or even understand natural language must come to grips with these fundamentals.

The approach to intelligent computing in the old paradigm is usually identified with the term "artificial intelligence" (AI). Practitioners of AI tend to think in terms of Boolean algebra, Turing machines and Markov models. They place great reliance on certain "existence" proofs which assure them that all solutions worth computing can be computed within their paradigm, but their methods only replicate isolated features of intelligence and have no plausible extension to anything as adaptable as the brain of a garden snake. I think that those who avoid the hard math and the hard physics will miss out on solving most of the really fun problems.

3. Quantum mechanics and the computation of consciousness.

The Quantum Neurodynamics approach to intelligent computing starts with the methods of modern systems theory and throws away the Kalman filter. Without the simplicity of Gaussian statistics, we had to find some other way to compute a probability density function -- some way that was not limited to Gaussian or even unimodal distributions; some way that could be computed with massively parallel asynchronous processors; some way that could be a part of *both* the system identification solution *and* the stochastic filtering solution. This latter goal is important, because it seems unlikely that the human brain and other allegedly intelligent systems solve the system identification problem prior to and separately from the solution of the stochastic filtering problem.

Unlike many writers, who start with quantum mechanics and try to fit it into consciousness somewhere (e.g., "The Emperor's New Mind", by Roger Penrose), we began with the fundamental mathematical requirements for conscious computing and found that the nonlinear time-dependent Schroedinger equation suited those requirements. Its solution provides a time-varying probability density function that is not necessarily either Gaussian or unimodal, though it can be both in a special case. Its scalar potential field, through the Ehrenfest theorem, provides a **control input** to the equation that enables one to incorporate it into a predictor-error-corrector loop (also known as the "innovations method", cf [15,16]) for stochastic filtering. And, most importantly, the fact that it is the most general differential model of Hamiltonian dynamics means that it can serve as an abstract "carrier" equation for modeling a large and useful class of physical systems. That is, by using associative memory techniques from the field of neural networks, we can build an explicit map between time series of observed patterns and the trajectories of the abstract wavefunction.

There were only two theoretical problems with the equation: There was no known form of the nonlinearity that would produce soliton solutions in two or more spatial dimensions; and whenever a wave packet is caught in a stationary potential well, it oscillates in the well rather than settling asymptotically to the bottom where the minimum prediction error is located. In practice, however, we found that the nonlinearity that was induced by the feedback in the predictor-error-corrector loop produced both nondispersive propagation and particle-like interaction of wave packets in two spatial dimensions which, to our knowledge, is the first demonstration of soliton properties for the Schroedinger equation in two dimensions. Furthermore, when the equation is integrated on a discrete spatial lattice, one has to work very hard to prevent one of the known artifacts of spatial quantization: **viscosity**! Ha! We can be lazy, keep the viscosity, and we get not only asymptotic convergence in the stationary case, but also a generalization of the model to accommodate both Hamiltonian and mildly dissipative dynamical systems!

The reason for dubbing this the method of Quantum Neurodynamics should now be obvious. The details of the mathematics and its computational implementation are described in [6-11]. It is important to note that QND is not just an abstract curiosity. It has been implemented in parallel on a parallel workstation and it is tracking multiple targets simultaneously in simulated video signals. Moreover, we have explained how QND serves as a quantitative, computable model for a number of neural and psychological phenomena, ranging from certain visual deceptions to multiple personality (wherein quantum tunneling provides the transition mechanism from one personality to another) and inferential reasoning [8]. In this respect, it is useful to think of a conscious train of thought as a traveling Schroedinger wavepacket in the neural medium, whose real and complex parts are mediated by concentrations of a pair of ion species like calcium and potassium.

4. The Equations of Quantum Neuro-dynamics

We now summarize the equations and provide a brief explanation of the neurocomputing architecture, called the Parametric Avalanche (PA), which implements the computations. Figure 1 is a functional diagram of the algorithm, which should be referred to while studying the following equations. This description is adapted from [7]. In what follows, the "state" variable x is customarily taken to range over the phase space of the observed system, but in Quantum Neurodynamics, it ranges over the manifold of the cortex in the observer's brain. That is, every x is associated with a processor that does a few simple things well, e.g., scalar products, Laplacians, and adaptations.

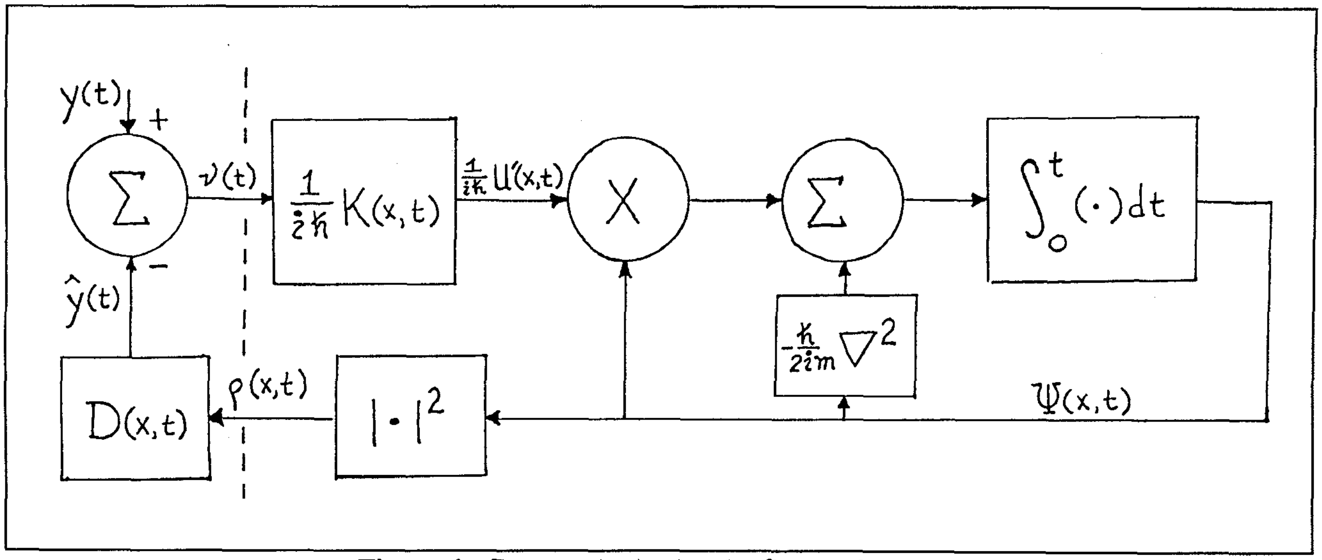

Figure 1. Parametric Avalanche System Diagram

Equation (1) is the nonlinear time-dependent Schroedinger equation from quantum mechanics. An excellent review of its properties may be found in [4].

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\left(\frac{\hbar^2}{2m}\right)\nabla^2\,\Psi(x,t) + \left(U(x,t) + G\left(|\Psi|^2\right)\right)\Psi(x,t) \qquad (1)$$

Here, $\hbar$ is Planck's constant divided by 2π, i is the imaginary unit, m is the mass of the particle and ∇^2 is the Laplacian differential operator. The real-valued function $U(x,t)$ is the scalar potential energy field, which we use via (5) as a control input to the equation. This equation has solutions $\Psi(x,t)$ in the form of complex-valued wavefunctions whose envelope (modulus squared) localizes the position of a particle in that this envelope represents a probability density function for the location of the particle in the vector space:

$$\rho(x,t) = |\Psi(x,t)|^2 \qquad (2)$$

This probability density function is used to compute an expectation of the current observation vector $y(t)$ by computing a ρ-weighted average over the state domain Ω of an adaptive pattern-vector-valued function $D(x,t)$:

$$\hat{y}(t) = \int_\Omega D(x,t)\,\rho(x,t)\,dx \qquad (3)$$

More generally (and more recognizable to a physicist), the expectation can be written as a vector of "observables", i.e.,

$$\hat{y}(t) = \int_\Omega \psi^*(x,t)\,D(x,t)\,\psi(x,t)\,dx \qquad (3a)$$

where we take $D(x,t)$ to be a vector of Hermitian operators indexed over the observation space. For example, if the observable $y(t)$ were a retinal image, then D would map each pair (x,t) onto an array of Hermitian operators, one for each "pixel" on the retina.

The difference between the actual observation vector $y(t)$ and the expectation is the prediction error vector,

$$\nu(t) = y(t) - \hat{y}(t), \qquad (4)$$

which under certain conditions of optimality would be called the *innovations process* [15] of $y(t)$. Such optimality is not practically attainable with this algorithm except, perhaps, after a long sequence of successive applications of the algorithm to the prediction error vector that results from each preceding stage. Such a sequence would implement a constructive Doob-Meyer decomposition of the input stochastic process, and would be a stochastic filtering equivalent of the Gram-Schmidt orthogonalization procedure. Whether optimal or suboptimal, however, $\nu(t)$ is that part of $y(t)$ that cannot be anticipated by this method of computing probabilistic combinations of patterns previously mapped as $D(x,t)$, and it is therefore the part of $y(t)$ that is *novel* with respect to this method at time t.

The linear part of the scalar potential field is the scalar product of another adaptive vector field $K(x,t)$ over Ω with the observation $y(t)$:

$$U(x,t) = -\left(K(x,t) \mid y(t)\right). \qquad (5)$$

The nonlinear part of the scalar potential field is obtained from

$$G\left(|\Psi(x,t)|^2\right) = \left(K(x,t) \mid \hat{y}(t)\right) \qquad (6)$$

The resulting scalar potential field controls the evolution of the Schroedinger wavefunction through the Ehrenfest theorem of quantum mechanics. According to the Ehrenfest theorem, a gradient in the potential field induces a local flow of the probability in the direction of the gradient with an acceleration that is proportional to the magnitude of the gradient. The nonlinearity arises from the fact that the estimate $\hat{y}(t)$ is dependent on the wavefunction through the feedback loop. This nonlinearity produces wavepackets having the properties of solitons. The vector field K(x,t) is the QND equivalent of the Kalman gain matrix, but it is much easier to interpret: At each state x, the vector K(x,t) is a weighted average of previously observed patterns developed in accordance with the adaptation law (8) below.

The adaptation laws for D(x,t) and K(x,t) are just the antiHebbian and Hebbian laws, respectively, of neurocomputing:

$$\frac{\partial D(x,t)}{\partial t} = -\alpha \rho(x,t) \nu(t) \tag{7}$$

$$\frac{\partial K(x,t)}{\partial t} = \beta \nu(t) \rho(x,t) \tag{8}$$

Antihebbian learning is a well known component of novelty filtering [19], and can be derived as a special case of the delta rule learning law when the threshold function is linear and the desired output of the unit is zero. In Quantum Neurodynamics, the objective of learning is to minimize the novelty in the input stochastic process. That is, based on the fundamental assumption that the observed stochastic process is *ergodic*, learning seeks to ensure that if any previously observed spatiotemporal pattern should recur, then the flow of the wavefunction will be "better" in the sense that the ℓ^∞ norm of the prediction error is closer to zero than it was during the previous observation.

Note that both learning laws are intrinsically "supervised" by the probability density function, so that learning occurs only at those states x where there is substantial prior probability for the observed system to be found. Or, to put it another way, the amount of the current prediction error vector that is averaged into the pattern vector that is stored with the state x is (for each x) proportional to the a-priori conditional probability that the state x is the current state of the system, *given* the current historical ensemble of observations. In effect, since that conditional probability is given by ρ(x,t), fast learning occurs in the wakes of large-amplitude traveling wave packets of the Schroedinger equation, but slow learning occurs at any state where the product ρ(x,t)ν(x,t) has nonzero mean over time intervals that are large relative to the reciprocal of the learning rate constants.

5. Where is the State Vector?

Finally, let me mention that while quantum mechanics contributes to this theory of cognitive computing, it is also true that QND may have something to contribute to quantum mechanics. When von Neumann considered certain paradoxes of quantum mechanics, such as the Einstein-Podolski-Rosen paradox and the Schroedinger's cat paradox, he concluded that one could resolve them with the "absurd" mechanism of assuming that the state "of" a system does not belong "to" the system, but rather belongs in the mind of each observer. That was a bit much for most physicists to swallow, until Bell proved that Quantum Mechanics had to be either nonphysical or nonlocal.

The theory of Quantum Neurodynamics does indeed place the state of every observed system squarely and physically in the mind of the observer. In so doing, it places the observer in a symmetric, rather than a priveleged, relationship to the observed. That is, the act of observation by a cognitive system is a physical interaction and nothing more. But the nonlocality of the state of physical systems becomes, in Quantum Neurodynamics, an intellectually tractible, physically realistic concept. Thus, for example, if I measure the state of the proximal component of an E-P-R system the result has an immediate effect on the wavefunction of the distal component, but **both** wavefunctions are in my own head, so there is no problem of anticausal effects.

Note that in QND, observations do not "collapse" the wavefunction of the observed system, but rather they "modulate" their evolution strictly in accordance with principles of information theory and stochastic filtering. This comes about because the mathematical model of observation that is contained in the equations surrounding 3a is more complex, but more physically meaningful, than the simple application of a Hermitian operator to some wavefunction supposed to "belong" to and be localized at the observed system. In QND, an observation is an interaction in which a signal from the observed system (perhaps after considerable amplification) produces a response y(t) in a sensory transducer. That response drives a recurrent filter that produces a wavefunction $\psi(x,t)$, an estimate $\hat{y}(t)$, and updated associative transformations, K(x,t) and D(x,t). The response of the sensory transducer is the "passive component" that is produced by every measuring instrument, whether it is cognitive or not. But the other four products are the "cognitive components" of the observation and only a specially equipped instrument can construct them.

Does any of this change the fact that Quantum Mechanics is a demonstrably correct model of (most of) the universe? Not that I can discern. One just has to get accustomed to the fact that an experiment involving a quantum mechanical system is "out there in the world", but the mathematical and physical models by which we understand and predict their behavior, which includes the state vector, are "all in our heads".

6. Conclusion

Quantum Neurodynamics is perhaps not the customary way in which one thinks of the convergence of physics and computation, but I think that it is an important and, if I may say so, rather elegant synthesis of the two. It provides a model of cognitive sensory processing that not only illuminates the possible structure and function of neurological systems, but also contributes to the structure and function of the theory of quantum mechanics. And if this computational model develops as I think it might, it is almost certain that device physicists will soon be challenged to find structures in silicon (or perhaps in hydrocarbons) that will efficiently carry out the processing that it requires.

7. Bibliography

1. Ablowitz, M., and Segur, H., Solitons and the Inverse Scattering Transform, SIAM, Philadelphia, 1981.
2. Ames, W.F., Numerical Methods for Partial Differential Equations, Academic Press, New York, 1977.
3. Anderson, B.D.O., and Moore, John B., Optimal Filtering, Prentice-Hall, Englewood Cliffs, NJ, 1979.
4. Bialynicki-Birula, I., and Mycielski, J., "Nonlinear Wave Mechanics", Annals of Physics 100, 62-93 (1976).
5. Bucy, R.S., "Linear and Nonlinear Filtering", Proc. IEEE, 58, No. 6, June 1970.
6. Dawes, R.L., "Quantum Neurodynamics: Neural Stochastic Filtering with the Schroedinger Equation", Proc. IJCNN, Vol. I, p. 133, 1992
7. Dawes, R.L., "Advances in Quantum Neurodynamics", Proc. First Appalachian Conference on Behavioral Neurodynamics, (to be published by INNS), Radford, VA, October, 1992.
8. Dawes, R.L., "Inferential Reasoning through Soliton Properties of Quantum Neurodynamics", Proc. IEEE Conf. on Syst., Man & Cyber., Chicago, November, 1992.

9. Dawes, R.L., "Quantum Neurodynamics and the Parametric Avalanche", Technical Report MRC-NASA-89004, Martingale Research Corporation, Allen, TX, August, 1989.
10. Dawes, R.L., "Quantum Neurodynamics: A New Neural Network Paradigm for Knowledge Representation, Cognition, and Control", Technical Report MRC-NSF-89003, Martingale Research Corporation, Allen, TX, July, 1989.
11. Dawes, R., "Theory and Feasibility of the Parametric Avalanche for Image Understanding and Control", Technical Report MRC-DARPA-88-003, Martingale Research Corp., Allen TX, March 1988.
12. M. Goldman, "Langmuir wave solitons and spatial collapse in plasma physics", Physica 18D, 1986, 67-76.
13. Ho, Y.C., and Lee, R.C.K., "A Bayesian Approach to Problems in Stochastic Estimation and Control", IEEE Trans. Automat. Contr., vol. AC-9, pp. 333-339, Oct. 1964.
14. Jammer, M., The Conceptual Development of Quantum Mechanics, Mc.Graw-Hill, New York, 1966.
15. Kailath, T., "An Innovations Approach to Least-Squares Estimation, Part I: Linear Filtering in Additive White Noise", IEEE Trans. Automat. Contr., vol. AC-13, pp. 646-655, Dec. 1968.
16. Kailath, T., "The Innovations Approach to Detection and Estimation Theory", Proc. IEEE, vol 58, No. 5, May 1970.
17. Kallianpur, G., Stochastic Filtering Theory, Springer-Verlag, 1980.
18. Kalman, R.E., Falb, P., and Arbib, M., Topics in Mathematical System Theory, McGraw-Hill, New York, 1969.
19. Kohonen, T., Self-Organization and Associative Memory, Springer Verlag, New York, 1984.
20. Krishnan, V., Nonlinear Filtering and Smoothing: An Introduction to Martingales, Stochastic Integrals and Estimation, John Wiley & Sons, New York, 1984.
21. Mendel, J.M., "Computational Requirements for a Discrete Kalman Filter", IEEE Trans. Automat. Contr., vol. AC-16, pp. 748-758, Dec. 1971.
22. Newell, Alan C., Solitons in Mathematics and Physics, SIAM, CBMS-NSF No. 48, Philadelphia, 1985.
23. Sorenson, H.W. (Ed.), Kalman Filtering: Theory and Application, IEEE Press, New York, 1985.
24. Yashin, A., "Continuous-Time Adaptive Filtering", IEEE Trans. on Automatic Control, Vol AC-31, No. 8, August 1986.