Quantum Learning

Ronald L. Chrisley
School Of Cognitive & Computing Sciences
University of Sussex, United Kingdom
ronc@cogs.susx.ac.uk

June 21, 1995

Abstract

The notion of quantum learning machines – quantum computers that modify themselves in order to improve their performance in some way – is introduced. This is followed by a discussion of the advantages that quantum computers in general, and a quantum implementation of a neural network learning algorithm in particular, might bring, not only to our search for more powerful computational systems, but also to our search for greater understanding of the brain, the mind, and quantum physics itself.

1 Introduction

In both the search for ever smaller and faster computational devices, and the search for a computational understanding of biological systems such as the brain, one is naturally led to consider the possibility of computational devices the size of cells, of molecules, of atoms, or even the size of sub-atomic *quanta*. Thus, it should not be surprising to find that the idea of *quantum computation* is not new; in particular, Deutsch has published striking papers on the notion as far back as 1985; and there were speculations on the issue even before that.

This paper continues this speculative vein, but tries to be concrete in describing what a quantum computational system might look like. The most important specific difference from other considerations of quantum computation is a focus on quantum *learning*: quantum computers that modify themselves in order to improve their performance in some way. The type of learning that is considered here is that family of algorithms loosely known as *neural networks*, *connectionism*, or *parallel distributed processing*. Hopfield [10] popularized the idea that any physical dynamical system could be constrained to serve as a neural network, with fixed points of the system acting as "memories", which could be recalled associatively, in a content-addressable manner. In this spirit, this paper considers what advantages the quantum implementation of a neural network learning algorithm might bring, not only to our search for more powerful computational systems, but also to our search for greater understanding of the brain, the mind, and quantum physics itself.

2 What is a quantum computer?

2.1 Deutsch's notion of quantum computation

Deutsch[6] employs a very formal, behavioral notion of what computation is. Correspondingly, he presents a very formal and behavioral notion of quantum computation.

For Deutsch, every physical system is a computer. To understand any given system as a computer, one chooses three observables of it, to function as the input, output and halt observables. Distinct eigenvalues of the input observable are mapped to the intended inputs to the system; similarly for the outputs. The system computes the function f if, whenever the system's input observable is set to the eigenvalue corresponding to input x, it halts in a state with an eigenvalue of the output observable that corresponds to $f(x)$. This relies on a very impoverished notion of computation[1] , but as it is sufficient for our purposes, I will not dispute it here.

Furthermore, for any given computational system, one can construct a multi-dimensional input space, where each input of the system is a distinct dimension; similarly for the outputs. Thus one can view a computer as a matrix transformation from one of these spaces to the other. The difference between classical and quantum computers can then be seen as this: the matrix of a classical computer is sparse, and has exactly one "1" in each row and column (the other entries being "0"); while the matrix of a quantum computer has no such restriction[6, p 76]. Thus, the end state of a quantum computation is, typically, a superposition of classical outputs.

This characterization is useful, in that it helps make precise the notion of quantum computation, and allows Deutsch to derive some of the novel properties of such computers. However, it is so abstract as to be misleading; indeed, I am not convinced that quantum computers have all the properties Deutsch derives using this characterization (see section 4).

2.2 A concrete alternative: The barrier/slit/plate feed-forward network

For the discussion of the advantages of quantum computation on which I wish to focus, a more concrete notion of quantum computation will be helpful. Also, I wish to stress the further advantages of quantum learning over quantum computation that does not involve learning. Therefore, rather than formulating a definition, I will provide an example, and hope that it will make the general notion clear.

Although there are many different neural network learning algorithms, in my example I will use the common feed-forward back-propagation network, as it is the one that is most likely to be familiar to a general audience. The general principles and insight, however, is easily generalizable to other algorithms.

Similarly, many different quantum situations could be used to implement a learning algorithm; I will stick to the situation of a barrier with slits in front of a photo-sensitive plate, because of its familiarity. It is very likely that such a situation would prove impractical in order to obtain some of the computational advantages (especially those of speed and size) of quantum computing,

[1] For example, it doesn't make any distinctions between the possible *ways* that mapping is achieved, a difference of *algorithm*. Furthermore, we sometimes say that a system is computing f even if it sometimes produces an output for x which does not map to $f(x)$.

but as such advantages are relatively obvious, and are not the advantages on which I will be concentrating, this impracticality need not concern us much.

The details of feed-forward networks involve units, activation values, layers, and weights. But what is important at first is that:

- feed-forward networks realize parameterized non-linear functions from an input space to an output space;
- networks modify these parameters in response to interaction with their environments (usually via a "teacher") so that the function that each network yields better approximates some intended function.

One could set up a quantum implementation of this kind of network in the following way:

A particle beam is set up facing a barrier, behind which is a photo-sensitive plate. The barrier has several slits in it, which result in the famous interference patterns on the plates. Some of the slits are designated to be *input slits*, and the rest are the *weight slits*. The interference pattern that results from the beam being aimed at the barrier with a particular input slit configuration is the output of the system for that input. Thus, use of the system to perform some computation will require two mappings: one (I) from inputs (e.g., character strings, images, queries, etc.) to input slit configurations, and one (O) from interference patterns to outputs (e.g., classifications, stored data, etc.).

Assume that one already has a mapping from inputs to n-vectors, and a mapping from m-vectors to outputs. One could map these input n-vectors to input slit configurations by, for example, dividing the barrier into n sections, each having an input slit, with the relative position of the ith slit within the ith section of the barrier indicating the value of the ith coordinate of the input vector. For binary input vectors, this mapping could be even simpler: there is a slit in the nth section of the barrier if and only if the nth coordinate of the input vector is "1".

Less straightforward is the mapping O from interference patterns to outputs. Since interference patterns are of high-dimensionality, any dimension-reducing mapping will do. Perhaps the plate could be divided into m sections, and the average intensity within the mth region can serve as the mth coordinate of the output vector. For the case of binary output vectors, a soft (sigmoid) thresholding of this value would do (the thresholding must be soft so as to allow differentiation for back-propagation learning; see below).

The *error* of the system, E, is defined to be some function of the desired (d) and actual (a) output vectors, typically $\sum(d_i - a_i)^2$. If $S(x, w) = p$ is the function that yields interference patterns (p) given input (x) and weight (w) slit configurations, then $a_i = O(S(x_i, w)$.

Given some such setup, the system could learn an associative mapping f in the following way:

- a number of samples $< x, f(x) >$ are collected as input/output pairs for training;
- the system's weight slits are randomized;
- for each training sample $< x_i, f(x_i) >$, the following occurs:
  - the input slits are initialized according to $I(x_i)$; the plate is cleared or replaced;
  - the beam is activated, until an interference pattern of sufficient resolution is produced on the plate, and an output is produced according to $a_i = O(p_i)$;

- for each weight slit w_j, the partial derivative of E with respect to w_j is estimated. This is done by calculating $\frac{\partial E_i}{\partial w_j}(d_i - O(S(I(x_i), w))^2$;
- this estimate is used to calculate the change to be made to the control variables w in such a way that gradient descent in error is achieved: the change in w_j is proportional to the negative of the partial of E with respect to w_j; the slits are moved accordingly.

After several passes through the training set, this procedure will ensure that the system settles into a weight configuration that produces minimal error on the training set. In many cases, this will also result in good performance on other samples drawn from the same source as the training set (i.e., generalization).

This is enough to establish a correspondence between neural nets and a quantum system. The correspondence can then be used to suggest how the many variations on connectionist learning (e.g., recurrency, competitive learning, etc.) also could be implemented in a quantum system.

3 Implementation issues

3.1 Multi-layer quantum networks

It has been shown[13] that, even though feed-forward networks might contain a non-linearity, they must have at least two layers of non-linear units if they are to be able to compute non-linearly-separable functions, which can be as simple as the function XOR. Thus, for any reasonably powerful form of quantum learning, it might be better to think of a quantum beam/barrier/screen apparatus as implementing one unit in a layered network of units. Any units in a layer beyond the input layer would not have their input slits positions determined by the input sample x directly, but rather by the outputs of the units in the previous layer. In such a case, it would be typical to take the output of a unit to be uni-dimensional, usually soft-thresholded.[2]

This would differ substantially from standard feed-forward networks, in that each weight would modulate all inputs to a unit. In standard networks, each weight only modulates the input from one other unit. This has the result of making the derivative computation during learning a relatively local computation. Whether or not this difference can be avoided, or whether it would be disadvantageous if it could not, has yet to be worked out.

3.2 Two-way quantum networks

One limitation of the scheme so far is that the quantum system really only implements the forward phase of the network. The back-propagation, or learning phase must be calculated off-line, after which the slits are altered accordingly. It might be better if the learning phase were implemented directly in the quantum system as well, by having the desired interference patterns and actual interference patterns *directly cause* the changes in the control slits.

[2] Note that the function that the system realizes must lie between the two extremes of linearity and discontinuity: if the function is merely linear, then it lacks computational power (to avoid this, non-linearities may be introduced in between the barrier and plate); but if it is so non-linear as to be undifferentiable, the gradient cannot be followed during learning.

Imagine a setup similar to the one already described, but with the following additions. Behind the plate, there is another particle beam, directed back toward the original barrier. Furthermore, there are slits in the plate, which will allow the beam to pass through and hit the original barrier, which has a photo-sensitive plate mounted on its back. Thus, the second particle beam will cause interference patterns on the back side of the original barrier.

The goal would be to have the setup work like this: the plate itself would calculate the difference between the actual and desired interference patterns (perhaps by having something like the negative of the target pattern projected onto the plate), then this could cause certain slits to open in the plate, causing characteristic interference patterns on the back of the original barrier. These patterns would in turn cause the weight slits in the barrier to move according to whatever learning rule is being employed.

Of course, implementation-dependent speculations such as these may be premature, or irrelevant, since the principal reason for using both the barrier/slit/plate setup and back-propagating feed-forward networks was not for ease of implementation, but ease of explication.

3.3 Purely quantum networks

Perhaps it seems even more desirable to eliminate all macroscopic entities except those needed to fix inputs, and read outputs. That is, perhaps it would be better if the control variables and the mechanisms which manipulate them were not macroscopic slits, but themselves quantum phenomena.

Perhaps not. Penrose makes some interesting comments[14, p 403, 171-2] concerning the physics of computation that are relevant here. He argues that we can only have computers built out of macroscopic objects because of the discreteness of the quantum level; if there were no underlying discreteness, then there would be an unacceptable degradation of accuracy within any computational system. Furthermore, at least part of this discreteness is provided by the collapse of the superposed wave packet, so a purely quantum computer which does not have its superpositonal states collapsed now and again by macro objects, may be less powerful than a hybrid classical/quantum one.

On the other hand, this limitation of purely quantum systems may only be an impediment to traditional, von Neumann style computation. Quantum neural networks, in that they are more robust and noise-tolerant, do not require as haigh a degree of accuracy, and thus may be able to function adequately without frequent "observations", which collapse the superpositional states.

3.4 Computing with individual quanta as opposed to aggregates

A issue related to the above is this: could one get more computational power by not using aggregates of quantum phenomena, but by using individual quantum events? In the network I have described, the mapping is from *ensembles* of quanta hitting the plate to outputs. One could instead imagine a faster-scale form of computation in which individual quanta hitting the plate are interpreted as outputs. As part of an inherently stochastic process, each quantum hitting the plate conveys not determinate information about the slit configuration, but probabilistic information: what the likely configuration of the slits is. This kind of information may be used on its own during actual computation, but during learning, it seems most likely that many samples will have to be used in order for the network to learn the proper statistics, regardless of whether or not the weight changes occur after each quantum or only after an ensemble.

4 Computational advantages

Carver Mead, a visionary of computer hardware design, has pointed out[11] that until lately, advances in hardware have focussed on issues of *scale*; smaller is better, smaller is faster. But he points out that although the brain's hardware is of much larger scale, and much slower than current computer hardware, the brain can perform computations far beyond our fastest supercomputers. His recommendation, then, for advances in hardware design, is to look at how the brain is organized, for inspiration to come up with new forms of computation, rather than just try to make the kinds of computer we have now faster and smaller. Neural network algorithms are part of this search for novel kinds of computation.

Contrast this with Deutsch's proposals, which make rather radical suggestions concerning size and speed, but which make no specific mention of any non-traditional algorithms. A computer on the spatial and temporal scale of quanta is bound to have advantages over current hardware, but Mead's point still holds. So why not pursue *both* improvements in scale *and* alternative forms of computation? That is one of the reasons for looking at quantum learning; perhaps there are qualitative advantages of quantum learning that are not just the simple addition of algorithmic and scale advantages.

To be sure, Deutsch's discussion does involve the notion of parallel computation. In fact, most of the advantages he sees accruing to quantum computation are due to some sort of parallelization. But he considers only a classical (in an algorithmic sense), symbolic form of parallelization, and does not consider learning at all (a phenomenon which is not best understood as a static transformation of inputs to outputs).

4.1 Deutsch's list

Although Deutsch concedes that quantum computers are not more powerful than classical ones in the sense that they can compute non-recursive functions, he does claim that there are several computational advantages to quantum computers:

Randomness Quantum computers can generate true random numbers, although a classical computer can get arbitrarily close to doing this. Note, however, that a truly random output from a quantum computer requires other outputs with which it is correlated, and which must be hidden from the observer.

Cryptography Quantum correlations can be set up between different programs, and this can be harnessed for quantum cryptography, for example.

Universal simulation? Only quantum computers have a *chance* of being able to simulate every physical system. Whether they can or not must be discovered empirically, but we know that classical computers cannot.

These all seem very sound. But Deutsch goes on to claim other advantages to quantum computation:

Fault-tolerance Quantum computers are more robust than classical computers with the same amount of resources. This is because the superpositional "many-worlds" nature of quantum states supposedly adds an extra amount of redundancy for the same number of processors. I don't think Deutsch has made it clear whether this is more than mere semantics: the same observation could be taken to be an argument that the analogue of a classical processor is not one of Deutsch's processors, but its entire superpositional "family". It is also unclear whether the superpositional parallelism can really be harnessed in the way Deutsch imagines (see below).

Speed Quantum computations can be faster than classical parallel algorithms, because "it is always easier in practice to prepare a very large number of identical systems in the same state than to prepare each in a different state"[5, p 113]. But again, it is unclear that the parallelism can be harnessed in this way, as is discussed below.

Hyper-speed A superposition of n quantum states can be used to perform parallel processing, although only one of the n results will be accessible in a given universe. Although the expected *mean* running time of the computation is no better than a classical parallel version, Deutsch claims that some of the time the computation may take much less time than the fastest possible classical implementation. He reasons as follows: assume that a quantum computer has been set up to compute a task which classically takes at least two days; assume that there is a program that extracts the info from the superpositional state in negligible time, with a certain probability of success per unit time (per day, suppose). Then there is a non-zero probability that the information will be extracted from the superpositional state in just one day, faster than the classical limit. One can just check the halt bit to see if a two-day computation has occurred in one day. Deutsch uses the illustration of a Stock Exchange simulation program that predicts activity one day in advance, but classically takes two days to run; if run on a quantum computer, there will be lucky days where one manages to run the simulation in only one day, so one can actually use the predictions to invest successfully. But this just seems wrong: what reason do we have to believe that there is a program that can extract the information in "negligible time"?

Another problem with this suggestion concerns the "halt bit". Deutsch points out[5, p 104] that a quantum computer "must not be observed before the computation has ended, since this would, in general, alter its relative state. So he requires that there be a halt bit that can be observed, without affecting the operation of the quantum computer. But this seems paradoxical: if the halt bit depends on the computational state, then surely observing it will collapse the superpositional state, just as observing a light on Schrödinger's Box will either kill or save Schrödinger's Cat. On the other hand, if the halt bit *is* independent of the computation, then it isn't really a halt bit, any more than a flip of a coin would be: if the halt bit goes on, it can only be an accident that the machine has in fact halted.

The problem that seems to be lurking behind these exciting possibilities is with harnessing the quantum parallelization. First: even if the superpositional variation does correspond to some task decomposition, only one of the n parallel computational results can be observed in "our universe". So only $\frac{1}{n}$th of the computational task will have been solved in our world.

But it is worse than that: we don't know *which* nth of the work has been done – which is almost as bad as if no computational work had been done at all. It's like the old joke about the stopped clock: at least it tells the right time twice a day. It's a joke because a stopped clock doesn't tell the correct time at all, even when, by chance, what it says happens to match up with the actual time.

Another analogy might help: Deutsch's parallel quantum computations are like someone "computing" the truth of Goldbach's conjecture by tossing a coin (assume, in order to make it as much as possible like Deutsch's case, that the toss is purely random). Sure, one could take "heads" to mean it is true, and "tails" to mean that it is false. And one could claim that there is a superposition of coin states, and that the correct answer is in our universe 50% of the time, but that does not mean that the truth of Goldbach's conjecture is being *computed*. Neither does it mean, *a fortiori*, that the truth of Goldbach's conjecture is being computed *faster* (by the flip of a coin!) than it could be classically.

4.2 Computational advantages, take two

Although they may be rather obvious, mention should be made of the more mundane computational advantages of quantum computation: size and speed. Quantum computers have the potential to be very small indeed, allowing a lot of computational power in a very small space. This is not just because of the fact that quanta are small; it is also because of the nature of the physical forces involved.

The biggest stumbling block, in conventional hardware design, to greater and greater scales of component integration is *not* the size of the components, but the density of connections. Communication in classical systems is via wires, and as components get smaller, there is geometrically less surface area of the component to which one can attach connecting wires. Also, wires have to be insulated from each other, which takes up more space.

In a quantum computer, not only are the components small, but they communicate, not with wires, but with *forces*. The holistic nature of quantum phenomena means that a change in a slit position at one end of a barrier can have a differential effect on all aspects of the output. In a conventional computer or network, this "communication" would require wire connectivity between all inputs and all outputs, which would limit the scale of integration, as discussed above.

Furthermore, in quantum computers, this communication is instantaneous. This increase in speed may or may not be dramatic for the short distances involved in conventional ways of thinking of nano-scale integration. But one can imagine a macro-spatially extended array of quantum-scale computational processors, which could communicate instantaneously (or near-instantaneously if relativistic considerations demand such a restriction) across substantial distances.

One should not be surprised at the fact that Deutsch did not mention such advantages in his discussion, as he would not see them as particularly *quantum* advantages. That is, they do not exploit the superpositional aspect of quantum states, which, for Deutsch, is the defining characteristic of quantum systems (see section 2.1). Since I am not sure that such exploitation is possible, I will refrain from *defining* quantum computation in terms of it, and will instead use a more intuitive criterion based on scale. Roughly, a computer is a quantum computer if the highest level of description with which one can explain its operation is the quantum mechanical level.

5 Physics: Quantum computation and physics

5.1 Deutsch's list

Deutsch mentions some other advantages to quantum computation, which mainly have to do with how it can help us understand other physical phenomena:

Complexity measure Traditional computation-based complexity measures (e.g., the complexity of a string of digits is the length of the shortest computer program that can print that string) have the problem that they classify noise as complex. The stochastic nature of quantum computation allows one to use it to provide a complexity measure that will classify noise as non-complex, since it can be generated by a very simple program on a quantum computer.

Foundations Deutsch suggests that this complexity measure could be used in the following way: one can postulate that the universe moves from the quantum-simple to the quantum-complex, and derive the third law of thermodynamics, and the psychological arrow of time from that.

Experimentation Given that a quantum computer is a true quantum system, one could program it so that its operation actually tests various physical hypotheses.

5.2 The interpretation of quantum mechanics

The aspect of the relevance of quantum computation to physics on which I wish to concentrate has to do with the interpretation of quantum mechanics.

The standard interpretation of quantum phenomena is in terms of wave/particle duality: the quantum is both a particle (it hits the plate at a point) and a wave (the pattern on the plate is just the kind of interference pattern one would get from a self-interfering wave passing through the slits). But there are other interpretations. The two to be mentioned briefly here are Everett's *many worlds* interpretation[8], in which the superpositional state is actually a superposition of universes, one for each possible value of the observable; and Bohm's *ontological* interpretation[4], in which quanta are particles, and their wave-like behaviour is to be explained by the presence of an extra force due to the *quantum potential.*

Deutsch sees quantum computation as implying the many worlds interpretation of quantum mechanics. He uses the many-worlds interpretation freely in explaining his ideas, and although he admits that these explanations could be re-formulated for other interpretations, he feels this can only be done with some loss of explanatory power. Using the Stock Exchange simulator example (see the "Hyper-speed" section in 4.1) he asks: "On the days when the computer succeeds in performing two processor-days of computation, how would the conventional interpretations explain the presence of the correct answer? *Where was it computed?*"[5, p 114, emphasis his]

I'm not convinced that quantum computation supports the many-worlds interpretation, mainly becaused I am not convinced that Deutsch's account of the Stock Market simulator is correct (see 4.1). Furthermore, even if one is convinced (as Penrose seem to be) that quantum parallelism can do the work that Deutsch claims it can, it seems that one can dispute (as Penrose does) Deutsch's claim that this argues in favour for the many-worlds interpretation.[14, p 401, fn 9]

Despite these disagreements, I am intrigued by Deutsch's explicit endorsement of the idea that the various interpretations of quantum mechanics can be distinguished experimentally. While investigating quantum computation, I have been keeping an eye open for possible connections it might have with the ontological interpretation. Whereas Deutsch claims that Everett's interpretation is favoured by the phenomenon of quantum parallelism, I have been looking to see if the ontological interpretation is suggested by the phenomenon of quantum learning. Or, weaker: if the ontological interpretation might be useful in thinking about, and designing, quantum learning systems. I'm still looking.

6 Brain: Quantum learning in real neural networks?

Another possible use of quantum implementations of neural networks might be as a way to understand what the *brain* is doing. The hypothesis would be that we might get a better idea of the function of some of the brain's features if we view them as implementing a quantum learning machine.

But is there any evidence so far that the brain is sensitive to quantum effects? Not really. There is the well-known study[1] that shows that a single photon striking the retina of a toad is sometimes sufficient to trigger a nerve impulse, but in humans, this phenomenon seems to be suppressed by noise filtering[9].

But as Penrose[14, p 400] points out, this does show that there are some cells in the human body that are sensitive to individual quanta, and therefore the possibility of quantum-mechanical effects in the brain is still tenable. But we would be making the task unnecessarily difficult if we, like Penrose, required that we find neurons that are sensitive to a single quantum. As discussed in section 3.4, quantum computation can still occur in the cases where an aggregate of phenomena (an entire interference pattern, rather than one quantum) is required to yield an output. Many of the advantages of quantum computation would still apply in such a situation, such as communication via instantaneous forces, rather than wires.

If quantum learning networks are a more plausible model of brain activity than mere quantum computation, it may have little to do with the fact that the learning algorithm is currently called a *neural network* learning algorithm. Most likely, the bits of the brain that would correspond to the "neurons" in the algorithm would be sub-cellular phenomena. The extra advantage of quantum learning as a model of brain activity would rather derive from its sub-symbolism, and from the fact that it emphasizes learning: brains are learning machines!

7 Mind: Quantum learning and active information

One of the central obstacles to a complete, unified understanding of the world is the *mind-body problem*, which in recent years has been generalized to the problem of *naturalizing intentionality.* One traditional approach to solving this problem, one that is implicit in much work in cognitive science, is to close the gap by making the mind more like the physical, the mechanical. Thus, computers have played an important role.

An alternative hope is this: perhaps we can make the problem easier, not by seeing the mind as physical, as mechanical; but by seeing the physical as having mental aspects, even (or especially) at the lowest (quantum) level.

This kind of idea has been advocated by Bohm[3], and has been furthered by Pylkkänen[15]. Both are concerned with the idea of *active information*, the kind of information a quantum particle carries about its environment. On Bohm's ontological interpretation, there is no wave-particle duality; rather, there is a particle, with a determinate position and momentum. The interference patterns are caused by subtle variations in the quantum potential through which the particle moves, which is in turn the product of the experimental configuration (number and position of slits, etc.) Thus, the trajectory of the particle is such that the particle can be said to carry the information about its environmental configuration. For surely, the particle must "know" that there is more than one slit open if it hits the plate anywhere other than directly behind the slit. It hitting the plate at that place *means* that the slit configuration must be so-and-so. And, the idea goes, if even a lowly particle can be seen to involve such mental items as " meaning", "carrying information" and "knowing", then perhaps the physical/mental chasm can be crossed.

Or so it seems. Actually, this idea has difficulty. Causal or statistical correlation is not the same as knowledge or having meaning. Otherwise, we'd have to say a broken window means something about the stone that smashed into it. And if we have to say that, then it just looks like we're watering down the notion of "meaning" to the point where the chasm opens up again (to mix metaphors a little).

At this point in the argument, Pylkkänen attempts to establish a difference between quantum information and mere classical causal connection by claiming that the former is special in that it is sensitive to *form*, and not just magnitude. But I think attempts to patch the active information idea in such a way cannot save it from failure, because they still ignore the principal difference between our notion of the physical, and our notion of the mental: the mental is *normative.* Thoughts can be correct or incorrect, right or wrong, true or false. But the kind of information that quantum systems seem to have is of the boring non-normative kind; they can't be wrong. A quantum state "means" just whatever caused it; there is no room for falsity or error. Contrast this with thoughts: I might have a thought "That is a horse", that was caused by me seeing something across a field. As it happens, the thing that caused my thought was in fact a cow. This does not mean that my thought was "That was a cow" and was true. No, it continues to have a meaning involving horses, and is therefore false. Unless we can make sense of such notions in a quantum system, then we will still have the large dualistic gap between mind and world. We will still wonder how a non-normative physical system can be the same as a normative mind.

This is where quantum learning might be able to help. Dretske[7, p 35-6] has attempted to naturalize intentionality with the notion of learning. As said before, we can't get a notion of falsity going for a state if we just equate its meaning with whatever causes it. But suppose that we equate the meaning of a state s with some x which comes to cause s during a *learning situation*, a situation in which a system could possibly learn a relationship between x and some relevant behavior. Then even if, later on, after the learning sitaution, some y different from x causes the state, the state will not mean y, because y was not a cause of the state in a learning situation. A biological example: on this account, a rat's brain state B, typically caused by a bell, means, for a rat that has undergone conditioning, that there is food present because food caused (or shared a common cause with) the bell during the learning situation. So if, aftre the learning situation, a bell rings because someone hit it accidentally, and causes state B (as is likely), then that state still means there is food present; it does not *mean* that someone hit it accidentally, even though that was the cause. Since there need not be food present in this case, we have the possibility of falsity. Therefore, a quantum learning system might acquire some form of intentionality, and begin the bridging of the physical/mental gap.

Other approaches to naturalizing intentionality may suggest other forms of quantum computation for those who wish to see quantum systems as intentional. For example, the evolutionary approach[12] proposes that a state s means that p if s is the product of a process of natural selection, and the explanation for why s was selected for was because it was present when p was true. Thus, a quantum implementation of *genetic algorithms*, with their computational version of natural selection, might be another way to get intentionality in at the most fundamental level.

Once one has a notion of quantum information, one might use this show how information can be *implicit*, yet causally potent. The information that a particle has about its environment is not explicitly represented in the particle, but it does have causal effect: it causes the particle to move in a particular way. Now one might think that one does not have to talk of information in this context at all; rather, we have action at a distance. The represented environment itself *directly causes* the action, so there is no need to invoke implicit information as a causal agent. But if we can make sense of a system being correct or incorrect, then we will not be able to say that what is represented is the direct cause, because the world might not be the way the particle's information takes it to be; the represented might not even *exist*. This idea of implicit but causally potent information has strong resonances with Bohm's idea of the *implicate order*[2].

8 Acknowledgements

This paper was originally presentated at the Brain, Mind & Physics workshop, Prague, September 14-17, 1993, organized by the Centre For Theoretical Studies at Charles University. I wish to thank Ivan Havel and CTS for making my participation there possible.

Many of these ideas benefitted from very fruitful discussions with Paavo Pylkkänen, Larry Gould, Ivan Havel, and with the partcipants in the symposium on "Physicalism, Connectionism, and Consciousness" held in Siuntio, Finland in September of 1992. I am very grateful, then, for the support of the Finnish Academy and the University of Helsinki, which made these interactions possible.

References

[1] Baylor, D., Lamb, T., and Yau, K. (1979). Responses of retinal rods to single photons. *Journal of Physiology*, 288:613–34.

[2] Bohm, D. (1980). *Wholeness and the Implicate Order*. Ark, London.

[3] Bohm, D. (1990). A new theory of the relationship of mind and matter. *Philosophical Psychology*, 3(2):271–286.

[4] Bohm, D. and Hiley, B. (1992). *The Undivided Universe*.

[5] Deutsch, D. (1985). Quantum theory, the church-turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A*, 400:97–117.

[6] Deutsch, D. (1989). Quantum computational networks. *Proceedings of the Royal Society of London A*, 425:73–90.

[7] Dretske, F. (1986). Misrepresentation. In Bogdan, R., editor, *Belief*. Oxford University Press, Oxford.

[8] Everett, H. (1983). 'relative state' formulation of quantum mechanics. In Wheeler, J. and Zurek, W., editors, *Quantum Theory and Measurement*. Princeton University Press, Trenton.

[9] Hecht, S., Shlaer, S., and Pirenne, M. (1941). Energy, quanta and vision. *Journal of General Physiology*, 25:891–40.

[10] Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the USA*, 9(2554).

[11] Mead, C. (1989). *Analog VLSI and Neural Systems*. Addison Wesley, New York.

[12] Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. MIT Press, Cambridge, Mass.

[13] Minsky, M. and Papert, S. (1969). *Pertceptrons.* MIT Press, Cambridge, Mass.

[14] Penrose, R. (1989). *The Emperor's New Mind.* Oxford University Press, Oxford.

[15] Pylkkänen, P. (1992). *Mind, Matter and Active Information: The Relevance Of David Bohm's Interpretation of Quantum Theory to Cognitive Science.* PhD Dissertation. Finnish Academy Press, Helsinki.