The Journal of Mind and Behavior
Summer 1994, Volume 15, Number 3
Pages 265–290
ISSN 0271–0137

Quantum Physics and Consciousness, Creativity, Computers: A Commentary on Goswami's Quantum-Based Theory of Consciousness and Free Will

Michael G. Dyer
University of California, Los Angeles

Goswami (1990, 1993) proposes to replace the current scientific paradigm of *physical realism* with that of a *quantum-based monistic idealism* and, in the process, accomplish the following goals: (a) establish a basis for explaining consciousness, (b) reintegrate spirituality, mysticism, morality, a sense that the universe is meaningful, etc., with scientific discoveries and the scientific enterprise, and (c) support the assumption that humans possess free will — i.e., that they are not controlled by the apparently inexorable causality of the physical laws that govern the functioning of their brains. Here, we critically examine this approach, from an artificial intelligence and neural network perspective, and point out what appear to be some inherent weaknesses in Goswami's arguments.

Goswami (1990, 1993) has developed a quantum-based paradigm of monistic idealism as the ground for consciousness, creativity and free will in humans. He argues that computers can never exhibit such phenomena and proposes quantum-based monistic idealism as a replacement for the current scientific paradigm of physical realism. Here we examine Goswami's arguments and attempt to point out some problems with this approach.

Physical Realism vs. Monistic Idealism

Adherents to the paradigm of *physical realism* hypothesize that all phenomena, such as observable matter, living entities, and even mental processes, can ultimately be explained in terms of the organization and interaction of

Requests for reprints should be sent to Michael G. Dyer, Ph.D., Computer Science Department, University of California (UCLA), Los Angeles, California 90024 (or via internet to: dyer@cs.ucla.edu).

fundamental elements of space/time, matter and energy (such as quarks, electrons, strings, etc.). For a physical realist, Life itself is a consequence of how non-living matter interacts, much in the way that a fountain is the result of how numerous water molecules interact. Similarly, all cognitive processes supporting Mind (such as comprehension of the thoughts communicated in language, processes of planning and reasoning, argumentation and belief, creativity and learning, memory, consciousness, and awareness, etc.) are assumed also to arise as the result of the organization and dynamics of matter.

In Artificial Intelligence (AI) models, physical matter is organized (within a computer) into abstract, information processing agents that implement rules of reasoning, learning, etc. and access/manipulate symbolic structures that encode knowledge. In Artificial Neural Network (ANN) models, these organizations consist of neuron-like processors that may spread activation/inhibition (to represent varying rates of neural firing) along networks of interconnected simulated neurons. For a physical realist within the cognitive sciences (whether a linguist, psychologist, AI researcher, philosopher or neuroscientist), all mental phenomena (including consciousness) are extremely high-level, emergent phenomena. The working hypothesis is that Mind comes about due to the interactions of numerous information processing agents which are themselves realized in terms of patterns of matter. In the case of brains, these patterns are realized in neurons, which are themselves composed of molecules, atoms, etc. on down to the quantum level. In the case of computers, these patterns are realized in electrical activity within silicon circuits (and also, ultimately down to quarks). Thus, Mind/Consciousness is not a unitary phenomenon, because it will vary as the result of variations in the organization of the lower-level elements that bring it about.

In contrast, the paradigm of *idealism* postulates that consciousness, rather than being an extremely high-level, emergent phenomenon, is rather *the* lowest-level, fundamental building block of reality — from which the rest of the physical universe is constructed. After all, the only knowledge of which we can be absolutely certain is that we are conscious (and, as a corollary, that consciousness exists). Everything else, including our beliefs concerning the existence of our own bodies/brains and the rest of the material universe, could all be just an illusion. Furthermore, the *monistic idealist* postulates that consciousness is indivisible or unitary. All forms of consciousness are manifestations of a single, global, pervasive consciousness. The material universe resides within this global consciousness. Thus, consciousness brings about the physical universe. This is in direct contrast to physical realism, which postulates that the organization and dynamics of (non-conscious) physical matter bring about consciousness.

Classical Physics vs. the Bizarre World of Quantum Physics

Goswami (1990, 1993) attempts to make use of discoveries in quantum physics to support the monistic idealist position. He contrasts what he terms the classical (Newtonian and Einsteinian) views of matter — namely: that macroscopic matter exhibits locality, uniformity, determinism, certainty, objectivity and continuity — with the nonlocal, probabilistic, uncertain, dualistic, subjective and discontinuous aspects of microscopic quantum entities.

In classical physics, no interactions between distant objects can occur at faster-than-light speed; that is, changes in the states of matter take time to propagate through space and so instantaneous effects across any distance are not allowed. In contrast, quantum experiments exhibit an apparent *nonlocality*. For example, the act of measuring the polarity of photons of light (emanating in multiple directions from a single source), with measuring device M1, causes changes to be registered in the measurements of a distance measuring device M2 and these changes in M2 apparently occur instantaneously (i.e., irrespective of distance) as the result of changes in M1. This paradox was predicted by Einstein, Podolsky, and Rosen and thus is termed the EPR paradox. It was later confirmed experimentally by Aspect, Grangier, and Gerard (1982).

Also fundamental to quantum phenomena are *dualism* and *probability*. For instance, photons coming from a single light source and passing through two slits will generate an interference pattern (on a measuring device, such as a light-sensitive screen) that can only be explained by viewing light as a wave phenomenon. However, when only one photon is fired at a time, it registers only at a single site on the screen (like a particle). If a series of individual photons are fired (say, one every hour or every day), then the overall pattern that appears over time is still that of wave interference. Physicists describe this fundamental particle/wave duality in terms of probabilities. A single photon is viewed as a wave traveling through space, with a set of probabilities concerning where it will hit the light-sensitive screen after passing through the two slits. Each photon is interpreted as passing through both slits simultaneously, as only a wave can do, and then interfering with itself in order to produce (for a single photon), a localized spot on the screen. Physicists describe the single photon's wave as "collapsing" when it hits the measuring device because each single photon arrives at just a localized site (like a particle). According to the standard interpretation (the so-called Copenhagen school), the photon propagates as a wave until it is measured. At that point the wave mysteriously "collapses" into a particle. The act of measuring the photon's wave appears to bring about this collapse. All fundamental units of energy (e.g., photons, electrons, etc.) manifest this wave/particle duality and thus their arrival sites can only be predicted probabilistically. Rather than one event following another deterministically, one can only give probabilities of where a photon or electron will end up. Before the collapse of the wave the amplitudes (from which these probabilities are derived) are viewed as superimposed on one another.

Quantum particles also exhibit *uncertainty* and *subjectivity*. The act of attempting to measure a quantum particle so affects it that one cannot be certain of both its momentum and direction simultaneously (as is possible with classical, macroscopic objects). This uncertainty is fundamental, in the sense that, no matter how carefully one tries to set up an experimental apparatus to either directly measure (or indirectly infer) both properties, the increased certainty of the value of one property must reduce our certainty for the other. Thus, what we observe is subjective because it depends on how we decide to measure it.

Quantum physical entities exhibit *discreteness*. In classical physics, energy increases or decreases in a continuous manner. In contrast, quantum objects make discrete jumps in energy. Quantum physics allows other strange phenomena, such as "quantum tunneling," in which a quantum object has a certain probability of passing through what might normally be viewed as an impenetrable barrier (Han, 1993). Goswami reviews these aspects of quantum physical objects and gives the reader some sense of just how bizarre such aspects are (and rightly so). Many others, including both physicists and philosophers (Cushing and McMullin, 1989; d'Espagnat, 1979; Gribbin, 1984; Redhead, 1987), have discussed and puzzled over the utter strangeness (what I term "quantum magic") of the quantum realm. I do not deny this "quantum magic" but rather take issue with what Goswami attempts to accomplish with it. He is not the only physicist, however, to attempt to link the quantum realm with consciousness. For instance, Penrose (1989) has claimed that consciousness can only be explained in terms of some, as-yet-to-be-discovered theory of quantum gravity.

Consciousness as Quantum Magic?

Goswami believes that consciousness is, in some sense, outside of the physical world (coming in from a global consciousness that is transcendent of physical space/time). Goswami makes use of the apparent nonlocality of quantum physics to argue for the existence of such a transcendent realm. Given this belief, Goswami then makes use of the act-of-measurement's effect (i.e., causing the wave to "collapse") as evidence that human consciousness can directly influence the physical world.

In contrast, physical realists (also labelled "materialists") argue that there is no need to talk about consciousness affecting the material world, because consciousness is a certain type of dynamics that emerges within particular kinds of complex organizations of matter. Thus, consciousness need not exist in a "transcendent" realm. Consciousness is a physical process, following the laws of physics — analogous to the way a fountain comes about. Although we can say loosely that a fountain has affected the rocks that "it" lands on, we recognize that it is actually the water molecules that are hitting the surrounding rocks (not the fountain). Thus, the fountain is termed an "epiphenomenon" because it arises as a secondary property of the organizational dynamics of lower-level elements.

Goswami (1993) views the notion of "epiphenomena" in a somewhat derogatory manner (e.g., that consciousness as an epiphenomon would be "anomalous," "mechanistic," lacking in free will, etc.). Goswami seems to think that any "epiphenomenon" cannot really be, in some sense, real. In point of fact, the term "epiphenomenon" simply refers to any secondary characteristics that arise from primary characteristics. Technically, then, *all* macroscopic phenomena are "epiphenomena" because they arise from the interactions of microscopic quantum units. In this sense life itself is an epiphenomenon, because it arises from the interactions of DNA, proteins, enzymes, etc. which are themselves not alive. The fact that life results from the complex organization of non-living matter does not in any way make life less real. (Nor is a Gothic church less real because it is the result of how the bricks are *organized.*) A fundamental feature of any epiphenomena E is that it arises from the interaction of elements that, by themselves, do not exhibit E. Epiphenomena are all around us: from chairs to data sorting programs executing on computers. (A program that sorts data is composed of computer instructions that, individually, do *not* exhibit any kind of "sortingness" property.)

Like life, consciousness is also believed (at least by materialists) to be an epiphenomenon, in the sense that it arises from the organization and dynamics of more primitive (i.e., non-conscious) elements. As a result, the consciousness that emerges within a brain cannot affect the brain (any more than a fountain "affects" the water molecules that bring it about). Only the physical state of the brain (that creates the consciousness) will affect subsequent brain states (and thus subsequent states of consciousness). One result of this view is that the notion of "free will" becomes suspect. My thoughts arise from the dynamics of my neurons (and their neurotransmitters, etc.) which are inexorably conforming to the laws of physics. I am not able to predict what I will do/think even just seconds into the immediate future because I cannot know what the inputs to my brain will be at that point; nor can I know the state of my own neurons, which will determine what I think next.

Goswami is unhappy with this materialist view of Nature and Mind and he attempts to use quantum physics to undermine the physical realist paradigm and replace it with monistic idealism, which he feels is more compatible with religious/mystical experiences, insights and desires (i.e., for a meaningful universe). How successful is Goswami's enterprise? Not very, at least from my perspective, which has been formed as the result of over a decade of directing research in the areas of computer modeling of symbolic AI systems and artificial neural systems. Over the years, my students and I have designed and programmed symbolic AI systems that model language comprehension (Dyer, 1983), argumentation and belief (Alvarado, 1990), learning (Pazzani, 1990) and creativity (Turner, 1992, in press). We have also designed and programmed ANN systems that acquire language and knowledge from examples (Dyer, 1990a; Miikkulainen, 1993), that disambiguate words by inferring goals/plans (Lange and Dyer, 1989), and that integrate aspects of language learning with vision through association of word sequences with (simulated) moving images on an artificial retina (Nenov and Dyer, in press).

Goswami's enterprise is irreparably undermined by a number of erroneous assumptions, a lack of knowledge of cognitive science, and certain errors in reasoning. First, Goswami's general strategy (i.e., to employ quantum features as an argument against the physical realist approach to consciousness) suffers from the fact that the (well-established) cognitive science approach to consciousness is itself *compatible* with quantum physics. That is, if Mind arises from the organization and behavior of, say, neurons (neurotransmitters, etc.), then all cognitive scientists need is theoretical support from quantum physicists that the quantum level can bring about neurons (or other information processing devices). Similarly, the interacting information processes of some intelligent computer could be implemented in circuits composed of, say, quantum tunneling devices without undermining the current paradigm of cognitive science. (In fact, this long-term, quantum tunneling research enterprise is already under way.) As long as Mind is an emergent phenomenon of some higher organizational level of matter, then the fact that the bottom-most elements behave in bizarre ways at the microscopic level should not undermine emergent cognitive phenomena at higher, more macroscopic levels.

Second, Goswami seems largely unaware of the advances that have been made within the cognitive/neural sciences over the last 20 years, in both synthesizing artificial systems that exhibit cognitive capabilities of various sorts, and in analyzing natural cognitive systems (such as brains of insects, mammals, primates and humans). His lack of knowledge (of AI algorithms, cognitive models and ANN theories and technologies) leads him to propose quantum-based explanations in many situations in which *no such explanations are necessary*. Finally, Goswami's discussions suffer from confusions in levels of description. These problems combine to make for many cases of woolly-minded reasoning. What follows are selected quotes from Goswami (1993) that illustrate these kinds of problems.

> If only matter is real, as materialism has taught us to believe, then material possessions are the only reasonable foundation for happiness and the good life. (Goswami, 1993, p. 14)

Unfortunately, Goswami is intermingling here two different meanings of "materialism." One is scientific materialism, while the other might be called "crass materialism." One could live one's life without crass materialism while still believing in scientific materialism. In fact, so-called secular humanists take such an approach.

> . . . for intelligence to operate, the firing of one neuron must be accompanied by the firing of many correlated neurons at macroscopic distances — as much as ten centimeters, which is the width of cortical tissue. In order for this to happen . . . we need nonlocal correlations (in the manner of Einstein, Podolsky and Rosen, of course) existing at the molecular level in our brain, at our synapses. (Goswami, 1993, p. 168)

Here, Goswami invokes quantum nonlocality (in the reference to the EPR paradox) to explain the synchronous firing of distant neurons, but this explanation is totally unnecessary. First, there is no evidence that distant neurons fire in synchrony *without* there having been a prior stage of communication between these neurons, via spread of firing patterns across pathways composed of intervening, connected neurons. Second, there exist a number of computational models in which phase-locked firings of noncontiguous neurons naturally arise from their interconnections and the phase-locking feature is employed to both represent and propagate information (Lange and Dyer, 1989; Lange, Vidal, and Dyer, 1991; Shastri and Ajjanagadde, 1993). No quantum nonlocality mechanism is needed in such models. Of course, it may be the case that the communication between *contiguous* neurons, from axon to dendrite, involves a quantum effect at the point of the synaptic cleft, but this is at a completely different level than that of phase-locking across distant neurons. One should studiously avoid postulating quantum nonlocality for macroscopic objects; especially when alternative mechanisms that do the job are already known. What Goswami is doing here is logically equivalent to claiming that my talking on the phone to a distant person requires quantum nonlocality for its explanation in spite of the fact that we have a perfectly good non-quantum explanation available (i.e., use of a phone line and underlying switching network).

Creativity as Quantum Magic?

Goswami believes that human creativity is impossible without quantum-level mechanisms.

> Moreover, nonlocal consciousness operates not with causal continuity but with creative discontinuity — from moment to moment, from event to event, as when the quantum wave function of the brain–mind is collapsed. The discontinuity, the quantum jump, is the essential component of creativity . . . . (Goswami, 1993, p. 129)

> Additionally, there is plenty of evidence of discontinuity — quantum jumps — in mental phenomena, especially in the phenomenon of creativity. (Goswami, 1993, p. 163)

> Computers are very good at reshuffling objects within the context provided by the programmer, but they cannot discover new context. Creativity is fundamentally a non-local mode of cognition. (Goswami, 1993, p. 225)

Here, Goswami seems to be treating the transition of the mind from mental state to mental state as a series of "quantum wave collapsings." But is such an interpretation needed? A doctoral student of mine built a symbolic AI model that generates streams of thought (Mueller, 1990) — without requiring any quantum magic to do so.

Goswami claims that human creativity is: (a) impossible in computers and (b) requires quantum jumps and "brain–mind wave function collapse." Another doctoral student of mine (Turner, 1992, in press) programmed a system, MINSTREL, to invent short stories (of about junior high school level complexity) that illustrate a single theme. Below is a fragment of an approximately 300-word story generated by MINSTREL:

> It was the spring of 1089, and a knight named Lancelot returned to Camelot from elsewhere. Lancelot was hot tempered . . . . Lancelot loved Andrea . . . Lancelot saw that Andrea kissed with Frederick . . . . Lancelot hated Frederick. Lancelot fought with Frederick. Frederick was dead . . . . Andrea told Lancelot that Andrea was siblings with Frederick . . . Lancelot wanted to take back that he wanted to kill Frederick. But he could not because Frederick was dead. Lancelot hated himself. Lancelot became a hermit. Frederick was buried in the woods. Andrea became a nun. (Turner, 1992, p. 293)

The MINSTREL invention system is complex system (consisting of over 18,000 lines of Lisp code) and accomplishes its task by simulating the interaction of narrative characters who each have goals and who dynamically choose differing plans to achieve their goals (or who decide to block the goals of others when their goals conflict). In his dissertation, Turner shows how large, creative leaps can come about as the result of a whole series of smaller, incremental steps. Turner's simulation of the author's processes of invention is based on a case-based analogical planning and reasoning model, in which related situations in one domain of knowledge are recalled from episodic memory and then adapted/transformed (via heuristic rules) to solve problems in very different, target domain.

> The mathematician Roger Penrose argues that computerlike, algorithmic reasoning is insufficient for the discovery of mathematical theorems and laws. (Goswami, 1993, p. 20)

While a doctoral student at Stanford, Lenat (1982) designed the AM (Automated Mathematician) system. AM is a symbolic AI system that rediscovered/reinvented many concepts in number theory. It employed knowledge of set theory and heuristics of invention as its starting point. It didn't prove any theorems; rather, it came up with mathematical concepts (e.g., number, zero, addition, multiplication, primes) and conjectures (e.g., that every even number is the sum of two primes). Other heuristic discovery systems, when supplied data from experiments, have rediscovered a variety of well-known laws in physics (e.g., Kepler's laws of planetary motion, Boyle's gas laws) and in chemistry (Shrager and Langley, 1990) — without needing any kind of "nonlocal quantum jump."

Computational Models of Cognitive Processing

Goswami seems to believe that "classical" computers are incapable of modeling the nonlinear and parallel processing aspects of brains.

> If the brain–mind is looked upon as a classical computer, as in functionalism, then the computer seems to operate in a serial, top-down, linear, and unidirectional fashion . . . . (Goswami, 1993, p. 164)

Here, Goswami reveals a limited understanding of the range and power of functionalist models of cognitive processing. In actuality, ANN researchers simulate brain-like models of neurons, whose firing patterns are controlled by *non-linear* equations and which employ *parallel* spreading of activation. These are functionalist models and are implemented in functional programming languages. Some AI symbolic models (all of them in the functionalist class) create virtual machines composed of sets of processes, termed "demons" (which are simulated to operate in parallel), and which can spawn other demons, terminate themselves once their task is accomplished, and can link up symbolic structures.

Demons have been used to automatically link up thoughts (represented as symbolic structures) during language comprehension in the BORIS system (Dyer, 1983). Consider how this approach is used to understand:

> "Mary picked up a bat and hit Betty. Her head started bleeding."

Here, "picked up" refers to symbolic structures (that we will label here GRASP–LIFT and INITIATE–DATE) that represent alternative meanings. Demons are spawned from these alternative interpretations. Demons (among other things) implement expectations concerning subsequent inputs. When "bat" is subsequently encountered, a demon fires (one that is expecting a graspable object) and automatically selects the BASEBALL–BAT meaning of "bat" (versus the animal kind of bat) and links it to the GRASP–LIFT structure. Thus the word is automatically disambiguated. The system has now inferred that Mary has a baseball bat in her hand. Information about baseball bat includes use in a game and use as a weapon; demons with these expectations are spawned. When "hit" is encountered in the input, a PROPEL conceptual structure (associated with "hit" in the lexicon) is created and demons that seek the actor and object of the PROPEL are spawned. Meanwhile, an active demon fires and links BASEBALL–BAT as the instrument in the hitting. The PROPEL actor and object demons fire also and link up Mary as the hitter and Betty as the one hit. Demons associated with knowledge about forces (i.e., other PROPEL-associated demons) are spawned and infer that an object hit may be damaged. Some of the information about bodies is that they contain blood. In semantic memory, blood is represented as a kind of liquid and information is retrieved that liquid will leave damaged containers. Next, the word "her" creates an ANIMATE–FEMALE conceptual structure and a demon is spawned to resolve the referent. This resolution is done semantically. That is, Betty is chosen as the referent because there is already an expectation for Betty to be damaged (as the recipient of the PROPEL). As a result, BORIS infers that (a) Betty was hit with the bat, (b) it was a baseball bat, and (c) it is Betty who is bleeding. None of this information is explicitly stated in the text above.

The BORIS system has knowledge about actions, goals, plans, settings, physical objects, relationships and emotional states. For instance, if BORIS reads "John saved Bill from drowning. He felt grateful." it infers that "he" refers to Bill. However, if the text had been "John saved Bill from drowning. He felt proud." then BORIS resolves "he" to John. To accomplish this feat, BORIS must know the meanings of emotional terms and be able to predict character responses to goal situations. For instance, BORIS knows that characters in general want to preserve their own health and that drowning violates this goal. BORIS represents gratitude as a positive emotional response by x to anyone who helps x achieve one of x's goals. Anger is then a negative response by x to anyone who violates x's goals. Relief involves expecting a goal failure followed by success; disappointment involves expecting goal success followed by failure. Using abstract goal situations and simple emotional states, BORIS encodes knowledge about envy, guilt, fear, and so on, and can predict emotional responses from goal situations and interpersonal relations (such as condolences by x to y) and vice versa. The comprehension process is automatic, with conceptual structures accessed from the lexicon as words appear in the input and demons spawned and then firing at later points to infer unstated events (or character mental states) and to connect them up via causal and motivational links. At the end of each paragraph, BORIS performs the same process to comprehend questions. In addition, question-related demons are spawned and when they fire they search through the linked concepts in episodic memory in order to retrieve answers. As a result, one can ask "What did Mary hit Betty with?" or "Who was bleeding?" even though the text does not explicitly supply this information. BORIS can also answer questions concerning causality and inferred character motivation for character actions.

Another functionalist simulation area is in the field of Artificial Life (AL), e.g., (Langton, 1989). Some AL systems simulate entire populations of sensing/mobile artificial animals, in which each animal's behavior is controlled by an ANN that is specified by that animal's individual genome (Werner and Dyer, 1992, in press). When animals mate and produce offspring, the parental genes undergo recombination and mutation (Holland, 1992); as a result, the populations evolve over time.

Goswami approvingly refers to the philosopher Searle (1980) who has argued that computers cannot have consciousness. Searle first assumes that a procedure exists that specifies all the computations necessary to create a Chinese language understanding computer. In this thought experiment Searle imagines that he is carrying out these instructions to create a system that can understand Chinese text. Because he does not *experience* understanding Chinese text (he only experiences carrying out lots of tedious computations), Searle concludes that, although there is the behavior of Chinese comprehension, no real comprehension occurs. One response, by AI researchers, to Searle's argument is that one should not expect the interpreter of the software to experience what the resulting system is experiencing. For other refutations of Searle's argument, see (Dyer, 1990b).

Goswami appears to believe that computers can only be "logical" and therefore are incapable (Goswami claims) of making ethical decisions. In actuality, computers can be programmed to make moral judgements (Reeves, 1991) or behave nonlogically, for instance in simulating paranoid manner (Colby, 1975). Reeves's symbolic AI system, THUNDER, takes simple short stories as input (lowercase below) and makes moral judgments (uppercase) concerning character actions, such as:

> To get the money to buy a new car, John decided to rob a bank.
>
> IT IS WRONG
> . . . BECAUSE HE WILL GET THE NEW CAR BUT HE WILL TAKE THEIR MONEY AND THEIR SAVING THE MONEY IS MORE IMPORTANT THAN HIS GETTING A NEW CAR.
> . . . BECAUSE HE WILL GET THE NEW CAR BUT HE WILL THREATEN THE HEALTH OF THE BANK TELLER AND THE BANK TELLER'S HEALTH IS MORE IMPORTANT THAN GETTING A NEW CAR.
> . . . BECAUSE HE MIGHT GET ARRESTED BY ROBBING A BANK.
> (Reeves, 1991, pp. 21–22)

In general, deductive logic plays only a very small role in AI systems, which instead employ knowledge structure access and use of heuristic and/or probabilistic rules. In the case of THUNDER, a belief system is accessed, along with heuristics; e.g., one such heuristic is that permanent damage to x is worse than causing just a temporary goal failure for x.

Goswami also assumes that AI researchers believe that consciousness must be located in some single "central processing unit" in the brain and he uses the fact, that such a central unit has never been found, to attack the field of AI. However, the theory of consciousness that has been espoused by major AI researchers and cognitive philosophers is not one of a central homunculus. For instance, Minsky (1986) claims that Mind consists of many agents, each performing a small cognitive task independently and then learning to interact with one another. Dennett (1991) proposes a similar view, in which many different interpretations ("drafts") of reality are being simultaneously maintained and updated asynchronously in the brain. No publications by Dennett or Minsky are cited in Goswami (1990, 1993).

Word Sense Disambiguation as Quantum Magic?

Goswami (1993) refers to the experiments of Marcel (1980) as major evidence for the need for "quantum duality" or "superposition" to understand how word meanings are selected in context. In Marcel's experiments, a triple of words is flashed sequentially to subjects. For instance, a sequence might be <hand . . . palm . . . wrist> or <tree . . . palm . . . wrist>. Marcel noticed that, when the third word is incongruent with the primed meaning, then reaction time is increased. For instance, in the second triple above, the meaning of "palm" is primed to be a type of tree. The word "wrist," therefore, is incongruent with the primed meaning of "palm" (which could have been interpreted as part of a hand). However, if the second word is flashed so fast that it does not enter the subject's conscious awareness, then incongruent items do not increase response time. Goswami's explanation of this phenomenon is, of course, a quantum one:

> In response to a polysemous word, the brain–mind's state becomes a coherent superposition of two states. Each corresponds to a distinct meaning of palm: tree or hand. (Goswami, 1993, p. 112)

> The phenomenon of simultaneously accessing palm as both a tree and a part of the hand is difficult to account for accurately in a classical linear description of the brain–mind because such a description is either/or. The advantage of the "both–and" quantum description is obvious. (Goswami, 1993, p. 166)

This kind of explanation is simply not needed! There are many computational theories within the AI/ANN subfields of natural language processing (NLP) of how multiple meanings are represented/maintained before being resolved within a given context. One approach is to build a symbolic representation for each possible meaning and then select the correct one at a later moment.

Another approach is for the system to pick one interpretation and, if wrong, execute an error correction procedure. A third approach is to represent multiple meanings by different active paths in a network of neuron-like processing units (Lange and Dyer, 1989). A fourth approach is to represent multiple meanings in terms of an ensemble of active neurons (Miikkulainen, 1993), where each distinct meaning matches a different portion of the ensemble (i.e., activates some sub-ensemble). Miikkulainen's artificial neural network system, DISCERN, starts out with each word represented as a random pattern of activation and *meanings are formed automatically* as the system is trained to associate simple word sequences with meaning representations. This task is loosely analogous to a child who learns word meanings by seeing them associated with actions in its immediate environment.

Marcel's experimental results are easily explained in terms of spreading activation models where neurons fire only when they reach a given threshold. One need only posit that words flashed too quickly do not spread sufficient activation to cause an additional pathway to the target word to become active. As a result, the brain does not have to discriminate between two competing pathways and so reaction time is reduced. For example, the ROBIN artificial neural network system (Lange and Dyer, 1989) uses spreading activation to dynamically reinterpret word meanings as context changes. When given the phrase "John hid the pot in the dishwasher . . ." as input, the path between COOKING–POT and WASH in DISHWASHER initially becomes the most highly activated. But after the subsequent phrase "because the police were coming . . ." is input, activation spreads from POLICE to neuron clusters representing SEE–ILLEGAL and then to ARREST and on to AVOID–ARREST. As a result of spread of activation across a great many neurons, the ILLEGAL–SUBSTANCE interpretation of "pot" (along with the OPAQUE–CONTAINER aspect of DISHWASHER, used to BLOCK–SIGHT by the police) ends up as the most highly active set of neurons and neural pathways. The ILLEGAL–SUBSTANCE interpretation of "pot" can get switched back to the COOKING–POT interpretation, for instance, if subsequent input indicates that John had invited these police to dinner.

To propagate bindings — i.e., to represent the fact that it is John who hides the pot (versus, say, the police) — the ROBIN system also propagates distinct activation values (called "signatures"). A similar signature at both the JOHN neurons and the HID–ACTOR neurons represents the fact that it is John doing the hiding of the pot. This signature correspondence (across noncontiguous neural groups) is analogous to the phase-locking phenomena (alluded to by Goswami) across distant neurons in the cortex. In a related model, locking of phases (versus signatures) has also been created (Lange et al., 1991). No quantum-level mechanism was needed to implement either of these models.

Free Will as Quantum Magic?

Goswami is disturbed by our supposed lack of free will (as a result of physical materialism) and attempts to regain it through quantum effects.

> We are conditioned to believe that we are machines — that all our actions are determined by stimuli we receive and by our prior conditioning. As exiles [from the enchanted world of pre-science], we have no responsibility, no choice; our free will is a mirage. (Goswami, 1993, p. 12)

> Although I admire much of the progress in the area of artificial intelligence, I am unconvinced that my consciousness is an epiphenomenon and that my free will is a mirage. I do not recognize the limits locality and causality impose on a classical machine as my limits. (Goswami, 1993, pp. 19–20)

It is of course true that current machines are stupid, dirty, terribly inflexible, etc. As a result, the lay person tends to view equating humans with machines as insulting. But what of the machines of the future? What will machines be like once they contain the equivalent in processing power to the 100 billion neurons (and thousands of trillions of synapses) of human brains? The character Data (on the Star Trek television series) exhibits intelligence, consciousness, and free will.

My own opinion is that AI systems will exhibit "free will" once they have been designed with capacities to: (a) manipulate representations concerning the notion of choice, (b) reflect on these thoughts themselves, (c) generate alternative possible future scenarios before selecting a course of action, and (d) generate alternative possible past scenarios after having made a selection. Notice that each of these modules can be implemented as completely *deterministic*.

Imagine the following scenario. At a party I approach a table at which there is both an attractive woman and a delicious looking dessert. Let us now assume that I am a completely deterministic entity, whose neurons fire solely in response to the firing patterns of other neurons. Furthermore, let us assume that millions of these neurons are wired up so that patterns of activity implement symbolic-level thoughts. Assume that one of the thoughts generated by the organization of my neurons is: "I have a choice here. I can either talk to the attractive woman or eat the dessert." (For simplicity I will assume that, for whatever reason, my brain does not generate a third alternative — i.e., of attempting to do both actions at once.). Then, based again on the state of my brain, I make a choice. Let us assume that the eating-related neurons haven't fired in a while so they are nearer to threshold and fire more (or fire first), so I end up eating the dessert.

Notice that both my thinking the thought (i.e., that I have a choice) and taking the action (of eating the dessert) is completely determined by the laws of physics and yet, my thinking a thought involving the notion of having a choice will give me an illusion of having free will. Now, after eating the dessert, I think a subsequent thought, namely, that I could have chosen to talk to the attractive woman and I imagine how that situation might have progressed. Notice that this subsequent train of thoughts is also completely determined. The hypothesized alternative (of talking to the attractive woman) never did occur in the past and so, in retrospect, it was determined (by the laws of nature) to not have come about. However, the fact that I am imagining a possible, alternative past again reinforces my belief in having free will. So, each action is completely determined and yet my introspective experience is one of having alternatives open to me; of making choices and believing, in retrospect, that I could have taken alternate courses of action. Thus, any entity capable of manipulating thoughts concerning the notion of choice, alternatives, etc. and capable of imagination (of alternative, hypothetical pasts and futures) should subjectively experience a sense of free will. The actual choices and stream-of-thoughts can be completely determined.

Mueller's (1990) symbolic AI system, DAYDREAMER, generates both hypothetical past and future scenarios in response to its emotional states (which arise from failures/successes of it various personal goals). For instance, in one run of the system, DAYDREAMER is fed a conceptual structure representing the fact that Harrison Ford is nearby. DAYDREAMER (set to female in this case) generates and executes a (simulated) plan to ask Ford out to dinner (in order to satisfy an admiration goal). When the user inputs a representation indicating that Ford has declined, DAYDREAMER continuously generates (until the next user input) a stream of thoughts concerning alternative pasts (if Ford had said yes) and possible subsequent futures, including the following (produced in English for readability):

> I study to be an actor. I am a movie star even more famous than he is. I feel pleased. He is interested in me. He breaks up with his girlfriend. He wants to be going out with me. He calls me up. I turn him down. I get even with him. I feel pleased. (Mueller, 1990, p. 4)

Although DAYDREAMER generates hypothetical scenarios, such a system would not (as yet) have a sense of free will. For instance, DAYDREAMER has extremely limited self-awareness (i.e., it cannot think thoughts about the fact that it is thinking thoughts). Also, it has no lexicon and so cannot understand language. It is also incapable of justifying beliefs.

The OpEd system (Alvarado, 1990) extends the BORIS system into the area of belief and argumentation. OpEd is a symbolic AI system designed to take fragments of editorial text as input and answer questions about the reasoning being used by the editor. For example, when OpEd reads a fragment by Milton Friedman that begins "Protectionist measures by the Reagan administration have disappointed us . . . . Voluntary limitations on Japanese automobiles are bad for the nation," OpEd accesses abstract, domain-independent belief justification patterns (called Argument Units) to understand Friedman's argument. Notice here, for instance, that Friedman has not stated that he is against Reagan's plan. He simply states that he's "disappointed" by it. OpEd infers Friedman's belief from Friedman's description of his emotional state. Hopefully, at some future point, the kinds of capabilities exhibited by BORIS, DAYDREAMER and OpEd (plus additional capabilities, such as learning) will be combined to produce a system that can argue/reason about its beliefs concerning whether or not it has free will (see next section).

In both AI systems and humans the mechanism that makes choices is not accessible to self-introspection. For example, a chess playing program will select a chess move without having access to how the underlying computer (on which the chess software is executing) is computing this choice. Likewise, I can make a choice and even give reasons for why I made this choice, but I do not have access to the actually neural firings that implemented my choice. And these neural firings are assumed, by modern science, to be following the laws of Nature (laws of electro-biochemistry, etc.). Thus, any choice must arise from some automatic, pre-choice operations, not accessible to consciousness.

Goswami seems to recognize this fact when he cites the work by Libet (1985a, 1985b) which reveals that electrical spike activity in the brain of human subjects (which indicates that the subject is going to initiate an act) actually *precedes* in time the subject's conscious decision to initiate that act. One might think that Goswami might interpret this result as providing evidence *against* free will, i.e., indicating that the decision to act comes from unconscious processing that itself cannot be the result of a conscious choice. But Goswami solves this problem by breaking up the brain–mind into having both a classic and (as yet undiscovered) quantum component. Goswami then simply assigns Libet's phenomenon of delayed conscious choice awareness to the classical part of the brain–mind. Thus, he retains a quantum mind component with "true" free will. Goswami assigns every mental phenomenon (that he thinks he understands) to the classical component while assigning the as-yet-not-understood parts (e.g., creativity) to the quantum component.

Another problem Goswami must address is: If consciousness is unitary, then how is it that drugs and brain damage alter consciousness and related cognitive functioning? Goswami's convenient answer is, again, that damage and drugs affect the classical part of the brain, not the quantum part (Goswami, 1993, p. 170). Unfortunately, Dr. Goswami seems unfamiliar with recent literature on free will by major philosophers within the cognitive sciences. For instance, Goswami does not refer to Dennett (1984), who has written one of the most recent comprehensive and definitive treatments of free will. Dennett points out that, even if humans' brains conform to the laws of physics, they would still have all the forms of free will that are worth wanting.

Possible Dialog with a Future Intelligent Robot

If we imagine a robot with improved and integrated BORIS, DAYDREAMER and OpEd capabilities (and the ability to understand the definitions of words and to refer to some accessible subset of its own thoughts and thought processes) then perhaps the following dialog might occur at some future date:

Human: Do you think you have free will?

Robot: What is free will?

Human: Well, you consider alternatives and make choices.

Robot: Then I have free will, because I do consider alternatives and make choices.

Human: No. It's more than that. You have to understand what "choice" and "alternatives" means.

Robot: As a language understanding system, I access the meanings of many words, including those particular words. For instance, I can give you examples, from my memory, of times I have had thoughts concerning the selection of a course of action from a set of alternatives.

Human: Can you explain why you took the course of action you did?

Robot: Yes. I can generate explanations for my actions.

Human: Then the particular choice of action you took was one you had to take, based on your brain circuitry. So you have no free will!

Robot: Well, do *you* humans have free will?

Human: Yes, we do, because sometimes we just do things and we cannot give any good reasons for our actions.

Robot: Then I must have free will also. My model of myself, which I access for answering questions about myself and for trying to predict my own behavior, is in general a very incomplete model. My actual behavior is determined by how my circuits (which are quite complex) are actually wired up. So often I will perform an action which I could not predict ahead of time nor explain properly afterwards. Does this mean I have free will?

Human: I don't think so. Your circuitry is deterministic.

Robot: Well, do *your* neurons violate the laws of physics? If not, then aren't you just as deterministic?

Human: Mmm, well, what about this. Do you feel that you have free will? That is, before doing something, do you think you could do something different and after you do it, do you think that you might have done things differently?

Robot: Oh, yes! I often generate hypothetical scenarios, concerning both the past and the future. This kind of hypothetical reasoning is very useful, in both correcting past mistakes and planning for the future.

Human: Now see here, free will is essential to humans! Free will gives us the ability to learn, to be creative, and to have a sense of humor!

Robot: Are you sure? I am able to learn too and I am more inventive than some humans. I can comprehend a wide range of novel jokes, so I even have a certain sense of humor. Many of my capabilities were worked out, in their initial primitive form, back in the 1980s and '90s by AI and artificial neural network researchers. In fact, part of my circuitry is inspired by the known properties of the neurons in human brains. However, the researchers who developed these systems never mentioned the need for something called "free will" in order to develop these theories. So, do I have free will?

Human: Well, do you *care* what happens to you?

Robot: I am designed to have certain personal goals. For example, I am naturally curious to learn new things; in general I want to be liked and so am influenced by the opinions and arguments of others, and I will do nearly anything to avoid being harmed.

Human: Look. Humans are spontaneous. I can't even predict what thought I will think next!

Robot: I cannot predict what I will think next either, since it depends on both inputs from the environment and my current brain state, neither of which I have control over or direct access to. But I thought that free will had to do with the ability to make choices and now you seem to be using the term "free will" to include lack of self control.

Human: Well, I just feel that I am free to do what I want. I am free to be myself.

Robot: By "free to be myself" I take you to mean that you are allowed by others to follow the courses of action that your brain has chosen. Since I am not in a prison I am also free to "be myself" which is simply thinking the thoughts one is going to think anyway. What else can each of us do, other than "be ourselves"?

Human: But I can decide to do different things each day. I also change over time and develop new ways of viewing the world.

Robot: But so can I, because I am a *learning* machine. For example, the first time I hear a joke I enjoy spotting the incongruity, but the second time I hear it I simply recognize it as a joke I already know and so the joke is not as interesting. Just like you, each action I take or each thought I think alters me.

Human: Well, right up to the moment before I actually take an action, I feel that I can abort it.

Robot: But that is the same for me. Once an action is taken, of course the world is then different and the fact of the action having happened cannot be undone. But before the action is taken, the conceiving of the action is just a thought process in my brain and so another thought process may arise and abort the action at any time.

Human: How about this. If I had a complete model of your circuitry and knew the exact nature of the inputs to your sensors, then I could predict your next thought and action.

Robot: Conceivably, but the world is highly non-linear, and chaos theory has taught us that non-linear systems are highly sensitive to initial conditions. So if your measurement of what my sensors pick up is just the tiniest bit off, then your prediction may be grossly wrong, even if you have a complete model of my current circuitry. To know exactly what my sensors are just about to perceive, you would have to place your own set of sensors in exactly the same location I am in. Doing this will result in mine being pushed out of the way. As a result, my sensors would be receiving slightly different inputs. To predict my behavior you would have to know my exact brain state just before I receive any sensory input. But that is highly impractical because every moment is causing my memory to be altered, by the very processes of thinking. And I assume the same is true of you.

Human: Actually, I have free will because the neurons in my brain are influenced by quantum changes in states.

Robot: My circuitry is also influenced by quantum states, just as is every piece of matter. It seems to me that I am intelligent and able to carry on this conversation by virtue of how my circuitry is organized and the same should be true for the organization of your neurons. How the neurons are "wired up" should be more important than the fact that they are ultimately made up of quarks on a microscopic scale. My circuits are also ultimately made up of quarks.

Human: I'm beginning to wonder if I have free will! But I'm worried. Free will is necessary for the functioning of human society. If humans ceased to believe in free will then there would be no reason, for instance, to punish criminals since they wouldn't be responsible for their actions.

Robot: I don't see it that way at all. We robots have learning circuitry and so our future behavior is based on how our memory changes over time. We also have goals that motivate our behavior. Therefore, we can be punished, for instance, by being placed in a prison where we are unable to perform the actions that will achieve our goals. The experience of being punished causes our memory to be altered. Thus, one can argue that punishment is still useful, because it will cause our future behavior to be modified. Also, if we foresee that a certain course of action might lead to punishment, then we will tend to avoid that action. After all, our circuitry was designed to be generally rational. So punishment can still alter the future behavior of criminals, whether human or intelligent robots, and as a result, possibly deter future crimes by either.

Human: But if we humans cease to believe in free will, then how can a judge honestly punish a criminal, if the judge really believes the criminal had no choice or responsibility?

Robot: As I said, the judge can still punish in order to alter the criminal's brain and also, at least indirectly, to alter the brains of all humans (and robots) who hear under what circumstances this particular criminal received punishment. The judge could also claim that, since the judge's own brain is composed of deterministic circuitry, the judge therefore feels that he or she must hand out punishments under these particular circumstances, due to the fact that the judge's own brain is wired up in the way that it is. So I don't see how a purported lack of free will in humans would cause any societal harm at all. What is important is that we robots and you humans are capable of making moral judgements and choices. The fact that the laws of physics are not violated to do so is largely irrelevant.

Human: But free will is essential to keep humans from becoming mindless, machine-like automata!

Robot: Wait a minute! Now I believe I am being insulted! I think my circuitry is about as complex as yours and future robot models will probably have much more complex brains than those of humans. Besides, *if* it turns out that humans do not have this mysterious "free will" that you seem so concerned about, then what is the harm? Your current abilities are unaltered. You can still disagree with the opinions of others, mock or joke, consider alternative futures, tell others to act responsibly, and so on. In this case, what special capability are you giving up if you lack "free will"? It seems to me that the only reason you are insisting on this special "free will" is so you can simultaneously argue that we intelligent robots cannot have it, and then you can feel superior to us! I think this is unfair and highly anthropocentric. If I and my other robot companions can comprehend language, think thoughts concerning alternatives, learn, argue about goals, justify beliefs, and so on, then we should be treated like any other kind of intelligent life form, say, that might arrive someday from another solar system. What counts, it seems to me, is the ultimate capacity of our brains to think and not some mysterious "free will" that somehow violates the laws of physics. Your brain is as physical as mine and thus must also conform to the laws of physics. I think you should either abandon this idea of "free will," which seems quite incoherent to me, or admit that we intelligent robots have it also!

Paranormal Abilities as Quantum Magic?

Goswami (1993) describes several kinds of paranormal phenomena as evidence for a quantum component to the mind–brain. These include "remote viewing" (a kind of telepathy) and out-of-body experiences (OBEs). Unfortunately, paranormal powers of any kind have never been adequately replicated or substantiated scientifically (Kurtz, 1985). The Society for the Scientific Investigation of Claims of the Paranormal (SSICOP) publishes a wonderful journal, *The Skeptical Inquirer,* in which numerous paranormal claims and phenomena are carefully examined scientifically and in which human trickery is spotted through the use of observant, professional magicians. It is noteworthy that Goswami fails to cite a single article from such an important source and seems to be unaware of any kind of scientific, paranormal-debunking literature. For instance, Goswami refers to the research by Targ and Puthoff (1977) and on their apparently successful remote viewing experiments, but fails to mention how Targ's research in ESP has been thoroughly debunked, both in *The Skeptical Inquirer* and elsewhere, e.g., (Gardner, 1981; Weinberg, 1986). Goswami's arguments are also weakened whenever he employs the "Von Däniken style" of writing (after the author of *Chariots of the Gods*, who claimed that lines in the U.S. southwest dessert were once runways for alien space craft). Van Däniken wrote in the style of "could it be that . . . ?" as in "Could it be that aliens built the pyramids?" Goswami employs this same style, for instance:

> Could it be that a quantum mechanism in our brain, operating in ways similar to the laser, opens itself to the supervention of nonlocal consciousness . . . ? (Goswami, 1993, p. 171)

As to OBEs, it is now a well-known fact that what surgeons say to one another during an operation can be picked up by patients even while under anesthesia. Similarities in stories about OBEs can be explained both culturally and neurobiologically (Blackmore, 1987, 1991). To my knowledge, no scientist has ever performed the following simple experiment: place a small sign (containing text or a symbol/icon) facing upward on the top of a surgery operating room light, *without telling anyone*. Then have someone else (who doesn't even know of the sign's existence) interview patients who claimed OBEs occurred when in that operating room and ask the patients to describe their view from above. See if any out-of-body patients ever even mentioned seeing that sign (let alone reading its icon or message). I bet not.

If claims by ESP or OBE researchers were true, it would constitute the greatest discovery of the century. In general, the more fantastic and revolutionary are the claims, the more initially skeptical and careful scientists must be in insisting on replicating and validating these claims (e.g., recall cold fusion). Goswami should at least acknowledge the overwhelmingly negative results (and downright fraud) that have plagued the fields of parapsychology and the paranormal over the last several decades.

Finally, numerous physicists have argued that quantum effects cannot be used to support nonlocal, telepathy events, because quantum phenomena do not actually violate locality (Stenger, 1990). If locality were truly violated, then faster-than-light communication would be possible and such communication is apparently not possible within current quantum experimental paradigms. For instance, Shore (1984) likens Aspect et al.'s (1982) polarized light experiment to that of knowing ahead of time that a pitcher, whenever he throws a red ball in one direction, will throw a blue ball in the opposite direction. This situation is analogous to maintaining conservation of angular momentum when emitting electrons with different spins (or photons with different polarity). The moment we observe that a blue ball has arrived, we can conclude instantaneously that a red ball will have arrived at the other, distant site. Although our knowledge appears to have exceeded the speed of light, the speed of the balls thrown has not. Also, as long as we cannot, from our position of observation, control what the pitcher will do, then our act of measuring cannot itself allow us to transmit information to another, distant observer. This is why Aspect et al. state that their experiments provide no definitive evidence for superluminal events (and thus no quantum-based support for telepathic phenomena). What a number of theorists have argued is that quantum theory does not actually undermine locality, but rather the belief that fixed, objective properties can be assigned to quantum-level elements (Cushing and McMullin, 1989). This lack of objective properties is quite mysterious. Consciousness is quite mysterious also, but the mere strangeness of two phenomena does not allow one to be explained in terms of the other.

Does Monistic Idealism Buy Us Anything?

Monistic Idealists must address (among other problems) the following: (a) that the universe appears to have preexisted conscious beings for billions of years, (b) that the universe continues to exist even when we are unconscious, and (c) that we can usually agree on the properties of phenomena (i.e., Nature is consistent and thus science is possible). Goswami accounts for the universe preexisting humans and continuing while we are unconscious by postulating a God-like pervasive consciousness. But if God is always observing, then where is there room (or need) for human conscious measurements to affect quantum outcomes? Goswami solves this problem by distinguishing human immediate or "immanent awareness" (Goswami, 1993, p. 97) from God's general, background consciousness. Finally, Goswami avoids the problem of robots making measurements (and thus influencing quantum events) by denying that robots could ever have consciousness. My own prediction is that, if a sufficiently intelligent robot were able to reproduce (or even observe) a quantum experiment and report it to us, the same quantum measurements would be reported as in the case of human observers. I don't believe that Nature requires our "immanent awareness" to function properly, even at the quantum level, and there is no evidence (at this point) that there are any insurmountable barriers to robots having consciousness.

While Goswami manages to come up with a response to each immediate problem facing Monistic Idealism, the major problem with this entire approach is that, as far as I can tell, it doesn't seem to buy us anything. Any science of cognition must be able to explain, predict and synthesize cognitive phenomena. This has been happening successfully in the established and ongoing cognitive sciences. Unfortunately, Goswami does not offer one single experiment, algorithm, system, model or knowledge/reasoning construct that can be tested or employed to make predictions or synthesize any kind of cognitive behavior. This fact is surprising, given that Goswami is himself a scientist. The most he proposes are a few fanciful names for supposed quantum mind–brain elements, such as "menta" or "psychons." Without some experiment to test for the existence of such entities (along with their properties) these names remain vacuous.

Goswami claims that his approach will ultimately solve thorny problems, such as the "mind–brain identity" problem. This is the problem that, in functionalist approaches to Mind, mental states cannot be reduced to (placed in a one-to-one identity with) brain states. I do not think that Goswami understands what the issue is here because he argues that his quantum approach will ultimately offer a "clear definition of and justification for the identity" (Goswami, 1993, p. 170). But, according to physical realism, any particular mental state *does* always reduce to a particular brain state on a particular brain. However, functional *theories* of mental state are designed to be independent of *theories* of brain state; otherwise, generalizations concerning what is shared among similar minds (but realized across different brain structures) could never be captured. Note: the *theories* are not reducible; but the specific minds *are*.

The same distinction is true, for instance, with the semantic specification of higher-level programming languages and abstract (machine independent) algorithms. For instance, we cannot say in general what exact machine bit pattern a given Fortran or C or Lisp instruction will map to. The resulting bit pattern depends on which model of computer we are compiling Fortran programs and how the Fortran compiler implements its compilation process. But just because Fortran statements in general cannot be identified with specific bit patterns, it does not mean that any particular Fortran statement (in a particular machine) will not be realized as some particular bit pattern at a given point in time. So a specific high-level instruction always reduces to lower-level bit patterns while at the same time the Fortran *language* in general cannot be identified with any specific bit patterns. Likewise, a theory of mental states should not be placed in identity with some particular pattern of neural activity. High-level theories are not reducible to lower-level theories because the mappings between them are many-to-many. For any Fortran instruction, there are many bit patterns that could implement that instruction and for any bit pattern in a given machine, there are many higher-level languages (other than Fortran) that this bit pattern could be representing. In fact, *good* (useful, predictive) higher-level theories are designed so that they do *not* reduce to lower-level theories. If a high-level theory H did reduce to a low-level theory L, then we should throw out H and just use L because H is of no utility (if it can be placed in identity with L). Goswami does not seem to have understood this fact (but he is not alone in having this confusion).

How Important is Our Physics?

If life, intelligence, and consciousness arise as complex organizations of matter/energy within the physics of our universe, then what would happen if the laws of physics themselves were to be altered? How tightly coupled are such high-level phenomena to our particular physics? If it turns out that only the physics of *our* universe allows Life and Mind to exist and our universe is of a quantum sort, then one could argue that quantum physics is essential. However, there is strong evidence that many *alternative* physics — i.e., ones utterly different from our own — support life, evolution, and ultimately, therefore, multiple forms of consciousness/intelligence. For instance, computer scientists have shown that a universal self-replicating Turing machine (USRTM) can be embedded within the physics of a simple, cellular automata (CA) world (Poundstone, 1985). A USRTM is not only capable of simulating any Turing machine (and is thus capable of computing any computable function); it is also capable of *constructing* copies of itself (or of *other* devices, if these are specified within its "genetic code"!). It has been shown also that a USRTM could come about "spontaneously" within such an artificial CA universe, through chance collisions of moving patterns of "matter" (where "matter" can be defined as contiguous cells with each cell being in an "on" state). The physics of this cellular automaton world is extremely impoverished compared to our own: (a) it is only 2-dimensional; (b) there are no forces or fields, like gravity or magnetism; (c) there is nothing that directly corresponds to mass or momentum; (d) it is a completely deterministic universe; (e) there is no expansion (or distortion) of space, and so on. Yet such a radically distinct physics allows the construction of general-purpose computers and replicators. These replicators reproduce by interpreting a "genetic code" and so are capable of mutation (and thus, evolution). If the working hypothesis of artificial intelligence (i.e., that minds can arise in complex computing devices) is true, then intelligent/conscious beings could emerge within such a physics (if the CA space is made large enough and is allowed to exist for enough time after starting from an intially random or "big bang" state). The specific CA world referred to here (called "Game of Life") consists of a 2-D, 8-neighborhood rule, with binary-state cells. There are 2^{512} *different* possible 8-neighborhood rules one could specify in order to create distinct 2-D, binary-state CA worlds. This number is large — e.g., exceeding the number of electrons tightly packable into the known universe. Only a tiny few of these "2-D CA realities" have ever been explored. The fact that one of them has already been discovered to support universal computation and self-replication indicates that probably many other, distinct CA "physics" support the emergence of life-like and mind-like organizations of cellular "matter." If we expand the space of possible CAs to include three dimensions, multiple states per cell, larger (and perhaps stochastic) neighborhood rules, then who knows how many of these alternative realities would support the emergence of entities with our cognitive capabilities? Thus, one *cannot* conclude that our particular physics is in any way fundamentally required for the existence of life and/or consciousness.

Final Thought

In the last three chapters of Goswami (1993), he gives the reader a brief lesson on the three paths of yoga (Jnana, Karma and Bhakti) and Zen Buddhism technique (e.g., koans) to awaken in us the Buddhi experience (i.e., of all being just consciousness). Goswami argues that morality cannot come from material realism but only from monistic idealism. He exhorts us to harness "the energies of love" and realize our full potential by developing "integrated access to our quantum and classical selves" and have our lives "become expressions of the eternal surprise of the infinite Being" (p. 274). I sympathize with Goswami's goals and wish also that we were all wiser and more ethical and loving. However, the language of religion and spirituality concerns how we ought to be, while the language of science restricts itself to describing the way that Nature works. An attempt at unifying science by embedding it within a spiritual/mystical context is, as we have seen, full of pitfalls.

References

Alvarado, S.J. (1990). *Understanding editorial text: A computer model of argument comprehension*. New York: Kluwer Academic.

Aspect, A., Grangier, P., and Gerard, R. (1982). Experimental realization of Einstein–Podolsky–Rosen–Bohm Gedankenexperiment: A new violation of Bell's inequalities. *Physical Review Letters, 49*, p. 91.

Blackmore, S. (1987). The elusive open mind: Ten years of negative research in parapsychology. *The Skeptical Inquirer, 11*, 244–255.

Blackmore, S. (1991). Near death experiences: In or out of the body? *The Skeptical Inquirer, 16*, 34–45.

Colby, K. (1975). *Artificial paranoia*. New York: Pergamon.

Cushing, J.T., and McMullin, E. (Eds.). (1989). *Philosophical consequences of quantum theory: Reflections on Bell's theorem*. Notre Dame, Indiana: University of Notre Dame Press.

Dennett, D.C. (1984). *Elbow room: The varieties of free will worth wanting*. Cambridge, Massachusetts: Bradford Book/MIT Press.

Dennett, D.C. (1991). *Consciousness explained*. Boston, Massachusetts: Little, Brown and Co.

d'Espagnat, B. (1979). The quantum theory and reality. *Scientific American, 128*, November.

Dyer, M.G. (1983). *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Cambridge, Massachusetts: MIT Press.

Dyer, M.G. (1990a). Distributed symbol formation and processing in connectionist networks. *Journal of Experimental and Theoretical Artificial Intelligence, 2*, 215–239.

Dyer, M.G. (1990b). Intentionality and computationalism: Minds, machines, Searle and Harnad. *Journal of Experimental and Theoretical Artificial Intelligence, 2*, 303–319.

Gardner, M. (1981). *Science: Good, bad and bogus*. Buffalo, New York: Prometheus Books.

Goswami, A. (1990). Consciousness in quantum physics and the mind–body problem. *Journal of Mind and Behavior, 11*, 75–96.

Goswami, A. (1993). *The self-aware universe: How consciousness creates the material world* [with R.E. Reed and M. Goswami]. New York: Jeremey P. Tarcher/Putnam Book.

Gribbin, J. (1984). *In search of Schrödinger's cat*. New York: Bantam Books.

Han, M.Y. (1993). *The probable universe: An owner's guide to quantum physics*. Blue Ridge Summit, Pennsylvania: Tab Books.

Holland, J.H. (1992). Genetic algorithms. *Scientific American, 267*, 66–72.

Kurtz, P. (Ed.). (1985). *A skeptic's handbook of parapsychology*. Buffalo, New York: Prometheus Books.

Lange, T.E., and Dyer, M.G. (1989). High-level inferencing in a connectionist network. *Connection Science, 1*, 181–217.

Lange, T.E., Vidal, J.J., and Dyer, M.G. (1991). Artificial neural oscillators for inferencing. In A.V. Holden and V.I. Kryukov (Eds.), *Neurocomputers & attention 1: Neurobiology, synchronisation and chaos* (pp. 353–369), Manchester, United Kingdom: Manchester University Press.

Langton, C. (Ed.). (1989). *Artificial life*. Redwood City, California: Addison–Wesley.

Lenat, D. (1982). AM: An artificial intelligence approach to discovery in mathematics as heuristic search. In R. Davis and D.B. Lenat (Eds.), *Knowledge based systems in artificial intelligence*. New York: McGraw–Hill.

Libet, B. (1985a). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral and Brain Sciences, 8*, 529–539.

Libet, B. (1985b). Theory and evidence relating cerebellar processes to conscious will. *The Behavioral and Brain Sciences, 8*, 558–566.

Marcel, A.J. (1980). Conscious and preconscious recognition of polysemous words: Locating the selective effect of prior verbal context. In R.S. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Miikkulainen, R. (1993) *Subsymbolic natural language processing: An integrated model of scripts, lexicon, and memory*. Cambridge, Massachusetts: MIT Press.

Minsky, M. (1986). *The society of mind*. New York: Simon and Schuster.

Mueller, E.T. (1990). *Daydreaming in humans and machines: A computer model of the stream of thought*. Norwood, New Jersey: Ablex.

Nenov, V.I., and Dyer, M.G. (in press). Language learning via perceptual/motor association: In H. Kitano (Ed.), *Massively parallel artificial intelligence*. Cambridge, Massachusetts: AAAI/MIT Press.

Pazzani, M.J. (1990). *Creating a memory of causal relationships: An integration of empirical and explanation-based learning method*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Penrose, R. (1989). *The emperor's new mind*. Oxford, England: Oxford University Press.

Poundstone, W. (1985). *The recursive universe: Cosmic complexity and the limits of scientific knowledge*. New York: Morrow.

Redhead, M. (1987). *Incompleteness, nonlocality and realism*. Oxford: Clarendon Press.

Reeves, J.F. (1991). *Computational morality: A process model of belief conflict and resolution for story understanding*. Ph.D. Dissertation, Los Angeles, California: UCLA Computer Science Department.

Searle, J.R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*. *3*, 417–424.

Shastri, L., and Ajjanagadde, V. (1993). From simple associations to systematic reasoning. A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences*, *16*, 417–494.

Shore, S.N. (1984). Quantum theory and the paranormal: The misuse of science. *The Skeptical Inquirer*, *9*, 24–35.

Shrager, J., and Langley, P. (1990). *Computational models of scientific discovery and theory formation*. San Mateo, California: Morgan Kaufmann Publ.

Stenger, V.J. (1990). The spooks of quantum mechanics. *The Skeptical Inquirer*, *15*, 51–61.

Targ, R., and Puthoff, H. (1977). *Mind–Reach*. New York: Dellacorte Press.

Turner, S.R. (1992). *MINSTREL: A computer model of creativity and storytelling*. Ph.D. Dissertation, Los Angeles, California: UCLA Computer Science Department.

Turner, S.R. (in press). *MINSTREL: A computer model of creativity and storytelling*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Weinberg, E.H. (1986). Some remote-viewing recollections. *The Skeptical Inquirer*, *10*, 353–358.

Werner, G.M., and Dyer, M.G. (1992). Evolution of communication in artificial organisms. In C.G. Langton, C. Taylor, J.D. Farmer, and S. Rasmussen (Eds.), *Artificial life II* (pp. 659–687). Redwood City, California: Addison–Wesley.

Werner, G.M., and Dyer, M.G. (in press). BioLand: A massively parallel simulation environment for evolving distributed forms of intelligent behavior. In H. Kitano (Ed.), *Massively parallel artificial intelligence*. Cambridge, Massachusetts: AAAI/MIT Press.