Physical computation and the design of anticipatory systems

Sunny Bains
Intelligent Systems and Networks Group
Department of Electrical and Electronic Engineering
Imperial College London, UK
and Department of Design and Innovation
The Open University, Milton Keynes, UK
sunny@sunnybains.com

Abstract - *For a system to learn from experience and predict the future, it must have access to the appropriate information. In conventional computational systems engineering, signals of specific types are sought at specific resolutions. Though this is appropriate for many well-understood problems, it may be inappropriate for other, less-well-defined, applications. These include multi-tasking anticipatory systems that interact with the real physical world. Here, the engineer may not have enough advance knowledge to determine what information will be relevant and what will not, particularly in the context of systems that perform complex tasks. Here we discuss how our physical model of computation can allow information to be used more effectively by either postponing, or forgoing, analogue-to-digital conversion. Further — based on evidence from neuromorphic engineering and theoretical computer science — we suggest that systems designed using the approach will be more efficient, both energetically and computationally, than their conventional counterparts.*

Keywords: Neuromorphic engineering, analogue electronics, embodied artificial intelligence.*

1 Introduction

We have developed a model of physical computation that is intended to do three things. First, we wanted to find a way of thinking about building embodied, multi-sensory, intelligent systems — such as we are — that could encompass both theoretical and physical approaches. In particular, we wanted to build a bridge between analogue neuromorphic engineering, as pioneered by Carver Mead [23] and theoretical work by Hava Siegelmann [27] suggesting that recurrent analogue networks (like those in our brain) have super-Turing properties: that is, they could do everything a Turing machine could do and more. Second, we wanted a model that our model had an interface with physics, so that as our knowledge improves, we can simply choose the instantiations of our model that best fits our newfound understanding. Finally, we wanted a model that could be considered to unify the symbolic with the physical: i.e. the Turing-machine with its implementation.

* **0-7803-8566-7/04/$20.00 © 2004 IEEE.**

In this paper, we intend to describe the new model only briefly: full details and motivation can be found in [2]. Instead, we intend to concentrate on the potential utility of the model, how it can change the way we look at anticipatory systems, and what its potential drawbacks are in real application. In particular, we will consider how thinking about computation as physical, rather than informational, may give us a broader picture of what is going on: one that is more consistent with the existence of emergent properties in evolutionary settings, whether external (evolution of species) or internal (evolution of a single brain). On the negative side, we will consider the difficulty — for engineers — of both designing and predicting the behavior of systems using this approach.

2 What is physical computation?

Physical computation is a simple model that can be used to analyze any kind of physical object: from a human to a glass of water to a laptop computer. The assumptions upon which it is based are simple: that any change in the environment (everything except our physical object) will cause some kind of change in the object itself, and *vice versa*. As both object and environment are restricted to obeying the laws of physics, any intelligent or anticipatory behavior on the part of the object may only be attributed to its physical structure, which changes in a seemingly intelligent way based on the energy, matter, fields, etc., that impinge on it. When the object is coupled to the environment in this way, we call it *real physical interaction.*

However, in engineering — and even in science sometimes — we do not consider all inputs and outputs when interpreting a behavior. So, for instance, pushing the key for the letter A in a hard or soft way is considered the same input although, physically, it is completely different. Likewise, the size, brightness, and font of the letter on the display do not alter our interpretation of it. This we call *virtual interaction* because we are interacting with a virtual world, not a real one. Virtual interaction requires that the physical inputs be filtered, or otherwise mediated. Because this mediation takes place, virtual interactions are not constrained in the same way that real interactions are: to an extent, they may be completely arbitrary. Thus, a programmer can sit down and, within the limits of the Turing-machine, do whatever he or she wants. The laws of physics don't apply.

To reconcile the virtual machine with its physical implementation — which is forced to undergo real interaction — we can consider what we call the *complementary function*: literally, this complements the virtual machine, allowing it to perform while the whole thing obeys the laws of physics. In practice, the complementary function is the combination of sensors, actuators, heat-sinks, casings, and so forth: everything that a symbolic or information-based machine lacks.

Our contention is that the better a system is coupled to the environment — the more the changes in the outer world change the structure within — the better will be the adaptation and learning.

3 Benefits of the approach

3.1 Understanding approximation

In our analysis (in [3]) one of the key results was the potential advantage of analogue technologies over their digital (symbolic) counterparts for learning and discrimination in intelligent machines. This is because of the necessary elimination (filtering) of potentially-useful information before it can be assessed for relevance by the intelligence. This process is known as analogue to digital (a/d) conversion. Where a problem is well-understood, such conversion poses few difficulties: engineers simply have to consider what level of approximation will produce results of the desired accuracy.

However, the approximation process in real to virtual conversion will — in poorly understood applications — cause an error whose effect cannot be predicted and therefore cannot be guaranteed to be within specific bounds. This is true whether this conversion is from the real numbers to the rational (continuous environment to Turing-machine input), or a large finite set to a smaller one (discrete-state environment to a strictly-limited-precision digital computer). One of these situations is where the function being performed by the embodied agent is chaotic: in this case, inputs may be arbitrarily close and outputs arbitrarily far. Another is where the error function changes over time in a way that is not known. Since the agent's function is not only determined by it's own learning function (the way it is changed by inputs given its current state) but by what it has experienced, the machine's whole lifetime would have to be predicted in advance in order for the error function to be known. Only in this case (and barring the function becoming chaotic at any point) could the precision required be set in advance with confidence that the error would be within a specified bound.

This issue is somewhat reminiscent of the halting problem, in that the only way to know what happens is to run the system. In this case, instead of trying to find whether or not the algorithm halts in a given time, the designer is trying to discover whether or not the error goes above a given threshold. Of course, it is not possible to know this unless an analogue machine is running at the same time for comparison.

3.2 Taking the system as a whole

Computational neuroscience (the study of biological neural structures using tools more commonly associated with computer science and engineering) often uses information theoretic approaches to understand the brain. In a paper like [1], for instance, Shannon's communication theory (see [26]) is used to improve our understanding of the vision system of a blowfly. The pattern of incoming light is considered to be a *message* from the environment. It travels through the system (modeled as a noisy channel) and, based on the total transformation that occurs, it can be calculated how much information is lost on the way.

This method is very helpful in that it allows us to think about individual neural systems at the microscopic level. As we look at them and see structures that encode, and thus retain, information well, we can learn a great deal about why evolution has designed them this way. In areas of the brain that are chiefly concerned with light, for instance, it is not at all surprising to find structures that, from an information-theoretic standpoint, are particularly good at handling light signals.

However, the brain is not a well-insulated, serial electronic, system, and the problems that it solves are not well-insulated either. For instance, mutual information (see e.g. pp. 40-41 of [10]) often exists between several different sensory modalities, and the information coming into the system through a single channel may have more than one cause. Related to the latter phenomenon, independent component analysis (ICA, see, e.g. [8]) is a well-known means of separating different signals traveling down a single channel into the constituent sound sources. It generally requires a number of different sensors in order to work (say a microphone per source in the audio case). Essentially, ICA is a more thorough version of principle component analysis which works through correlation of the incoming signals. In the case of multi-sensory information, there are artificial systems (based on mutual information, see, e.g. [4]) that can attribute a single cause for information coming through different channels [9]: in this case, audio and visual.

Given these two abilities, both attributed to neural systems, a third naturally arises that has yet to be fully explored: the ability to separate information from several different sensor modalities travelling down channels together. In other words, imagine that the neural response of the eye is somewhat modulated by pressure waves from sound, and the response of the skin and ears are modulated by temperature, and so forth. Each of the minor signals (those that are not of primary interest for a given sensor) would normally be seen as corrupting noise. However, there is no reason in principle why these socalled noise components cannot be extracted out: that this noise might eventually be considered useful signal.

Thus, communication theory — though helping us to understand the way various structures work (and so helping us to think about how to design artificial versions) — limits us to a relatively narrow view of the functionality. We can think, for instance, of the visual system as purely dealing with incoming light as a kind of first-order approximation. Combining this with issues related to the other senses might be considered a second-order approximation. Thus, the dangers inherent in this type of analysis to define global properties of the brain are clear: more kinds of information may be processed than we realize. What we label 'noise' may, in fact, be information added, not taken away.

This is an argument that suggests that the whole will always be greater than the sum of its parts. Not because the analysis of any of the parts are incorrect, but because they cannot all be combined into a single model: in simple terms, the number of simultaneous equations would make the problem intractable. Such limitations are on the model, however, and not on the physical brain itself. This would be true whether the brain were real or artificial.

3.3 Providing footholds for evolution

The complex structures that exist in biology can, as discussed in the last section, be considered in information-theoretic terms. But how did these structures evolve? Mammals with sophisticated light or sound-sensing organs evolved from single-celled organisms. We know that this happened through evolution. But how does a sense organ evolve when there wasn't one there before? By considering the entire physical object as a sensor, just by virtue of the fact that it obeys the laws of physics, this model allows nominal sensor effects caused by physical structures to be exploited.

This can be thought of in two ways. First, you can consider the effect in a single animal. Physical computation, for instance, explains the ability of deaf people to hear through their bodies as the Scottish percussionist Evelyn Glennie has learned to do [11]:

"Evelyn spent a lot of time when she was young (with the help of Ron Forbes her percussion teacher at school) refining her ability to detect vibrations. She would stand with her hands against the classroom wall while Ron played notes on the timpani (timpani produce a lot of vibrations). Eventually Evelyn managed to distinguish the rough pitch of notes by associating where on her body she felt the sound with the sense of perfect pitch she had before losing her hearing. The low sounds she feels mainly in her legs and feet and high sounds might be particular places on her face, neck and chest."

Thus, physical features that are not designed as sensors can nevertheless be exploited as such. This is only possible, because — in the physical computation model — no explicit measurement path is required for these non-sensors. If it were, in an engineered system there would almost certainly be no such path. The kinds of vibrations involved in hearing, even at low frequency, would be unlikely to be considered of primary importance in an information-theoretic analysis of a leg: they would be considered too marginal (in terms of purpose) and too small (in terms of signal) to be worth collecting. Nevertheless, the brain can make use of the information that the legs provide as hearing organs: and in Glennie's case it actually does.

Such small effects can also be exploited by changing the entire morphology of a species. For instance, there is much debate about the reason for the shape of the hammerhead shark [19]. Unlike an ordinary shark, this animal has an oblong-shaped head much wider than its body — wide ends at the front and neck — with the eyes on the front two corners of the oblong. There are various reasons why biologists believe that this shape evolved. First, there is an obvious sensory advantage in that binocular vision (depth perception) improves with the separation of the eyes. Another explanation is that this head shape has hydrodynamic advantages including increased stability while turning. Finally, sharks have electrosensory pores that assist them in catching prey. The wider head means there can be a greater number of these pores for the same density.

With the physical computation model, evolution could select for any of these properties (i.e. they would supply some kind of advantage that would allow the animal to survive and procreate better than its neighbors) because any physical changes caused by mutation, no matter how small, or how seemingly unrelated to either sensing or locomotion, could be exploited. For instance, a slight widening of the head produced by evolutionary accident, could have improved the survival of the hammerhead shark precisely because all three abilities (electro-sensing, depth perception, dynamical stability) were enhanced. Similarly, we can imagine that a hair-like structure — one that happens to vibrate at a frequency of interest and happens to be located near some kind of nerve ending — could eventually evolve into a hearing organ.

Another example of this might be the vision system of the sunfish [7]. This creature lives in areas where the water is relatively cloudy, causing much diffusion. However, the morphology of the eyes allows polarization-difference imaging: where images are acquired of each polarization (due to the shape and orientation of the photoreceptor anatomy) and then subtracted from each other. This removes most of the diffuse light, leaving a useful image for the fish to interpret. Again, in terms of evolution, the physical computation here is crucial.

From an engineering perspective, development of the physical computation approach could eventually allow the use of genetic algorithms (GAs) to select for success in a much more powerful way. Though evolutionary algorithms initially focused software agents in artificial environments [20], the idea of engineering hardware this way has been around for some time: initially working with simulated physics [14]. Later, from the same group, Thompson [29] showed that small hardware effects can be exploited by GAs to solve problems in ways that would not have been possible using conventional circuit design techniques or software-based design packages. Also, Sims [28] demonstrated that it is possible to engineer brain and body together, with physical and informational attributes evolving together in a competitive genetic selection process (he used a virtual physical model of the environment to do this). By combining all three approaches, we may be able to evolve embodied agents that can multi-task, using their physical attributes to best effect in terms of sensing, actuating and processing.

3.4 Processing the unmeasurable/computing the incomputable

If the most optimistic scenarios for analogue computation are applicable (which depends on answers, as yet unforthcoming, from physics), physical computation has advantages over Turing machines. Literally, this means that analogue could compute the incomputable, where computable is defined as computable on any discrete-state machine. Also, real-interaction-based physical computation does not require the making of measurements in the way that virtual computation does. Consequently, physical parameters that could never be known, even in principle, can nevertheless contribute to the physical processing taking place. In an application where sensitivity to the environment is an advantage, this is an important property.

It is important to stress at this point that this advantage cannot be easily harnessed (if at all) for conventional computing applications like word-processing or database management. The advantage is derived by two things that make it incompatible with conventional sensor-processor-actuator systems. First, the ambiguity of the sensors (the whole device is a sensor) means that all physical influences can be taken account, but their sources cannot all be separated. For instance: we are subject to gravity, and any small changes in its force on us will affect the way we behave. However, that does not mean we know about all the different masses that have combined to create this particular force. Such ambiguity — while acceptable for applications where the macroscopic physical behavior of a system is what's important — cannot supply the information required for conventional informational tasks.

3.5 The analogue shell and nested virtual machines

Chris Toumazou [30] has talked about the analogue shell of a device: the outer sensors and actuators that allow the inner digital electronics to interact with the outside world. The physical computational model fits well with this metaphor. Though Toumazou does not work on artificial intelligence, the systems that he designs — which include mobile phones and biomimetic sensors — fit well into a hierarchy that is suggested by the model. There are three obvious examples we can discuss: the digital virtual machine, the analogue virtual machine, and the analogue real machine. The latter should be well-understood by now as it has been the subject of much of the discussion so far. However, the other two cases are worth further consideration.

For virtual interaction, the first extra cost is that of shielding: in order to be selective about the information acquired (information implies selection) the object must be resistant to the other physical influences around it. Second, the object must interpret the information coming in through its sensors (and information it is, in this context) and assign it some kind of meaning (even if this is purely in the context of which bit comes before which) before it can act: this tends to mean dramatically reducing the types of information that can be considered and the amount of information that can travel through those channels. For instance, we might consider visual information (which is really a kind of map of our local electro-magnetic field) important, then reject certain wavelengths for consideration because their source, for instance, is ambiguous. Without the relevant knowledge the signal, or information, becomes meaningless. We can understand these shielding and interpretive functions as being part of the complimentary function.

In fact, the electronics industry is entirely structured on the basis of this duality. On the one hand, designers treat resistors, transistors, capacitors, etc., in terms of mathematical functions that they will perform on an expected (in terms of dynamic range, frequency, etc.) signal from a known source. Software engineers do not even consider the implementation: they simply assume that the high-level functions they need will be permitted on the machine they use. At the lowest level, a completely different set of designers look at the physics of the systems that will make both of these higher levels work efficiently. You can think of this as the analogue shell (complimentary function) with a first virtual layer of the designed electronics and the second virtual implementation of the software. These levels have clear demarcations.

What is potentially interesting about an all-analogue implementation is that there may be many levels of virtual computation rather than the two I've just described for digital. In an information-theoretic sense, we can think of many layers of analogue circuits, each one throwing away more information from the outside world in favour of those signals that are deemed (by virtue of, for example, neural network learning) important. It could be argued that the visual system, as described, for example, by Hubel and Wiesel [18] does something like this. Rather than encode all information, retinal processing elements seem to break down images into lines of different orientations, and so forth. This concept of filtering is also used in pattern matching through Fourier-transform-based image correlation. Instead of comparing the images, elements of the image are filtered out and considered as a collection of elements. The more closely the feature collection of one image matches another, the more closely the images themselves are expected to match.

We can consider adaptive multi-level filtering as a kind of nested virtual machine: where an agent takes as much information into the system as is feasible given the sensors and then deciding what to throw away based on experience. Work by Geoff Hinton (see e.g. [17]) points to how this might be done: he devised a network that creates its own ideal set of features to discriminate between relevant objects. In a fully-adaptive system, not only the features (and how they are compared) might start by being unknown, but also the very question of what constitutes a relevant object. These might then be used as primitives for higher-level abstractions. In other words, as we penetrate deeper — past the analogue shell to the core — the computation looks increasingly symbolic and representational. Crucially, however, this is a much more flexible system than one that requires immediate (and irrevocable) analogue-to-digital conversion. And yet it is compatible with conventional artificial intelligence in that it allows for a pseudo-symbolic core to emerge. Thus, conventional algorithms that people have devised to explain intelligent behavior in people may be considered accurate: the question is how to implement them within a nested virtual system implemented on an analogue machine and without the usual interface of software.

3.6 Intelligence is more than thinking

The concept of the complimentary function may be helpful also when considering the mind-body problem. Our bodies not only provide our conscious mind (the intelligence function in this case) with sensors and actuators, but also the un/subconscious mind which is not normally considered in AI. Other bodily systems — the limbic system, the spinal chord, even basic organs like the heart and lungs — all contribute to our behavior to a greater or lesser extent. Using the physical computation model, we can see that what many people consider to be artificial intelligence, particularly purely software-based approaches, concerns only the virtual machine.

Striking a better balance may be said to be the concern of the neuromorphic engineers: their work often goes much further, blurring the interaction boundaries. Projects like Harrison and Koch's all-analogue fly vision system and robot controller [13], Hasslacher and Tilden’s analogue walking/light sensitive robots based on the nervous systems of small animals [15,16], or Lewis’s bipedal robots based on central pattern generators [21,22], all have in common that there cannot be said to be a clear line between sensor, processor, and actuator. Indeed, there cannot be said to be a clear line between software and hardware.

Pushing this to an extreme, one could argue that full understanding of real interaction, virtual interaction, and the complementary function might go some way to addressing Brooks’ recent question about the relationship between matter and life [6]. In essence, we may just need to go further to try to understand a whole brain and body in order to understand how a creature works, rather than considering mind or brain processes alone. Considering this issue in the 1940s, Schrödinger suggested that the secret to the difference between what is and isn't alive lies in the nature of its reproduction and the coding of its construction in genetic code [25]. More recently, Pattee has suggested the same thing, and pointed to the importance of self-knowledge (in the sense of what is, and isn't, the self: see e.g. [24], and the importance of physical symbols (sets of physical states that essentially 'mean' the same thing). Since the integrity of both the genetic code and the self (whether at the level of single-celled organisms or animals with a large and complicated structure) is based on physics and chemistry, a more material, less abstract, understanding of these issues may be helpful.

4 The difficulties of the approach

Although it may provide us with new conceptual avenues down which to pursue embodied artificial intelligence, the adoption of this approach has three major problems associated with it.

4.1 Design and fabrication

Complicated physical systems are notoriously difficult to design and can be expensive to build. According to industry experts, too few analogue engineers are being trained [5]. There may be several reasons for this. First, the mathematics involved in analogue electronic design is extremely difficult. Second, even if a designer (particularly at the academic level) has the right analogue skills, he or she may have to wait a long time (weeks) for designs to come back from a VLSI foundry. Third, a relatively small design mistake or miscommunication could render the manufactured chip useless. Finally, the process is as expensive as it is time-consuming. A similar argument could be made for the work of mechanical engineers and optical engineers. By contrast, programming (software engineering) is relatively easy, inexpensive, and can be done by a single researcher without assistance or delay.

Though research progresses towards simplifying the design process of optical, mechanical and analogue systems (see, e.g., [12]) , the level of difficulty is currently still vast in comparison to the easy of sitting down to programme a computer.

4.2 The unpredictable black box

There is another problem that comes through working with any machine that has an interface to the real world and which is not well understood: unpredictable behavior. In a desk- or laptop computer this is not a problem because, in most circumstances, it poses no safety issues. In an embodied intelligent agent — whether one designed to have the same physical abilities as a human, or an autonomous aircraft or power plant — this is not the case. The larger and more potentially destructive a machine is, the more predictable we generally want it to be.

While any embodied AI is likely to be complicated and therefore difficult to predict, a neural-network-based analogue AI is likely to be much more so. This is because the behavior of the machine is determined by its history alone, and not by algorithms designed by software engineers. In a sense, this could make the machines more human, as people are also considered to be black boxes, conditioned by their experiences. On the other hand, it would not seem to be worth risking such an unpredictable approach for straightforward engineering applications or those in which it is easier to simply have a human operator in control.

4.3 Implementation, not explanation

The physical computational approach is intended to allow researchers to think about the most efficient possible solution to a problem in a given physical environment. However, like other approaches such as genetic algorithms (natural selection), it doesn't seek to explain why a particular solution is better than another. In fact, the parallel with natural selection is very close. Both it and physical computation explain the mechanism by which intelligent evolution takes place in an environment, not what form it should take. Thus, conventional models such as those based on information theory have two advantages. First, they can explain why a solution is a good one. Second, if they are sufficiently small problems that they can be adequately computed on digital hardware and with a reasonable analogue to digital conversion, such models are much more cheaply and easily implemented than the physical computational approach would be.

This ease of implementation, along with the ability to devise and then test hypotheses about the way intelligence works, makes the conventional approach very attractive. It is only where efficiency and sensitivity (either physical, informational or both) are crucial that the physical computational approach becomes worth the extra effort.

5 Conclusion

Embodied human-like artificial intelligence is a particularly good example of an application where first-order information theory would not seem to provide a sufficient model of what's going on. Likewise, looking at all possible uses of an information channel is not an efficient way of considering a system. A machine's embodied nature, encompassing its brain, means it has a myriad of ways of interacting with the world. Looking at a few of these will only tell a small part of the story. They may provide a helpful explanation, but will not be able to explain global behavior.

The physical computation model is useful for such demanding applications because it reminds us that — quantum mechanics aside — noise doesn't exist in the real world: everything is signal until we start to put an interpretation on what is important and what isn't. In order to build systems that can fully exploit their interactions with the environment, we must avoid restricting the information that they can take in: instead, we should allow them to develop, evolve, their own filters as their priorities emerge.

6 References

[1] P. Abshire and A. Andreou, "A communication channel model for information transmission in the blowfly photoreceptor", *BioSystems,* Vol. 62, pp. 113–133, 2001.

[2] S. Bains, "Intelligence as Physical Computation", *AISB Journal,* Vol. 1, pp. 225–240, 2003.

[3] S. Bains, *Physical computation and embodied artificial intelligence,* Doctor of Philosophy Thesis, The Open University, 2004.

[4] S. Becker, "Mutual information maximization: models of cortical self-organization", *Network: Computation in Neural Systems,* Vol. 7, pp. 7–31, 1996.

[5] S. Briggs, "ON Semiconductor Provides $1.5 Million to Support Analog Electronics Program", *ECE News,* Fall, p. 1, 1999.

[6] R. A. Brooks, "The relationship between matter and life", *Nature,* Vol. 409, pp. 409–411, 18 Jan 2001.

[7] D. A. Cameron and E. N. Pugh, "Double cones as a basis for a new type of polarization vision in vertebrates", *Nature,* Vol. 353, pp. 161–164, 12 Sep 1991.

[8] P. Comon, "Independent component analysis, A new concept?", *Signal Processing,* Vol. 36, pp. 287–314, 1994.

[9] V. R. de Sa and D. H. Ballard, "Category Learning Through Multimodality Sensing", *Neural Computation,* Vol. 10, pp. 1097–1117, 1998.

[10] G. Deco and B. Schürmann, *Information Dynamics: Foundations and Applications,* Springer-Verlag, *2001.*

[11] E. Glennie, "Evelyn's Hearing", on her website at http://www.evelyn.co.uk/hearing.htm, 1996.

[12] D. Haigh, "Systematic synthesis method for analogue circuits (Parts I-III)", Proc. IEEE International Symposium on Circuits and Systems, 2004.

[13] R. R. Harrison and C. Koch, "A silicon implementation of the fly's optomotor control system", *Neural Computation,* Vol. 12, pp. 2291–2304, 2000.

[14] I. Harvey, P. Husbands, and D. Cliff, "Seeing the light: Artificial evolution, real vision", *From Animals to Animats 3: Proceedings of the 3rd International Conference on the Simulation of Adaptive Behavior,* MIT Press, pp. 392–401, 1994.

[15] B. Hasslacher and M. W. Tilden, "Theoretical Foundations for Nervous Networks", *Applied Nonlinear Dynamics and Stochastic Systems Near the Millenium,* American Institute of Physics, pp. 179–184, 1997.

[16] B. Hasslacher and M. W. Tilden, "Living Machines", *Robotics and Autonomous Systems,* Vol. 15, pp. 143–169, 1995.

[17] G. E. Hinton, "Training Products of Experts by Minimizing Contrastive Divergence", *Technical Report GCNU TR 2000-004,* Gatsby Computational Neuroscience Unit, University College London, 2000.

[18] D. Hubel and T. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex", *Journal of Physiology (London),* Vol. 160, pp. 106–154, 1962.

[19] S. M. Kajiura, "Head morphology and electrosensory pore distribution of carcharhinid and sphyrnid sharks", *Environmental Biology of Fishes,* Vol. 61, pp. 125–133, 2001.

[20] S. Levy, *Artificial Life: The Quest for a New Creation,* Pantheon Books, 1992.

[21] A. M. Lewis, M. J. Hartmann, R. Etienne-Cummings, and A. H. Cohen, "Control of a robot leg with an adaptive aVLSI CPG chip", *Neurocomputing,* Vol. 38-40, pp. 1409–1421, 2001.

[22] A. M. Lewis and L. S. Simó, "Elegant Stepping: A Model of Visually Triggered Gait Adaptation", *Connection Science,* Vol. 11, pp. 331–344, 1999.

[23] C. A. Mead, *Analog VLSI and Neural Systems,* Addison Wesley, 1989.

[24] H. H. Pattee, "The physics of symbols: bridging the epistemic cut", *BioSystems,* Vol. 60, pp. 5–21, 2001.

[25] E. Schrödinger, *What is Life,* Cambridge University Press, *1944.*

[26] C. E. Shannon and W. Weaver, *The mathematical theory of communication,* University of Illinois Press, *1949.*

[27] H. T. Siegelmann, *Neural Networks and Analogue Computation: Beyond the Turing Limit,* Birkhäuser, Boston, *1999.*

[28] K. Sims, "Evolving 3D morphology and behavior by competition", *Artificial Life IV,* MIT Press, pp. 28-39 1994.

[29] A. Thompson, "Explorations in Design Space: Unconventional Electronics Design Through Artificial Evolution", *IEEE Transactions on Evolutionary Computation,* Vol. 3, pp. 167–196, 1999.

[30] C. Toumazou, "The Bionic Man", Royal Society, 9 Oct 2003.