CITE AS IN IEEE PROCEEDINGS SYSTEMS, MAN & CYBERNETICS (CHICAGO), IEEE, NEW YORK, 1992.

Neural Networks & the Human Mind: New Mathematics Fits Humanistic Insight

Paul J. Werbos
Room 1151, National Science Foundation*
Washington, D.C., USA 20550

Abstract **The past two years have seen substantial progress in artificial neural networks(ANNs). New ANN control designs, combining reinforcement learning and generalized backpropagation[1,2], have demonstrated success on large-scale real-world problems which could not be solved by earlier designs, neural or nonneural. There now exists a ladder of designs, rising up from simple designs (of limited capability, but a good starting point), through to complex proven designs (which give more power and flexibility after they are mastered), up to new untested designs and ideas which should be able to replicate and explain human intelligence at the highest possible level.**

After a brief review of neurocontrol, and an explanation of reinforcement learning, this paper asks what implications these designs have for our understanding of the human mind. It argues that this new mathematics is fully compatible with older deep insights into the human mind, due to humanistic thinkers East and West. One may therefore hope that this new view of the human mind may be of value in unifying important strands of human culture.

I. A Review of Neurocontrol

From an engineering point of view, the human brain is simply a computer, an information processing system. The function of any computer, as a whole system, is to compute its outputs. The outputs of the brain are control signals to muscles and glands. Therefore the brain as a whole system is a neurocontroller (a neural net control system) [3]. To understand the brain in functional, mathematical terms we should therefore focus on the subject of neurocontrol.

ANNs have been useful in four kinds of control task: (1) as subsystems of larger systems, where the controller itself is not an ANN; (2) in "cloning" applications, to copy what an expert does (unlike conventional expert systems which copy what an expert says); (3) in tracking applications, such as holding a plant to a fixed setpoint or making it follow a pre-specified reference trajectory; (4) in reinforcement learning or optimization over time. The first three of these four clearly have nothing to do with human intelligence.

Some biologists once argued that lower-level functions in the brain, such as the control of arm movement by the cerebellum, might be based on a simple tracking system; however, Kawato et al[4] have done impressive experiments proving that these parts of the brain actually optimize movements over time. They do perform tracking, but only as part of an optimization task. A few authors have questioned these conclusions, but they still hold up quite well.

In brief, the reinforcement learning or optimization designs are the only designs of value in understanding the human mind. They are the only designs capable of meaningful planning or foresight. They are also the designs which have led to the most exciting real-world applications in recent years. Therefore, this paper will focus exclusively on them. To learn about other designs, stability, etc., see [1].

Within the field of optimization over time, two classes of design have proven useful in practice: (1) direct optimization, using generalized backpropagation to calculate derivatives of utility or performance or cost; (2) adaptive critic designs, which approximate dynamic programming.

Direct optimization based on the backpropagation of utility was first proposed in 1974[2]. By 1988, there were four significant working examples[4], including two model robot controllers, one controller of a simulated truck-backer-upper, and a US Department of Energy official model of the natural gas industry[5]. By now, dozens of examples have appeared, including a Model-Predictive Control scheme now used to improve efficiency and reduce waste in profit-making chemical process plants, and an optimal tracking scheme of Narendra[1]. However, none of these designs are plausible as models of the brain. Some of the designs require calculations backwards through time; others require huge computational costs for large problems; and others simply cut off the key calculations which account for the long-term effect of present actions. These designs have great value in engineering (including an ability to reduce pollution while saving money), but they are not directly relevant to understanding the human mind; therefore, I will not discuss them here. Optimization methods derived from static function maximization are even less relevant here.

II. Reinforcement Learning and Adaptive Critics in General

Adaptive critic systems can perform "reinforcement learning." In reinforcement learning, an ANN system receives a vector of sensor inputs $\underline{X}(t)$. It outputs a vector of control signals or actions $\underline{u}(t)$. Then it receives a "reward" or "punishment" U(t). In reinforcement learning, the system must somehow learn to output actions $\underline{u}(t)$ which maximize future rewards U, summed across all future time, from U(t) to $U(\infty)$.

*The views herein are certainly not official views of NSF.

Many thinkers throughout history have argued that the human mind is a reinforcement learning system. The behaviorist psychologist Skinner built his theory around "primary reinforcement," which is exactly like U(t). Many ancient Greeks argued that the human mind maximizes "hedony" or "pleasure versus pain." Aristotle argued that the "telos" or organizing principle of the mind is simply the maximization of "happiness." John Stuart Mill and Von Neumann claimed that we maximize "utility."

In psychology, it is important to ask what variables actually enter into the U(t) which we maximize. In biology, it is clear that there are fixed centers, such as the hypothalamus and epithalamus[3,6], which generate primary reinforcement. From a Darwinian perspective, E.O.Wilson has developed deep insights[7] which suggest that U(t) -- for a human individual -- would represent some kind of weighted sum of the well-being of people whom the individual cares about. Wilson's insights fit well with ancient Confucian notions, which stress that well-adjusted humans give great priority to family values; however, they also allow for tribal feelings and for the possibility of explicit bargains[8] or social contracts which permit cooperation at higher levels across different families or tribes, especially in "pioneer" environments which may permit growth. The Darwinian approach also suggests that humans do not really "discount" the intrinsic importance of U in the far future; this fits with the Confucian notion that wise and well-adjusted humans think ahead for many generations. These ideas (similar to some ideas from Christianity and Islam) can be reconciled with the requirements of a market economy, but they still have major policy implications[9].

Critics of reinforcement learning argue that humans often make suboptimal decisions, and often put great energy into activities like exploration which do not lead to direct rewards. However, working, real-world reinforcement learning systems share these characteristics[1,6]. In Darwinian evolution, one would expect nature to converge on systems which do the best possible job of maximizing some measure of success, subject to the constraint of what is actually possible for a physical learning system; in neurocontrol, we try to do the same. In addition, later parts of this paper will suggest that the true U(t) of human beings may favor exploration more than simplified versions of it would.

In real-world engineering applications[1], we use adaptive critics to maximize $U(\underline{X}(t))$ rather than U(t), because of the great efficiency we can get by exploiting knowledge of what we are trying to maximize. One would expect a similar arrangement in biological systems[6].

III. Backpropagation, Freud and Advanced Adaptive Critics

Adaptive critic systems try to approximate dynamic programming.

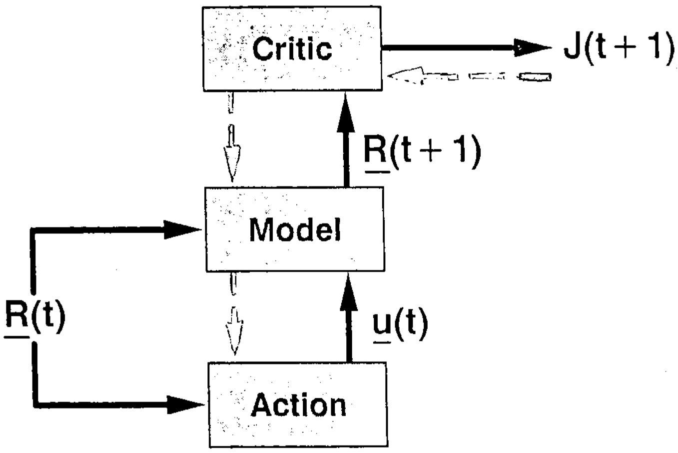
Dynamic programming is the only exact and efficient technique for maximizing U across future time, in the general case, where noise and nonlinearity may be present. To use dynamic programming, we proceed as follows. First of all, we, the users, must supply a utility function $U(\underline{X})$ and a model of the external world that we want to optimize. Then dynamic programming tells us how to solve for another function, $J(\underline{X})$, which can be used in choosing $\underline{u}$. $J(\underline{X})$ may be thought of as a secondary or strategic utility function. It represents a kind of strategic assessment of any possible situation $\underline{X}$. The basic theorem in dynamic programming is as follows: by picking $\underline{u}(t)$ at each time t so as to maximize $J(\underline{X}(t+1))$, we automatically maximize the sum of U over all future times. Dynamic programming converts a difficult problem in long-term optimization into a straightforward problem in short-term function maximization.

Dynamic programming becomes impossibly expensive to use in its pure form, even for medium-sized problems. Therefore, we cannot hope to solve such problems exactly; we will never play a perfect game of chess, and we should not expect our ANNs to do so either. To approximate dynamic programming, we can try to adapt "Critic" networks -- networks which try to approximate the J function, or something similar.

These functions U and J show up very clearly in tasks like playing chess. In formal chess, the goal is to win; thus $U(\underline{X})$ is zero except at the end of the game, where it is +1 for a win and -1 for a loss. But beginning chess players often learn an old rule of thumb to measure their progress before the end of the game. They count 9 points for a queen, 5 for each castle, and so on. This is a simple approximation to J. Better chess players learn to place value on holding the center, etc. Some analysts argue that the best human chess players really look ahead only one move; their apparent foresight may be due to a very careful, in-depth strategic assessment (J) of the near-term alternatives. In other words, their success may be due to a better approximation to the true function J.

In human psychology, the output of a Critic network corresponds exactly to Skinner's idea of "secondary reinforcement" -- a learned reinforcement. Biologists have shown that the limbic system of the brain[1,6] generates such reinforcement; the limbic system is often described as the "emotional" system of the brain. Just as U includes inborn feelings like pleasure and pain, J represents learned responses like hope and fear. When I argue that the human mind is an adaptive critic system, I am simply saying that the human mind is governed by hopes and fears which it learns through experience. The development of a more accurate J function is crucial to the development and intelligence of a human being; this corresponds to the Confucian idea that balanced judgement (J) is a crucial basic faculty for humans to develop.

Adaptive critic ANNs were first implemented by Widrow in 1973, and improved upon in a famous 1983 paper by Barto, Sutton and Anderson; however, neither of these designs can handle truly large, brain-like control problems[1,4]. The figure below illustrates an advanced adaptive critic design which I proposed briefly in 1977[10]



Backpropagated Adaptive Critic

and at greater length in 1981[11]. In this design, the overall system is made up of three ANN components. One component, the Action network, actually generates the vector $\mathbf{u}(t)$. Another network, the Model networks, is adapted to predict or explain the external world. The Critic network, at the top, is needed in adapting the Action network. The upwards arrows show how to predict $J(t+1)$ for any given vector of actions $\mathbf{u}(t)$. See [1] for all the mathematical details. The vector $\mathbf{R}(t)$ represents an image of reality at time t, based on $\mathbf{X}(t)$ plus other information[1].

This figure also illustrates how the Action network can be adapted through the backwards, broken arrows, which represent the derivatives of J. These derivatives are important mathematically, but they also have great significance to psychology and economics.

In studying human values and fears, it is not enough to study functions like $U(\mathbf{X})$ and $J(\mathbf{X})$ which represent global measures of happiness, etc. Humans also place values on specific objects and specific variables, values which are crucial in governing our behavior. For example, in economic systems, there are values or prices put on specific goods, which are crucial to the efficiency of such systems.

In economics, the value of a specific good, X_i, is defined by its "marginal utility." The marginal utility of X_i is defined as the proportionate increase in $U(\mathbf{X})$ which would result from a small increase in X_i. For example, the value of a peanut to you equals the increase in utility which would result from your consuming one additional peanut. Mathematically, the "marginal utility of X_i" is just a synonym for the derivative of $U(\mathbf{X})$ with respect to X_i. The value of a good to society as a whole over the long-term future is usually discussed in terms of "Lagrange multipliers," λ_i, which are actually just the derivatives of $J(\mathbf{X})$.

Years ago, Sigmund Freud made a persistent effort to understand the underlying laws of human learning, laws which could explain the deep and rich experience he acquired regarding human thought over many decades. The theory he arrived at was essentially a neural network theory, motivated by his earlier study of neurophysiology in medical school. He argued that human behavior and human feelings are dominated by "psychic energy," by a system of emotional charge or values placed on specific objects. (An "object," in his terminology, could be an "object of affection," like a loved one.) He proposed that humans "first" build up knowledge of cause and effect, through experience; thus, for example, we may learn that object A at time t is associated with object B at time t+1. He proposed that such causal associations are represented, in the brain, by a forwards connection or synapse from the neuron representing A to the neuron representing B. Then came his crucial insight: he proposed that "psychic energy" flows backwards from cell B to cell A, with a connection strength proportional to the forwards association from A to B. At the time, he proposed that the backwards flow of psychic energy is represented in the brain by chemical flows inside of cells, running backwards compared to the usual flows of electrochemical information in the cell membranes.

Backpropagation -- the most widely used and successful algorithm in the ANN literature -- was originally developed as part of a conscious effort to translate Freud's theory into working mathematics[6]. The figure above illustrates the original idea. In this figure, the dashed lines represent the derivatives of J, which represent the "psychic energy." A backwards flow of calculations -- matching Freud's idea exactly -- is used to work out the derivatives of J with respect to $\mathbf{R}(t+1)$, then the derivatives of J with respect to $\mathbf{R}(t)$, then the derivatives of J with respect to $\mathbf{u}(t)$, and then -- finally -- the derivatives of J with respect to the weights in the Action network. These weights are adapted in response to those derivatives. This arrangement exactly fits the prescription of dynamic programming, which tells us to pick $\mathbf{u}(t)$ so as to maximize $J(t+1)$.

Again, this idea was published in 1981[11]. That same paper also discussed backpropagation in general terms (as in [2]), and proposed the use of differentiable model neurons to permit the use of backpropagation in adapting ANNs. It also discussed links to the brain. Over the years, I have also developed other ways to use backpropagation in adaptive critic systems[1].

As recently as 1990[4], there were no published working examples of adaptive critic systems really exploiting backpropagation. (There had been delays, of course, even in simpler uses of backpropagation.) As of now, there are at least four, of which two are practical real-world systems with multiple applications. BehavHeuristics of College Park in Maryland has used a three network design, similar to the figure above, to perform optimal seat allocation and scheduling, so as to maximize airline profits; this system has passed large-scale tests, and USAir has signed a contract to apply it to optimize their global network of flights. The system developed by White and Sofge at McDonnell-Douglas and Neurodyne has been used to make high-quality composite parts in a continuous production system, to build an F-15 controller able to adapt in two seconds to major changes in the airplane, and to develop a prototype thermal controller for an airplane (NASP) designed to reach escape velocity[1]. New applications by AAC of Tennessee are also important, but the reinforcement learning aspects are not yet published.

Some researchers argue that backpropagation has not been found yet in the brain. However, there are fundamental reasons why it would be difficult or impossible to build brain-like intelligence without such backwards flows of information[6]. There are many biological mechanisms which could implement such a flow, but no one has looked at these systems very carefully yet[6]. This past year, it has been discovered that nitric oxide does act as a "backwards transmitter." Karl Pribram, in conversation, has suggested that his classical experiments on the limbic system may already demonstrate such a backwards flow of information, inconsistent with classical neuron dogmas. Links to the cerebellum are discussed in [12]. There is a great need for new research to explore these biological issues, perhaps using new instrumentation.

IV. Language, Planning, Dreaming and Confucian Ethics

Within the Confucian scheme, the first and deepest imperative is to be true to oneself. The most fundamental virtue is "integrity" -- a state in which you tell the truth to yourself, and not just to others. However, to be true to yourself, you must make some effort to understand yourself in the most accurate way possible. This is one motivation for studying neural networks and other ways of studying the human mind.

If human beings are born as reinforcement learning systems, doing their best to maximize their personal sense of what seems good (U), how could they be capable of anything but integrity? Why should integrity be such a big issue for human beings, but not for other animals?

The answer comes from the role of language.

The adaptive critic systems now in use cannot replicate or explain language, because they are not even capable of high-order planning or "chunking," which is a prerequisite to language. However, recent research suggests that true planning and chunking emerge naturally and automatically from adaptive critic systems which use appropriate types of ANN as Critic networks. (See chapters 10 and 13 of [1].) Biological research suggests[1,12] that the brains of all mammals do use the appropriate types of network.

Another prerequisite to language is "dreaming." In any kind of adaptive critic system, there is much to be gained by exploring or simulating possible future states of reality **R** which have not been experienced yet as actual states. This point was stressed in [13], and demonstrated graphically by Sutton in his "Dyna" simulations[4]. Current sleep research appears to be fully consistent with this interpretation of dreaming as a kind of offline simulation[14]. In classical dreaming, of course, one would adapt the Critic network and the Action network so as to handle the hypothetical states **R**; one would probably not adapt the Model network, because the Model network is the system used to generate the simulations in the first place. There is reason to believe that all types of mammal are capable of dreaming.

Even if we add dreaming and an appropriate form of Critic, the designs now in use would still not have a true ability to learn to use language. They would still adapt their components based on their own individual experience. They could respond to words as sounds, in a highly intelligent way, but this is not the same as learning from the experience of those who are talking to you. There would be a tremendous evolutionary advantage to learning from the experience of other animals, as if that experience were one's own experience. I have argued [6] that humans have evolved such a unique ability, very recently in evolutionary time, and therefore very imperfectly. This ability may have begun with a kind of trance-like state, similar to dreaming, which allowed members of human groups to experience vicariously the memories of other group members returning from a hunt and dancing out their memories; unlike dreaming, this state would allow adaptation of the Model networks. The ability to use language led to another skill -- symbolic reasoning -- only in historical times.

If the ability to use language is new, it is not surprising that humans face transitional problems or instabilities in using language. These problems force us to learn certain skills through experience -- such as the skill of integrity -- without the biological support we have in learning older skills like vision. Still, the natural equilibrium of the system -- the natural equilibrium state of a well-adjusted person -- would involve a balance or match between the verbal side of our thinking and the nonverbal side; it would maintain the Confucian ideal of integrity. In matters concerning U -- as in matters concerning visual perception -- we would constantly be aware of what is coming to us from the nonverbal level, constantly ready to articulate those inputs as accurately as possible, and constantly ready to analyze what we see and feel by making full use of verbal and nonverbal skills, on an integrated basis. There is a close analogy here between the Confucian ideal of integrity and the Freudian ideal of sanity.

Social forces often encourage subtle forms of dishonesty and dogmatism which make it more difficult for humans to learn effective symbolic reasoning and "integrity." Primitive corporate cultures (like animal societies) often do not value honest symbolic reasoning, or mental productivity in general [9]. Philosophers like Nietzsche and Ayn Rand have presented desperate (and valid) pleas for individual humans to overcome such phenomena, in order to achieve greater integrity. When the doors of perception and sensitivity inside the self have been locked and sealed, the emotional violence described by Nietzsche may be necessary at times to open them up; however, the reintegrated self need not be violent or antisocial. Irrational fears to that effect are one of the many factors one must overcome, in gradually developing more productive corporate cultures throughout the world, cultures which encourage individuals to live up to their full potential.

V. Beyond the Brain: How Far Can the Mathematics Go?

The preceding part of this paper is fully consistent with the classical materialistic idea that the human brain and the human mind are essentially identical. ANN research -- like modern neurophysiology -- suggests very strongly that the brain itself can generate virtually all aspects of human intelligence which are generally agreed upon. By Occam's Razor, this would tend to suggest that we should throw out classical views of the mind which try to go beyond the brain. The logic of that viewpoint is unquestionable. The people who agree with that viewpoint should stop reading here, having found a mathematical framework which is fully capable of sustaining that view.

Nevertheless, the most creative ANN researchers (in my experience) do not all subscribe to classical materialism. So far as I can tell, their views seem to be similar in character to the views of the four most famous physicists of this century -- Einstein, DeBroglie, Schrodinger and Heisenberg. Like those four, they have a great deal of diversity. Like them, a majority seem to be open to the possibility that the human mind may possess capabilities or attributes which go beyond those of the known brain. This is paradoxical, since these are precisely the people most conscious of the points in the preceding paragraph. They are not the kind of people who believe things simply because their parents did. One might speculate that this surprising situation is due to these people observing capabilities or phenomena in their own unusually capable minds which they find it difficult to explain. Or, as creative people, they may open themselves up to connections and feelings which more constrained technicians and apparatchiks may tend to block out. Or perhaps this is just a way of distancing themselves emotionally from the current state of what is known. In any event, a balanced and complete account of humanistic views of the mind must make some allowance for ideas and possibilities which are taken seriously by a significant portion of humanity.

Could the mathematics of neurocontrol still be useful in describing the human mind, if in fact the human mind were larger than the human brain?

There are two reasons to be optimistic here: (1) if the human mind were larger than the brain, then mathematical insights would be all the more important, to help minimize the incredible confusion and chaos which this viewpoint would otherwise permit; (2) the mathematical issues which led to our designs are not limited to systems made up of wet neurons and atoms; they should be applicable even to systems built up from other kinds of devices.

As an example, there is a large, old literature suggesting that individual human beings have some kind of symbiotic relation with or membership in some kind of large, collective intelligence. Hinduism and Buddhism have promoted the idea that we are all part of some great Mind. People like Teilhard de Chardin and Lovelock have promoted similar ideas in the West in recent decades. Carl Jung -- borrowing heavily from Buddhism -- has published very extensive studies of human psychology, supporting his notion of a "collective unconscious." Even the New Testament talks about a "true vine" which (to me) sounds like a neural network -- a vast, invisible neural network, held together by invisible connections between people of all sorts across the world.

If those ideas were correct, I would predict that backpropagation in some form would have to play a crucial role in organizing this collective intelligence. Based on casual observation, I would claim that this collective intelligence would have to be in a relatively young or immature state; therefore, to understand or assist this intelligence, we would need to understand the forces permitting or encouraging greater maturation. We would need to reconsider Freud's discussion of the evolution of ego through appropriate flows of raw psychic energy, as it would apply within a larger intelligent system.

In the developmental process, the development of critical new gestalts or variables R_i plays a central role in channeling psychic energy in the service of the ego; it is particularly important for us to crystallize out new concepts which support greater foresight, greater subjective appreciation of the reality of a larger universe in space and time, and -- in Freud's terms -- the resulting ability to delay gratification (even in terms of budget deficits, investment[9] and welfare[15]).

When we as individuals participate in such a system, I would predict that the issues of symbolic reason and Confucian integrity would still apply, both on the personal and on the collective level, perhaps with much greater force. The evolution of the greater system from a purely nonverbal level to more (self-)conscious, symbolic communication within itself would be crucial to its maturation. (Some old