Warren S. McCulloch [*]

"What's in the Brain That Ink May Character?"

SINCE WE have come together as scientists who would become a bit wiser as to the process of our art, it is proper for us to ask what are the enduring qualities of our activities and what are our present problems. Whether he would create poetry, fiction, or science, the American is apt to think first of Mark Twain's law: "You have to have the facts before you can pervert them." Which are the facts? They are those that puzzle us – and not even all of them, but those that arouse in us one and the same sort of uneasiness in various contexts of experience. From a vague sense of there being something similar in these facts, we become curious as to exactly what it is that is similar in them, and we define them with increasing clarity, doing all of this before we are able to phrase a single question to put to nature. At that stage we are uncertain whether we really have one question or several questions.

You will find this difficulty explicit in the writings of Galileo, who, in founding physics, speaks of two new sciences where we now find only one. Kepler, in the act of putting physics into the sky to produce elliptical orbits, was actually up against two questions, one in geometrical optics, and the other in mechanics, where he originally thought them one question. At the end of the last century, it looked as though physics was only a matter of pushing one decimal point to have a tidy theory of the universe. Only three awkward items had to be explained. These were the precession of the perihelion of Mercury, the drag of a moving medium on refracted light, and the absence of an aether drift. They raised three apparently separate questions, and no one expected that he had a single answer in the theory of relativity before that answer was forthcoming. Today there is a similar uneasiness in physics, perhaps foreshadowed by the want of a general field theory. It arises from the multiplication of the strange particles of subatomic physics, from the behavior of ballistic missiles, from transitions from streamline to turbulent flow, and from reports of an enormous object, a-fifth-of-the-age-of- the-universe away, which pulsates so fast that it requires a physical transmission immensely faster than light to keep it going.

The role of the projectile, and of its impact, in the development of physics may be of more than historical importance. In Galileo's hands, it proved fatal to the Greek doctrine of natural places. It disproved Descartes' attempted solution in terms of a plenum with a conservation of motion, and Leibniz' plenum with a conservation of force. It now threatens Newton's conservation of momentum. For macroscopic projectiles and their impacts, there seems to be an intrinsic time, or τ, during which they absorb or deliver energy but during which they are incapable of a conservation of momentum in the macroscopic sense, and thus require a third temporal derivative. Its introduction has also served to explain both varieties of turbulence, the quasiperiodic and the hyperbolic, or explosive, in our rockets. Davis has pointed out that, without this assumption, these can

---

* Presented at the International Congress for Logic, Methodology and Philosophy of Science, Jerusalem, Israel, August 28, 1964.
This work was supported in part by the Joint Services Electronics Program under Contract DA36-039-AMC-03200(E); the National Science Foundation (Grant GP-2495), the National Institutes of Health (Grants NB-04985-01 and MH-04737-04), the National Aeronautics and Space Administration (Grant NsG-496); and in part by the U.S. Air Force (Aeronautical Systems Division) under Contract AF33(615)-1747.
published in : Warren S. McCulloch, "Embodiments of Mind", The MIT Press, 1965, p.387-397.

only be explained away by distinct hypotheses *ad hoc*. Several of my friends have been asking whether or not atoms and particles may have a τ that accounts for some of their strange properties; and, at the other extreme of size, whether or not the gravitational field, like the electromagnetic field, may propagate, thereby giving a τ to gigantic structures. In short, it looks as though physics is again about to enjoy a new resolution, or at least a new revolution, and whether there be one question or many remains to be seen.

Since this is so in the most advanced of sciences, there is no need to apologize for the state of our own, for we are Johnnies-come-lately into the hypothetical and postulational stage of knowledge. Just as chemistry got off to a bad start in the rigid doctrine of alchemy and was saved only by the "puffers," so psychology was hindered by doctrinaire epistemology and saved only by biologists. To make psychology into experimental epistemology is to attempt to understand the embodiment of mind. Here we are confronted by what seem to be three questions, although they may ultimately be only one. It is these which we should like you to consider.

The three exist as categorically disparate *desiderata*. The first is at the logical level: We lack an adequate, appropriate calculus for triadic relations. The second is at the psychological level: We do not know how we generate hypotheses that are natural and simple. The third is at the physiological level: We have no circuit theory for the reticular formation that marshals our abductions.

Logically, the problem is far from simple. To be exact, no proposed theory of relations yields a calculus to handle our problem. When I was growing up, only the Aristotelian logic of classes was ever taught, and that badly. The *Organon* itself contains only a clumsy description of the apagoge – perhaps from the notes of some student who had not understood his master. Peirce says that when he was making the *Century Encyclopedia*, he understood the passes so badly that he wrote nonsense. "The apagoge," ordinarily translated "the abduction," is explained by Peirce as one of three modes of reasoning. The first is *deduction*, which starts from a rule and proceeds through a case under the rule to arrive at a fact. Thus: All people with tuberculosis have bumps; Mr. Jones has tuberculosis; *sequitur* – Mr. Jones has bumps. The second, or *induction*, starts from cases of tuberculosis and patients with bumps and guesses that the rule is that all people with tuberculosis have bumps. Peirce calls this "taking habits"; and properly it leads only to probabilities, coefficients of correlation, and perhaps to factor analysis. The guess at the rule requires something more – a creative leap – even in the most trivial cases. The third, or *abduction*, starts from the rule and guesses that the fact is a case under that rule: All people with tuberculosis have bumps; Mr. Jones has bumps; perhaps Mr. Jones has tuberculosis. This, sometimes mistakenly called an "inverse probability," is never certain but is, in medicine, called a diagnosis or, when many rules are considered, a differential diagnosis, but it is usually fixed, not by a statistic, but by finding some other observable sign to clinch the answer. Clear examples of abduction abound in the Hippocratic corpus but are curiously absent in Aristotle's own writings, where one finds only genus, species, and differentia.

What seems even stranger in the Greek writings is a total absence of our notions of a priori or a posteriori probability. The ancients had only a possibility and a guess. Probability as we know it was still nearly two thousand years in the future. Possibility appears in Aristotle's problematic mode but was even more sharply handled by the Stoics and by the physicians. Both groups questioned whether a possible proposition can be said to be true if it never happens to be fulfilled. One thing is clear, then – the mind makes a leap from the cases and facts to the rule, and Mill's attempt to bridge this gap, and the attempts of all of his followers; slur over it too easily. We do not know how we even make the jump and come up with a simple and natural hypothesis – certainly not from probabilities.

When I was young, it was fashionable to sneer at Stoic logic as mere pettifoggery; at that very time it was being slowly and laboriously re-created under the alias of the logic of propositions. Thanks largely to Northrop and to Sambursky, I have recently become familiar with its tenets. Had I known it forty years ago, it would have saved me much wasted labor. In the first place it is, as Peirce points out, both pansomatic and triadic in its propositions. There are always three real related bodies: One is the utterance, the *flatus vocis* of Abelard; one is that which it proposes; one is something in the head like a fist in the hand called the *Lekton*. Shakespeare, at about the age of twenty-five, had it clear and wrote for a lawyers' club:

> What's in the brain that ink may character,
> Which hath not figur'd to thee my true spirit?
> What's new to speak, what new to register,
> That may express my love or thy dear merit?

The lawyers for whom he wrote it were concerned with writing lawyers' law, which grows out of Stoic logic, giving us our contracts, corporations, and constitutions, created as postulated entities and hypothetical relations, much as we inherited this structure from the Greeks to start the renaissance of science. What's in the brain is the Stoic *Lekton*. Stoic law contemplates possible alternatives but never probabilities, and time enters, allowing no contract without date of termination, no bond without date of redemption, and no elected office but for a limited term.

Time appears in Stoic logic in the relation of the necessary to the possible, and I have heard lawyers discuss this as a probable source of this aspect of contractual law. There are three statements attributed to Diodorus, called the Master, of which any two may be true and the third false: Every possible truth about the past is necessary; an impossible proposition may not follow from a possible one; there is a proposition possible that neither is true nor will be true.

Diodorus rejected the third and defined the "possible" as that which is or will be true. This is in keeping with his notion of implication, which is concerned with time. He held that *A* implied *B* only if, for all time, *A*, as a function of time, materially implied *B*, as a function of that time. For the last of the great Stoic logicians, Philo, implication was our material implication. There were at least two other forms of implication used by the Stoics, one resembling strict implication, and the other perhaps requiring analyticity. Unfortunately, none of these is the implication that we really want for our purposes, and, as you will see, we have had to turn to biology for the notion of a bound cause. A signal should be said to imply its natural cause, which is bound, and not its casual cause; for when it arises ectopically, it is false for the receiver. The communication engineer calls such a false signal "noise." Again, the trouble is that we are dealing with a triad of Sender, Signal, and Receiver, and with the Stoic triad: *A* means *B* to *C*. The signal means to the receiver what the sender intended.

In order to avoid paradoxes and ambiguities, the Stoics not only would not allow any self-reference, as in the famous Cretan's "This statement is a lie," but would not allow a proposition to imply itself and, as an added precaution, would not allow a negation within a proposition. This left them with implication, and an exclusive or, and with a *not both,* the last of which is one of Peirce's *amphecks,* or a version of Sheffer's stroke. Hence, they needed exactly five figures of argument to form a complete logic of atomic propositions.

In about 1920, I attempted to construct a logic to handle the problems of knowledge and action in terms of a logical analysis of propositions involving verbs other than the copulative, and found it worse than modal logic. One has to distinguish those verbs in which the physical activity described by the present tense begins in the object and ends in the subject, such as verbs of sensation, perception, etc.; those in which it begins in the subject and ends in the object, such as the verbs of action; the group of so-called intransitive and reflexive verbs in which the events begin in the subject and end in the subject, called the verbs of behavior; and finally, a group of verbs that in the present tense refer to no action but define some kind of action that will be taken if thus-and-so happens – verbs of sentiment, which are like propositional functions rather than propositions. In perception, time's arrow points to the past; in action, to the future; in behavior, it becomes circular; and in sentiment, it simply does not exist. Literally, one deals with a state. I gave up the attempt because I realized that I had been trapped by the subject-predicate structure of language into supposing I was dealing with diadic relations, whereas they were irreducibly triadic. My hypothesis was simple and natural, but I had mistaken the *flatus vocis* for the *Lekton.*

I next attempted to construct for myself a simplest psychic act that would preserve its essential character; you may call it a "psychon" if you will. It was to be to psychology what an atom was to chemistry, or a gene to genetics. This time I was more fortunate, probably thanks to studying under Morgan of fruit fly fame. But my psychon differed from an atom and from a gene in that it was to be not an enduring, unsplittable object, but a least event. My postulated psychons were to be related much as offspring are to their parents, and their occurrence was in some sense to imply a previous generation that begat them. There is perhaps no better understood triadic relation than family structure. Even the colligative terms are clearly specified. There is scarcelya primitive tribe but has a kinship structure. So I was fortunate in this hypothesis in the sense that it gave a theory of activity progressing from sensation to action through the brain, and even more so in this, that the structure of that passage was anastomotic, whereby adherents of any sort could find their way by intersecting paths to any set of efferents, so relating perception to action. The implication of psychons pointed to the past, and their intention foreshadowed the proposed response. In those days the neuronal hypothesis of Ramòn y Cajal and the all-or-none law of axonal impulses were relatively novel, but I was overjoyed to find in them some embodiments of psychons. There was a *Lekton* in the head like a fist in the hand, but it took me out of psychology through medicine and neurology to ensure my pansomatism. Thereafter, in teaching physiological psychology at Seth Low Junior College, I used symbols for particular neurons, subscripted for the time of their impulse, and joined by implicative characters to express the dependence of that impulse upon receipt of impulses received a moment, or synaptic delay, sooner.

But even then I could not handle circularities in the net of neurons, for which I lacked a genetic model. They were postulated by Kubie, in 1930, to explain memory and thinking without overt activity in the supposititious linked-reflexes of the behaviorists. Circles were well known as regulatory devices, as reflexes, in which the action instead of being regenerative was an inverse, or negative, feedback. My major difficulty was having insufficient knowledge of modular mathematics. This, Walter Pitts could handle, and we published our paper on a logical calculus for ideas immanent in nervous activity. Chicago in those days was under the spell of Rudolf Carnap, and we employed his terminology, although it was not most appropriate to our postulates and hypotheses. Quite apart from misprints, this has made it unduly difficult for all but a few like Bar-Hillel, who worked with Carnap, and we shall always be grateful to Kleene for putting it into a more intelligible form. I still feel, however, that he treated closed loops too cavalierly and so left open questions that we had raised, and neglected certain distinctions that, in Papert's hands, may prove a source of new theorems relating nets to the structure of the functions that they compute. The history of the ensuing developments in automata theory is certainly familiar to you.

As geometry ceased to be the measurement of the earth, so automata theory is ceasing to be a theory of automata. Recently, in Ravello, I was told that an automaton or a nerve net, like me, was a mapping of a free monoid onto a semigroup with the possible addition of identity. This is the same sort of nonsense one finds in the writings of those who never understood the *Lekton* as an embodiment. It is like mistaking a Chomsky language for a real language. You will find no such categorical confusion in the original Pitts and McCulloch of 1943. There the temporal propositional expressions are events occurring in time and space in a physically real net. The postulated neurons, for all their oversimplifications, are still physical neurons as truly as the chemist's atoms are physical atoms.

For our purpose of proving that a real nervous system could compute any number that a Turing machine could compute with a fixed length of tape, it was possible to treat the neuron as a simple threshold element. Unfortunately, this misled many into the trap of supposing that threshold logic was all one could obtain in hardware or software. This is false. A real neuron, or Crane's neuristor, can certainly compute any Boolian function of its inputs – to say the least! Also, in 1943, the nets that we proposed were completely orderly and specified for their tasks, which is certainly not true of real brains. So, in 1947, when we were postulating a *Lekton* for the knowing of a universal, we began with a paragraph of precautions, that the function of the net be little perturbed by perturbations of signals, thresholds, and even by details of synapsis. All of this underlies the beginnings of a probabilistic logic to understand the construction of reliable automata from less reliable components, as is apparent in the work of Manuel Blum and Leo Verbeek. Finally, in the work of Winograd and Cowan, it is clear that, for an information-theoretic capacity in computation in the presence of noise, the logic has to be multiple-truth-valued, and the constructions require, for coding without fatal multiplication of unreliable components, not threshold elements, but those capable of computing any Boolian function of large numbers of inputs – that is, they must be somewhat like real neurons. The facts that worried us over the years from 1947 to 1963 were simply that real brains do know universals, are composed of unreliable components, and can compute in the presence of noise. The theory of automata has proved more provocative than the automata theory divorced from the automata.

Please note that, to this point, we have considered only deductive processes. The automata were not "taking habits." Our group has not been concerned with induction, either experimentally or theoretically. Soon after World War II, Albert Uttley produced the first so-called probabilistic perceptive artifact. It enjoyed what is now called a "layered computation" and could be trained to classify its inputs. He has stayed with the problem, and I happen to know that he has written, but not yet published, an excellent theoretical paper based on a specific hypothesis as to the events determining the coupling of neurons in succession; and, moreover, that physiological experiments performed by one of his friends indicate that his assumption is probably correct. I take it you are familiar with the writings of Donald MacKay, Oliver Selfridge, Marvin Minsky, Gordon Pask, Frank Rosenblatt, and a host of others on perceptrons, learning and teaching machines, etc., and that you know of the numerous studies on the chemical nature of the engram, which certainly involves ribose nucleic acid and protein synthesis.

The next step would obviously be to postulate a process of concept formation. This is the very leap from weighing probabilities to propounding hypotheses. Marcus Goodall, Ray Solomonoff, Marvin Minsky, and Seymour Papert, among my immediate friends, are all after it, and I think they all feel that it requires a succession of subordinate insights organized at successive superordinate levels or types. This is what Hughlings Jackson called "propositionalizing." This certainly cannot be left to variation and selection as an evolutionary process starting from chaos or a random net. That would be too slow, for it can be followed in ontogenesis, as Piaget has shown. The child does form "simple" and "natural" hypotheses, as Galileo called them. "Simple" and "natural" are evaluative terms and are based upon the evolution of the organism and its development in the real world, the natural world in which it finds itself. There again we come up against our logical limitation, for there simply does not exist any proper way to handle the triadic, or n-adic, relations of such relata. We cannot state our problem in a finite and unambiguous manner.

That man, like the beasts, lives in the world of relations rather than in a world of classes or propositions seems certain. He does not know the relative size of two cubes from a measurement of the lengths of their edges, or even from the area of their faces. If he can just detect a difference of one part in twenty of a length, he can do the same for areas and also for volumes. I happen to have spent two years in measuring man's ability to set an adjustable oblong to a preferred shape, because I did not believe that he did prefer the golden section or that he could recognize it. He does and he can'! On repeated settings for the most pleasing form he comes to prefer it and can set for it. The same man who can only detect a difference of a twentieth in length, area, or volume sets it at 1 to 1.618, not at 1 to 1.617 or 1 to 1.619. So the aesthetic judgment bespeaks a precise knowledge of certain – shall I say privileged? – relations directly, not compounded of the simpler perceptibles. A sculptor or painter has sometimes told me he had added enough to a square so that the part he had added had the same shape as the whole. This example is pertinent here, for in this case we do have an adequate theory of the relations, namely ratio and proportion. But these apply only to the perceived object, not to its relation to the statement or the *Lekton* in the brain of the aesthete. Clearly, the concept of a ratio must be embodied before the concept of a proportion can be conceived as the identity of the ratio. Once formed, the concept endures in us as the embodiment of an eternal verity, a sentiment, like love. To quote from the same Shakespearean sonnet CVIII:

> What's new ...
> ....
> Nothing, sweet boy, but yet like prayers divine,
> I must each day say o'er the very same,
> Counting no old thing old, thou mine, I thine,
> Even as when first I hallowed thy fair name.

So that eternal love in love's fresh case,
Weights not the dust and injury of age,
Nor gives to necessary wrinkles place,
But makes antiquity for aye his page,

Finding the first conceit of love there bred
Where time and outward form would show it dead.

Such is the beauty we still find, the pure form, the golden section, in the ruins of a Greek temple.

The golden section is a ratio that cannot be computed by any Turing machine without an infinite tape or in less than an infinite time. It is strictly incomprehensible. Yet it can be apprehended by finite automata, including us. Nor does it arise from any set of probabilities, or from a factor analysis of any data or correlation of observations, but as an insight – a guess, like every other hypothesis that is natural and simple enough to serve in science. It is nearer to the proper notions of classical physics than to the descriptive laws, the curve-fittings, that bedevil psychology.

This brings us to the problem of abduction, the apagoge. Evolution has provided us with reflexive arcs organized for the most part by what are called "half-centers," whose activities may alternate, as in breathing or walking, or synchronize, as in jumping. These are then programmed for more complicated sequences, and all of these are marshaled into a few general modes of behavior of the whole man. Psychologists and ethologists count them on their fingers or at most on their fingers and toes. These modes of behavior are instinctive, and only the manner of their expression and their manner of evocation are modified by our experience. The structures that mediate them have evolved in all linear organisms, like us, from an original central net, or reticulum, and while they may be very dissimilar from phylum to phylum, the central core of that reticulum, has remained curiously the same in all of us. It is distributed throughout the length of the neuraxis and in each segment determines the activity of that segment locally, and relates it to the activity of other segments by fibers, or axons, running the long way of the neuraxis. The details of its neurons and their specific connections need not concern us here. In general, you may think of it as a computer to any part of which come signals from many parts of the body and from other parts of the brain and spinal cord. It is only one cell deep on the path from input to output, but it can set the filters on all of its inputs and can control the behavior of the programmed activity, the half-centers, and the reflexes. It gets a substitute for depth by its intrinsic fore-and-aft connections. Its business, given its knowledge of the state of the whole organism and of the world impingent, upon it, is to decide whether the given fact is a case under one or another rule. It must decide for the whole organism whether the rule is one requiring fighting, fleeing, eating, sleeping, etc. It must do it with millisecond component action and conduction velocities of usually less than 100 meters per second, and do it in real time, say, in a third of a second. That it has worked so well throughout evolution, without itself evolving, points to its structure as the natural solution of the organization of appropriate behavior. We know much experimentally of the behavior of the components, but still have no theory worthy of the name to explain its circuit action. William Kilmer, who works on this problem with me, is more sanguine than I am about our approach to the question. Again, the details of our attempts are irrelevant here. The problem remains the central one in all command and control systems. Of necessity, the system must enjoy a redundancy of potential command in which the possession of the necessary urgent information constitutes authority in that part possessing the information.

The problem is clearly one of triadic or n-adic relations, and is almost, or perhaps entirely, unspecifiable in finite and unambiguous terms without the proper calculus.

We see, then, the same theme running throughout. We lack a triadic logic. We do not know how to create natural and simple hypotheses. We have, at present, no theory to account for those abductions which have permitted our evolution, ensured our ontogenesis, and preserved our lives. The question remains:

What's in the brain that ink may character?

The text was originally edited and rendered into PDF file for the e-journal <www.vordenker.de> by *E. von Goldammer*

ISSN 1619-9324

How to cite:
Warren St. McCulloch: "What´s in the Brain That Ink May Character?", in: www.vordenker.de (Edition: Winter 2008/09), J. Paul (Ed.), URL: < http://www.vordenker.de/ggphilosophy/mcculloch_whats-in-the-brain.pdf > — orginally published in: "Warren S. McCulloch, Embodiments of Mind", The MIT Press, 1965, p.387-397.