vordenker
Winter-Edition 2008/09

Warren S. McCulloch [*]

# Through the Den of the Metaphysician

## ARGUMENT

Epistemic questions raised by early physics are theoretically answerable in terms of communication by means of least signals that are propositions on the move in computing machines. We can know only the past, and information, being negative entropy, can only decrease in passage. Requirement of coincidence of signals at a relay decreases the logical probability of a signal in it and so increases the chance that its signal corresponds to something in the world. But to suppose input signals true is the superstition of causality. Attempts to extend the domain of implication into the future may always be thwarted, hence the notion of the will, which for moral responsibility must pass into the deed. Among deeds, words stand for universals which correspond to configurations of signals; without closed circuits, they refer only to one past instant. But reverberations patterned after some fact preserve its form and introduce existential operators for times past. Thus, active memory recognizes things new to us in the world. Reflexes, say by centering forms to be seen, or automatic volume controls within the brain, say by bringing afferent signals to a fixed mean frequency, and appetitive circuits over targets outside the body, all reduce the given through a series of transformations to the canonical one among its many possible exemplars. This one serves as the original in the cave of the sun. Finally, the brain, given any item, may form all of its transforms belonging to some group and assigning arbitrary values to presence or absence of excitation at all points and times in the structure wherein the transforms are made. Sum these values. A set of such invariant sums constitutes an Aristotlean abstraction. By detecting noncongruence of apparent continua of sense awareness, even a system constructed of leasts can form the universal of "in-between" and propose the Eleatic riddles.

By projecting universals as expected regularities of all future experiments, it can frame hypotheses in order ultimately to disprove them. It can build bigger and better brains, but what it has are sufficient to guide it through the very den of the metaphysician.

ANY AMERICAN who comes to the University of Virginia with matter philosophical must feel a bit as if he were bringing coal to Newcastle. Fortunately, most of this matter has been mined rather recently in America. We are again in one of those prodigious periods of scientific progress – in its own way like the pre-Socratic period to which we are still indebted for the crisp formulation of out physical problems and, consequently, for out epistemological quandry. Anyone who has had the good fortune to listen to Wiener and von Neumann and Rosenblueth and Pitts wrestling with the problems of modern computing machines that know and want has a strange sense that he is listening to a colloquy of the ancients. But they would be the first to tell you that they themselves were drunk with an American wine of an older vintage; they quote liberally from Charles Peirce and from Josiah Willard Gibbs. These men have altered out metaphysics by altering out physics. It is epistemology that is most affected, for it is the physics of communication which is today receiving an adequate, theoretical treatment. For the first

---

* Warren S. McCulloch, in: "Warren S. McCulloch: Embodiments of Mind", The MIT Press, Cambridge, Mass. 1965, pp. 142-156. (first published under the French title "Dans L´antre du Métaphysicien" in: Thales, vol. 7, 1951, pp. 35-39.)

time in the history of science we know how we know and hence are able to state it clearly.

Physiologists, working on the central nervous system, have long had such a goal in mind. Rudolph Magnus, inspired by Immanuel Kant, made his last great lecture one on "the physiology of the *a priori*", by which he meant the go of those mechanisms that determine for us the three-dimensional nature of out world, its axes and its angles, and that give to us out sense of velocity and acceleration, from which he held out notion of time to be in large measure derived. Perhaps the most notable attempt of this sort was by Sir Charles Sherrington, entitled *Man on His Nature*, for, near the end of a life spent on studying the ways of the brain, he was forced to the conclusion that "in this world, Mind goes more ghostly than a ghost." The reason for his failure was simply that his physics was not adequate to the problem that he had undertaken. That has so regularly been the shortcoming of scientists who would have approached this problem, that even Clerk Maxwell, who wanted nothing more than to know the relation between thoughts and the molecular motions of the brain, cut short his query with the memorable phrase, "but does not the way to it lie through the very den of the metaphysician, strewn with the bones of former explorers and abhorred by every man of science?" Let us peacefully answer the first half of his question "Yes," the second half "No", and then proceed serenely.

Our adventure is actually a great heresy. We are about to conceive of the knower as a computing machine. That is not a new heresy. It has already been prejudged by Dryden in *The Hind and the Panther*, when he says,

And if they think at all, 'tis sure no higher
Than matter, set in motion, may aspire.

I believe that he is correct, but I am not sure that that may not be high enough. I have no intention of burdening you with the detail of the construction of the computing machine, whether these be man-made or begotten. The latter are my daily business. My problem differs from that of the men who build computing machines only in this – that I am confronted by the enemy's machine. I have not been told and must learn what it is, what it does, and how it does it. It is a complicated computing machine consisting of $10^{10}$ relays. Each of these relays receives signals from other relays. Each on receipt of an appropriate signal – or group of signals – emits a signal. It is my business to learn how these relays are connected one to another, what it takes to fire a given relay, how long after receipt of a signal it will send a signal, and how a signal received can prevent a relay from responding to a second signal otherwise sufficient.

But we must first be prepared for the kind of world we now invade. It is a world, for Heraclitus, always "on the move". I do not mean merely that every relay is itself being momentarily destroyed and re-created like a flame, but I mean that its business is with information which pours into it over many channels, passes through it, eddies within it, and emerges again to the world. Surely Heraclitus would feel at home with such a knower. Paradoxes raised by such a conceit of the world have always led, through Parmenidean unity and Eleatic riddles, to a Democritean multiplicity, that is, to one in which the stuff of the world is a set of atoms – of indivisibles, of leasts – which go batting about in the void. Whenever they make their appearance they bring with them Chaos. Therefore, it may at first sound paradoxical that every modern computing machine of any great size or scope works in a Democritean manner. Again, I do not merely refer to its being constructed out of chemical atoms, but I mean that it is a machine with a least count. Its signals are quantized. Each either happens or does not happen. It does not half happen. To this general rule, the nervous system is no exception. Its least signal, or nervous impulse, is an none event. If a neuron emits a signal, it does all that it can then do. Thus, not merely is the structure of the nervous system quantized in neurons, but its action is quantized in their impulses, or least signals: Surely Democritus would claim kinship with a knower whose actions were thus atomic – perhaps more readily because there is at present no reason to suspect that these atomic signals differ significantly save as to when and where they occur. Hence, all that a least signal can say to the next relay is that the relay that emitted that signal had been adequately excited. From this it follows that whatever quality may be anywhere detected in the universe by our knower must depend upon the *figure* of these least signals in time and space. Moreover, since every relay has a characteristic delay, given enough relays, it is always possible to convert a figure of impulses given simultaneously in space into a figure of excitation in time over a single relay, or vice versa.

Consider any one relay. At any given time, say a millisecond, it can be in either of two states; that is, it can be transmitting one signal or none. Hence, two independent, or unconnected, relays can be in any one of four states; three in one of eight, four in one of sixteen, etc. That is, of n neurons, the number of possible states is $2^n$. It takes one signal, or one unit of information, to determine in which of two states any one relay is in any one relay time. Now if we have a single relay but consider it at two times, then it can be, in the two times together, in any one of $2^2$; in three times in $2^3$; in four times, $2^4$; and so on. Clearly, the same amount of information can be conveyed by n independent relays in one unit of time as can be conveyed by one relay in n units of time. Hence, if we desire to convert a given figure at a one time into a series in time, we need as many units of time as we had independent relays. Note that the unit of information appears with a negative sign in the exponent; that is, given n independent relays, or one relay at n times, one unit of information, by fixing the state of one relay, subtracts one from n, leaving $2^{n-1}$ states possible.

The *a priori*, or logical, probability that a neuron is in a particular state at a particular time is one-half; that two are in a given state, one-fourth; and so on. Hence, information is exactly the logarithm to the base 2 of the reciprocal of the probability of the state. But this has a peculiarly familiar sound. Gibbs had defined entropy as the logarithm of the probability of the state. In Wiener's words, entropy measures chaos, and information is negative entropy. So, corresponding to the second law of thermodynamics, that entropy must always increase, we can write for any computing machine the corresponding law – information can never increase. This ensures that no machine can operate on the future but must derive its information from the past. It can never do anything with this information except corrupt it. The transmission of signals over ordinary networks of communication always follows the law that deduction obeys, that there can be no more information in the output than there is in the input. The noise, and only the noise, can increase. Therefore, if we are to deal with knowers that are computing machines, we can state this much about them. Each is a device, however complicated, which can only corrupt revelation.

In order to preserve a correct sense of proportion, let me be technical for a moment. The human eye has about one hundred million photoreceptors, whereas it has but one million relays to carry that information to the brain. The whole body contributes another million channels. Thus we may figure approximately three million relays putting information into the nervous system simultaneously. Let us next evaluate the output of the nervous system. To do so we have conceived a piano player performing at top speed, given him a keyboard of one hundred keys, let him strike ten times per second with each of ten fingers with any one of ten strengths, and let his hands each span ten keys. No man can do so much. Yet, when we translated this into the number of all-or-none signals per second, we discovered that it was only three units of information per millisecond.

Today we can estimate the amount of information conveyed by sounds, speech or music, for there is a device which samples the sound every millisecond and sends, at most, three all-or-none signals according to the instantaneous amplitude of the wave. These three decisive signals per millisecond convey the full information, for waves reconstructed from them are indistinguishable from the original sounds.

Thus the over-all reduction in information from input to output of brain is a million to one if we neglect the eyes proper, and a hundred million to one if we include them. What becomes of all that information?

In large measure we use it this way. It is easy to have a relay which will fire only if impulses from two sources arrive almost simultaneously. Such a neuron detects the coincidence of information over the two channels to it, and it responds only in that case in which they agree that something happened. Therefore, the logical, or *a priori*, probability of finding an impulse in it is the product of the probabilities of finding one in each of the afferent pathways singly. That is to say, it is more *improbable*. The chief reason for the enormous reduction from afferent signals to efferent signals is the requirement of coincidence along the way. Every such requirement of coincidence, by reducing the *a priori* probability of a signal in the output, increases the assurance which can be placed in any subsequent signal, for that signal must then be due to coincidence in the world impingent upon our receptors. In short, by throwing away all information that fails to agree with other information, we achieve an immense certainty that what we do observe is due to something in the world. Clearly, we should follow the same procedure in forming our hypotheses. Each should be of minimum logical, or *a priori*, probability so that if it is confirmed in experience, then this is because the world is so constructed.

At this point I should tell you that the limitation of the information in the output is in large measure determined by the effectors themselves. This is the ineluctable corruption of thought in deed. As there remain $10^{10}$ neurons, or relays, in the central nervous system, it is obviously impossible for any man ever to convey so much as one part in $10^{10}$ of what is going on in him. Even though he is a poet, the rest of his soul remains his private property. Now we habitually think of our sensations, or any knowledge of the world so derived, as an activity going on in those places in which afferent channels end, but we can demonstrate it only by output over some efferent channel. Consequently, the answer to any particular question, as to whether we did or did not know something, will be dependent upon the particular efferent channels we choose to examine. This is painfully obvious to anyone who deals with diseased brains, for it is often a matter of diagnostic or forensic importance to know whether a man is or is not conscious of his deeds or of things about him. *Conscious* in this context means that he can, then or later, bear witness to those events to which we can also bear witness. In malignant stupors the patient is conscious, in benign, unconscious, but we can determine this only when he is no longer stuporous. The epileptic who is unconscious of his acts is never responsible for them at law, although the acts are so complex and so adjusted to the world about him as to ensure that he is then and there aware of much related detail in what is then adoing.

With all of these limitations and hazards well in mind, let us ask whether a knower so conceived is capable of ,constructing the physics of the world which includes himself, But, in so doing, let us be perfectly frank to admit that causality is a superstition. By causality I mean any law of necessary connection between events.

Let me put that this way, in the old quillet. Nothing is true and not true, nothing false and not false, nothing true and false. Hence, whatever is is either "true or else false"; or else it is "neither true nor false." In the first case it is a proposition; in the second case it is not. It is then impossible truthfully to deny that propositions exist, for either the denial is a proposition, in which case it is false, or else it is not a proposition and, hence, neither true nor false. Yet, that it is impossible to deny its existence truthfully is no assurance that any proposition exists. It remains an assumption that propositions exist. Later it will be apparent that even that assumption is inadequate. If there is to be in the world any knowledge about the world, there must be true propositions about the world. In the world of events a true proposition is an event which materially implies another event – that is, in the simplest case, one which happens only if that other event happened. In the world of physics a true proposition implies what it asserts.

The world of physics is fairly described by Whitehead's "Aether of Events." The world for him is the whole that happens only once; and to be an event is to be some part of that whole. In his description of the relation of whole to part, there are, however, two assumptions which we have ignored in our theory of the action of the computing machine, and it is these assumptions which permit of analytic continuation, a going toward continuity. He supposes that there can exist an event which has no least part, and that if A is a part of C, then there is always an event B, such that A is a part of B and B is a part of C. Apparently a computing machine, quantized as ours, is not built on either of these assumptions. We have supposed a least signal, that is to say, a signal which either occurs or does not occur. That least signal is a proposition "on the move." It is true or else it is false, and it occurs at some particular time and at some particular place. Since the number of relays is finite, these can be ordered and a number assigned to each, and since we can quantize time in units equal to relay time, we can start counting at any particular instant and, by subscripting the number of a relay by a number representing the time of a signal there, we can construct statements in which the signal of a given relay is expressed in terms of those signals which reach it. The resultant calculus is the calculus of atomic propositions of Whitehead and Russell, subscripted for the time of the occurrence of the propositions. Each of these signals is, strictly speaking, an atomic propositional event, which can only be or not be, and, if it is, may be either true or false. Each materially implies its proper antecedents.

Consider for a moment a computing machine in which there are no closed paths, that is, no circuits around which signals may chase their tails. In such a system each signal, implying its antecedents, implies a signal of a relay nearer to the receptors until we arrive ultimately at them. Their signals likewise imply the world impingent on our sense organs. In the strictest sense of the word, what goes on in such a nervous system implies the world impingent upon its sense organs. But note that the domain of implication extends only backward in time. Even if the threshold of every relay were fixed, between the moment at which a given set of impulses started from a spot in our brains to our hands and feet, there might intervene other impulses coming, by shorter paths from the outside world, to our effectors. Thus, in the forward direction, the relation falls short of implication inasmuch as aught else intervenes. In short, our thought does not imply our action but, as we say, only intends it. Perforce we distinguish between futurity and intention. Our notion of our wills has, I believe, arisen from this enforced distinction, and its perennially questioned "freedom" presumably means no more than that we can distinguish between what we intend and some intervention in our action. If I shall do what I will do, then my will is free.

To make this clear, let me return to the quillet. Even an atomic signal asserts that such and such is the case, and it is true only if such and such is the case. If we assume that there are such things as true atomic propositions of the kind called signals, then these are significant propositions as Wittgenstein uses the term; for the truth or falsity of each signal depends upon whether or not that which it asserts occurred. If significant propositions of this kind are to be true, there must be a law of necessary connections between the event which is the proposition and the event which it proposes. But a law of necessary connection among events is causality. In such a world the superstition of causality must first be assumed by anyone who wishes so much to deny it. But, if we have once admitted causality in this sense into the workings of our brains, our significant propositions are determined by our past, and freedom from the past would make no sense. We really want freedom toward the future – freedom from affairs intercurrent between our ideas and our deeds. This is all that is needed to fix our responsibility for those deeds.

Among those deeds are our words, and every word has the flavor of being given "once for all." It bespeaks a universal, idea, or quality, appearing for us in the world. In order for the activity of a nervous system to imply a universal, or idea, we need not merely the calculus of atomic propositions but also logical quantifiers which assert that *all* or *some* x's are such and such; and the question at once arises how these are introduced. Again in Whiteheadian terms, the problem is not one of how we apprehend events but of how we recognize objects. For me, the most difficult part of Whitehead's theory of percipient events is concerned with the notion of primary recognition. How can we come by an idea we did not have before? You will notice that, had our nervous system no closed paths, a signal anywhere within it would mean that something had impinged on a receptor at some particular millisecond prior thereto. Per contra, if there exist closed paths around which signals may reverberate, and if the sequence of these signals is patterned after some fact, then the pattern persists in us as long as the signals continue to reverberate. Each time they circle they literally know again, or re-cognize, that which was given but once in their input. This is memory of a kind, and it suffices to free the signals of their reference to a particular time. That circuit knows that such and such happened at some previous time but not at what time. This introduces the existential operator for time – namely, there was some time such that at that time so and so happened. Notice that the tense is past.

Not all human memory is of this kind. While we are young, we may grow new connections. By means of them the ways that led us to our ends become embroidered into the warp and woof of our nervous net, but no kind of memory can do anything which cannot also be done by activity circling in closed paths. Pitts and I have already demonstrated that a machine constructed with such closed paths, as well as open afferents and efferents, can produce any consequence which can be produced from its afferents, that it can compute any computable number, or arrive at any conclusion which follows logically from any finite set of premises. Moreover, to satisfy any doubting Thomas who would thrust his electrodes into the brain, we have shown that any wound or any other alteration of the net, say one acquired by use, can be replaced by the action of a hypothetical unalterable net of unalterable neurons. But injury and learning will differ in one respect; for, in the case of learning, the known afferent channels suffice, whereas, in the case of injury, we require a new afferent channel to initiate an action in our hypothetical net.

Conceive, then, our knower as an unalterable net of unalterable relays. To this net come impulses from the world, and from it impulses go to the world. Within it are not only thoroughfares but circles. Of these we have mentioned only one that will reverberate. It suffices to free us from one particularity, reference to one past time, but there are other closed paths which are important in our knowledge of universals. Their reaction was well described in 1817 when Magendie defined what he called the "reflex" as an activity which began in some part of the body, passed by way of nerves to the central nervous system, whence it was reflected to that same part of the body where it stopped or reversed that process which had given rise to it. Electrical engineers refer to such a circuit action as inverse, or negative, feedback. The anatomy of some of these negative feedbacks was exhibited in 1825 by Sir Charles Bell. The mathematics for handling many of such actions is to be found in Clerk Maxwell's paper "On Governors," given before the Royal Society in 1868. Their properties are well known. Every such circuit pulls toward some particular value of some variable. In the case of reflexes, something, say the length of a muscle or the temperature of the body or the pressure in the artery, is measured by a set of receptors which send impulses to the central nervous system, whence impulses are sent back to those structures to bring them back to the particular state established by the reflex arc. This particular state measured by those receptors is the goal, or aim, or end, in and of the operation of that reflex. By means of these reflexes we achieve dynamic stability in a changing world.

One such circuit frees us from the variations in intensity of stimulation and consequent variation in the number of impulses that would otherwise excite the bark of the brain, called the cerebral cortex. This circuit, unlike the reflex path, lies entirely within the brain. Impulses ascending to the thalamus, or last relay beneath the bark, are passed on to the cortex and from portions of the cortex descend, by a devious path, to the thalamus, where they inhibit the relaying of impulses to the cortex, thus keeping the cortical activity at a nearly constant level. This is an "automatic volume control." It leaves us with a figure of excitation in the cortex which is invariant under fluctuations of the intensity of peripheral stimulation, making it possible for us to determine some aspect of that stimulation regardless of intensity. Hence we detect that there was some intensity of stimulation which was of such and such a figure. In short, it is done for us with respect to intensity what simple reverberation did with respect to time.

Let us next consider the so-called appetitive circuit. It is inverse feedback over a path which extends beyond the body. In the external part of the path is to be found the goal, or target, and the circuit is said to be "inverse over the target." Today one of the best known of these circuits runs through the eye and the brainstem. It turns the eye automatically toward any object which appears in the periphery of the visual field. By turning the eye so as to bring the image of the thing seen to the center of the receptive surface of the eye, it translates an apparition to a standardized position and hence its projection by means of signals to a standardized place in the bark of the brain. In so doing, it rids the form to be seen of the gratuitous particularity of that position at which it happened first to appear. Because for us the centered apparition established by this reflex is a given one from among the many possible exemplifications, we refer to the process as that of reduction to the canonical position. Every reflexive circuit brings some apparition through a series of positions, intensities, or what-nots to the final canonical one of all the many through which it has progressed. So, in mathematics, if we are given the Pythagorean theorem, we reduce it to the canonical form of the axioms, postulates, and the definitions at the beginning of the book. Having chased it thus back to the cave of the sun, we cry, "Ah-ha," for we have recognized it.

We pass now to another way of securing universals. At first sight it looks altogether dissimilar. It is, in fact, obviously a method of averaging. The epicritic modalities of sensation, for example vision, map the impingent activity of the world on a fine mosaic of cortical relays. The centered form, given by the reflex of the eyes, appears in the input to the visual cortex as the distribution of impulses over a fixed area and about a fixed center. In this portion of the cortex of the brain are actually made all of the possible dilatations and constrictions of the forms impingent upon it. They are limited by the grain of its mosaic. These are then added and relayed, as a set of averages, to the next portion of the brain. Clearly, these averages do not depend upon the size of the particular object that was seen. The image may have been big or small, but since from any given one of them we made all sizes, the size of the original does not affect the averages. The output is size-invariant. This size-invariant has been abstracted from the entire group of dilatations and constrictions and corresponds to the shape regardless of the size. As the former method conforms to Plato's notions as to the origin of ideas, so this recalls Aristotle's notion of the abstraction of ideas from many particular exemplars. Those of you who desire a clear and rigorous statement of this process will find it in the *Group Theory and Quantum Mechanics* of Weil.

Let me oversimplify that statement. We have obtained a group invariant as follows: We have calculated a set of numbers, each of which is, for all transformations belonging to the group, the average of numerical values arbitrarily assigned to the presence or absence of signals at particular points or at particular times. That in which one mechanism obtaining these invariants, or universals, differs grossly from another mechanism lies in that arbitrary manner in which the numerical values are assigned. For example, in the case of the centering reflex, the value *zero* is actually assigned to all translations of the apparition until it is centered. Whereas, in the case of the dilatations and constrictions in the visual cortex, all evidence points to the value one being assigned if a cell situated there was fired then – otherwise zero. We have, then, in the general statement all ways of obtaining universals. Their differences are accidental and depend upon the arbitrary way in which the mechanism assigns the values. At this level, Aristotle's method and Plato's method are peas in the same pod.

There is, of course, one limit to the number of times this process can be repeated upon the output of a previous structure. We have only a finite number of relays and finite span of life measured in their unit times. Short of that limit we are, of course, at liberty to have the idea of ideas and the idea of the ideas of ideas, etc. In short, our knower is, in the Spinozistic sense, conscious. Also, he may discover himself among things known to him in his world, and be self-conscious in exactly the same way in which he is conscious of anything else in the world. But all of his ideas, no matter how high-flown, are ultimately reducible to a logic of the lowest level and, even at that lowest level, to a finite number of atomic propositions. He can no more know the infinite than he can know the future. Of course, he may guess the future. He has only to run a correlation (over time) of past events, abstract a universal, and project it upon the future. The projection is always but a guess. The best guess may go wrong. The scientist knows this to his cost; for every hypothesis is a guess as to the outcome of an infinite number of possible experiments. We expect that every hypothesis will be disproved. Surely none can be proved. This is, in fact, its glory; for we know, *when we have* proved it false, that it was a significant proposition.

To have proved a hypothesis false is indeed the peak of knowledge, for we extrapolated it out of atomic propositions in the input to our calculating machine. These signals in the input are independent propositions. Any one can be false or true and the rest remain the same, be they true or false. Of no one of them can we ever know that it is true or false. These are merely revelations, and we cannot look the giver of the data in the teeth. How, out of such data, finite and discrete, have we been able to construct the notions of analytical continuity which are so clearly requisite for the Whiteheadian analysis of the physicist's world?

The problem of analytic continuation is pointed up in the paradoxes of Zeno of Elea. We require a meaning of "in-between" quite different from that given in any single experience. The world given in experience is always this up to where it is that. The "continuum of sense awareness" is not continuous in the mathematical sense. Where we have no receptor, we pick up no impulse. That point has no connection to the cortex. It is not represented directly by any impulse arriving there. We are not directly aware of our ignorance. For a moment conceive the same portion of the world given in two of these speciously continuous modalities of sensibility. In each the world appears to be continuous. But there is no reason why that which is continuous in the one modality should fit that which is continuous in the other modality – no *a priori* reason! Most of the time the continua fail to be congruent, and any circuit, subsequent to these recipients, which can detect the failure of congruence will generate a meaning of in-between. This meaning of in-between is just as good a universal as any other; for it we can construct an invariant. We have but to give it voice and, like God or a baby, cry "Do it again, do it again, don't stop," and we have generated the fundamental notion required for analytic continuation. For the paradoxes of Zeno of Elea, seen motion appears continuous through seen space likewise continuous. We have but to abstract the required universals, and we are able to propose the paradoxes.

I do not believe that anything more is expelled of us. A stick thrust in water felt straight and looked bent to a Greek. The sun moved for the inquisition, the earth for Galileo. Light is a wave for Schrödinger and a particle for Heisenberg. But even the last have had their Dirac. The seeming contradictions vanish in the grace of greater knowledge. We have learned that the answer depends upon how we ask the question. And we have learned to ask the question so as to get an answer of a kind that we can use. Knowledge itself presented no great problems to the Greek mind until it had invented its theoretical physics in terms of insensible reals. It was these that bothered Galileo and Descartes.

To suppose the world continuous instead of discrete did not help Descartes. He was willing to admit that an automaton might be constructed which could do everything that a man does. Yet, to this automaton he would deny Mind. Doubting, knowing, and thinking seemed to him beyond its scope. Now that we have constructed automata which, like us, can compute any computable number, can formulate clear ideas, and, by inverse feedback, have purposes of their own, built into them as ours are born in us, we are confronted with the humbling prospect of the work of our own hands – machines more steadfast in their purposes, more supple in the execution of these purposes and in their modifications for good cause, capable of learning and thinking far beyond us, at present in certain fields only but, in time to come, in any field for which we care to construct them. As yet we have not made them capable of multiplying their kind. That would be for us the final mistake.

In closing, let me remark that I am not using knowledge in any restricted sense. The propositions with which I have concerned myself are the significant propositions of machines, those that propose something external to the event which is the proposition. But if instead of relays that wait for signals to trip them, I install in any computing machine a relay which will fire itself each millisecond, I can introduce into it all tautological propositions. To my mind, they fall short of being knowledge, which I look upon as an activity that says "such and such is the case" and such and such is the case! These propositions are primary and atomic. We cannot know that we know them. A lie is as truly a proposition as is a truth, and their independence prevents any test of truth. In hypotheses, proved false, we are at least aware of the error of our ways. For a scientist that must be sufficient. If we cannot rest content with our brains, we can at least construct machines which, like them,

> If they think at all 'tis sure no higher
> Than matter put in motion may aspire.

I am convinced that that will be sufficient to guide me through the very den of the metaphysician, strewn with the bones of the former explorers. One of these is surely the femur of Immanuel Kant his confusion of the empiric with the epistemological ego. This supported him on the solid ground of science while his skull was highest in the realm of theory. Another is certainly his skull, which housed his computing machine, for the net of his relays embodied his "synthetic a priori." If my bones are to fall beside them, I hope aftercomers will recognize my spine. Its joints are the superstition of a necessary connection between events – called causality. I humbly submit that it is but a reincarnation of Saint Thomas' faith that God did not give us our senses to fool us. It is enough that this trust in the goodness of God cannot truthfully be denied. So, at least in Virginia, the den of the metaphysician seems curiously like the cave of the sun, and hence like home.

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



vordenker
ISSN 1619-9324

**How to cite**:
Warren St. McCulloch: Through the Den of the Metaphysician, in: www.vordenker.de (Edition: Winter 2008/09), J. Paul (Ed.), URL: < http://www.vordenker.de/ggphilosophy/mcculloch_through-the-den.pdf > — published in: "Warren S. McCulloch: Embodiments of Mind", The MIT Press, 1965, p.142-156. (first published under the French title "Dans L´antre du métaphysicien" in: Thales, vol. 7, 1951, pp. 35-39.)