PRELIMINARY DRAFT

Mathematical Principles of Cognition

Neil W. Rickert

ABSTRACT. under preparation

KEYWORDS: under preparation

1 Introduction

My aim, in this report, is to provide a mathematical model of human cognition, in order to present what I believe to be the important principles that underlie cognition. As with most mathematical models, this one oversimplifies and idealizes the problem of cognition. The aim is not to provide a blueprint for designing an artificial person, but rather to highlight the important principles involved as a guide to those seeking to better understand cognitive systems.

Although the model depends on some advanced mathematics, I shall attempt to present it in a way that does not too seriously stress the mathematical knowledge of the reader. For the main part, I shall attempt to give straightforward descriptions of the mathematics that is used, with references provided for those who want to more fully acquaint themselves with the underlying mathematical theory.

I shall be primarily concerned with modeling perception and its role in cognition. We can think of perception as a system for getting information about the world, and we can think of science as a kind of extended perception for gaining information (through it measuring instruments) that are not directly available to our sensory cells.

2 Our perceived world

For the sake of being able to provide a mathematical model, we shall denote by W, the world that we perceive. As a first approximation, we may think of W as consisting of points in space time. I may perceive that the traffic light is green. But, a little later, I see that it is red. Since I can distinguish between the two, on the basis of time, we must allow time as one of the 'dimensions' of our world W.

Our perception is not limited to the physical world. We perceive the moods of others around us; we perceive many social and cultural phenomena. By virtue of our ability to use measuring instruments such as a ruler or a thermometer, we extend our perception to physical phenomena. And, by virtue of our ability to read newspapers, watch television, read book, etc., we further extend our perception to include mathematical idea, and events in parts of the world that are far from where we are located. Thus we can think of W as a world that is rich with emotional, cultural and social dimensions, in addition to time, space and other physical dimensions.

Our perceptual ability can be thought of as functions over W. For example, our ability to recognize the color green can be treated as a function f, such that $f(x)$ is the intensity of greenness we perceive at location x in W.

I shall be thinking of W as a topological space, and of our perceptual functions as continuous functions over W. For convenience it is best to think of the perceptual functions as having values that are complex numbers, mainly because much of the mathematical literature is oriented toward complex valued functions.

For those with a technical (mathematical) interest, I am assuming that W is a normal Hausdorff topological space. That assumption indicates that there are sufficient continuous functions over W to enable us to distinguish between points in W. For terminology, see [5] or other books on point set topology.

Beyond its being a normal space, I make no other assumptions about the structure of W or its topology. Part of what a child has to do, is learn about the world, and learning something of the structure of W as a topological space is part of that learning. The underlying principles involved in obtaining that knowledge of W is one of the concerns of this report. And, of course, we should not assume that a cognitive agent ever achieves complete knowledge of W.

3 Perceptual functions

I shall assume that perceptual ability exists in the form of perceptual functions. I take these functions to be complex valued continuous functions over the world W. It is probably more realistic to assume real valued functions. However, a real number is just a special case of a complex number. It is reasonable to doubt that the neurons in the brain deal with numbers at all. However, as with most mathematical models, this is an idealization of what an actual perceptual system might do.

I have already suggested that we can consider the intensity of greenness to be a function. We can write that in a function-like notation as $\mathrm{greenness}(x)$, or as $\mathrm{greenness}()$ when we wish to omit mention of the specific point $x \in W$ where this function is evaluated. Similarly, we can have perceptual functions such as $\mathrm{distance}()$, $\mathrm{temperature}()$, and many others. We are able to perceive cats, so we can also treat that as a perceptual function. We might think of a function $\mathrm{felinicity}(x)$ which has a value 1 when point x is part of a cat, but which drops off toward a value of 0 as we look at points x that increasing distances away from any cat.

You may notice that these perceptual function find values for what we often consider to be properties. When we determine the values of those properties in science, we describe that as measurement. So it seems reasonable to assume that what the perceptual system is doing is something akin to measurement, and I shall sometime describe perceptual operations as measurment.

I should add that perception is not easy. You cannot just setup a green sensitive photocell, point it at x, and thereby find the value of x. The problem for us, as cognitive agents, is that we are moving around and changing our orientation. Thus that photocell doesn't just point at a fixed point in W. So in order to find $\mathrm{greenness}(x)$, we have to carry out some procedure where we control our location and orientation such that we can be sure we are picking up the reflected light from the point x. And if we are doing this before we have much knowledge of the world, then we don't know enough to do anything much in the way of controlled orienting of ourselves. So that makes the perceptual problem even more difficult.

Given a normal Hausdorff space X, mathematicians use the notation $C(X)$ for the system of all continuous bounded complex valued functions over the space X. Thus we may use $C(W)$ to denote the system of all continuous bounded functions over W.

There are infinitely many functions in $C(W)$. As finite beings, we can only expect to be able to evaluate finitely many of them. If we can evaluate $f(x)$ and $g(x)$, we can also evaluate $3f(x) - 2g(x)$ and we can evaluate $f(x)g(x)$. That is, we can, at least in principle, evaluate functions formed from our basic perceptual repertoire by combining them with simple arithmetic operations of addition, subtraction, multiplication by a constant, multiplication (of two or more functions). I shall denote by $P_f(W)$ the set of all functions that can be formed by such finite combinations of our perceptual repertoire. Here the subscript f is intended to stand for *finite*. I shall use $P(W)$ to denote the closure of $P_f(W)$ as formed by allowing also functions that are the limits of uniformly convergent sequences of functions in $P_f(W)$.

Note that $P(W)$ could be the entirety of $C(W)$, or it could be a smaller algebra of functions. If W happens to be a compact Hausdorff space, and if there are enough perceptual functions to distinguish between any two distinct points of W, then $P(W)$ will be identical to $C(W)$, as shown by the Stone–Weierstrass theorem.

4 Facts

Most accounts of what are facts describe them in ways such as ‘bearers of truth’ or ‘true statements’ or ‘statements of what is the case.’ To me, that kind of definition has always seemed unsatisfactory, because it has failed to indicate how facts are connected to the world. In this section, I plan to show how it is our perception (or extended perception) that links our facts to our world.

We start by looking at what appears to be a simple fact:

> The temperature is 75°F.

The problem with this, is that it is incomplete. It does not mention the location and time when the temperature was taken. If we add the missing information, we might have something like

> The temperature was 75°F at latitude 45, longitude 90 at 3 p.m. on June 2, 2000.

Temperature, latitude, longitude, and date-time are all things we would expect to be able to come from extended perceptual functions in P_f. Let's call the corresponding functions f_1, f_2, f_3 and f_4, and lets assume that the time and date can be represened as a number *nnn*, perhaps by representing as seconds past a specified time. Then the expressed fact is equivalent to

$$\exists x \in W; f_1(x) = 75, f_2(x) = 45, f_3(x) = 90, f_4(x) = nnn$$

The important point here, is that the fact is presented as a relation between values of functions in P_f at an unspecified point $x \in W$. If you think about the kind of facts we use, you will see that most of our facts about the world are such relations. The possible exception might be facts that use proper names, and even those still use some of our perceptual functions.

The conclusion we can draw from this is that, in order to have facts, it is necessary to first have some sort of perceptual system. We can, of course, write down a fact in a book. But what is in the book is a representation of the fact, and it takes us to interpret that representation in order for it to be a fact about the world.

5 Realism

The aim here is to understand how it is that we can have knowledge of reality as a whole, beyond isolated facts such as discussed above. That is to say, our primary concern is with the epistemic problem of reality.

To briefly review, we have looked at extended perception as based on a collection of information collecting functions (or measuring procedures) $P_f(W)$, which leads to $P(W)$ when we also include functions that arise as limits of uniformly convergent sequences of those in $P_f(W)$. We saw this as something like $C(W)$, the algebra of bounded continuous functions over the world W (thought of as a topological space).

Given a normal Hausdorff space X, there is a mathematical duality between the space X, and the function algebra $C(X)$ of bounded continuous functions over X. In particular, studying X can inform us about $C(X)$, and studying $C(X)$ can inform us about X.

In the case where X is also compact, thus a compact Hausdorff space, a knowledge of $C(X)$ allows us to reconstruct X, as shown in [3]. That is to say, the algebra $C(X)$ determines the topological space X. For an alternative way of deriving the relation between X and $C(X)$, see [1, pp. 261–281].

In the case where X is not compact, the same mathematics shows that $C(X)$ determines a compact Hausdorff space which has the same function algebra. The original X is embedded in that compact space, so that the space found by this method can be considered a compactification of X. That is, it is a compact space containing X, and we might say think of it as X augmented by some additional virtual points. As an illustration, when we look toward the horizon we see the sky seeming to connect to the ground. If it actually connected to the ground, then we would be in a cosmos. As it is, we can think of the points where the sky appears to connect to the ground as the virtual points that would result in a compactification of our cosmos.

What determines the topological space X is the algebraic structure of $C(X)$, and what determines W is the algebraic structure of $C(W)$. This is particularly interesting from an empiricist point of view. For it implies that, without having any a priori knowledge of the structure of W, simply by developing measuring procedures that can be reliably used on W, and finding the algebraic structure of the resulting algebra $P(W)$, we can hope to discover the structure of W itself. Even if $P(W)$ turns out to not be the whole of $C(W)$, we can still use the same methods to determine a working approximation of the world W.

We should note that the mathematics shows only that W is determined as a topological space. There is an old joke that a topologist cannot tell the difference between a doughnut and a coffee cup (because they are equivalent as topological spaces). What distinguishes a doughnut from a coffee cup are metric properties, rather than topological properties. However metric properties such as length are determined by functions that we would expect to be in $P_f(W)$, hence in $P(W)$, so these too should be determined by $P(W)$, though perhaps not by only using the algebraic structure of $P(W)$.

6 Learning, duality and emergence

In this section, I wish to discuss the mathematical duality between W and $C(W)$, and its relevance to the problem of learning.

If for some point $x \in W$, it turns out that $f(x) = a$, then that relation expresses a factoid about W. I call it a factoid, rather than a fact, since unless we have actual names for x and W it isn't something we could assert as a fact. This factoid, $f(x) = a$ indicates that a is the degree to which the property measured by f is found at point $x \in W$. We can think of f as associating a value a with the point $x \in W$. But we could also describe this in reverse. That is to say, we could look at it as x associating a value a with the point $f \in C(W)$. We could even write that in functional notation, as $x(f) = a$, where we think of x as a continuous function over the space $C(W)$, assigning values to points such as f that are in $C(W)$. It is this reversibility of roles that leads to a duality between W and $C(W)$.

As often happens with mathematical duality, it turns out that a local feature in W is reflected as a widely distributed global 'feature' in $C(W)$. Something of interest about a local point $x \in W$ can involve many different property functions in $C(W)$. Likewise, a local feature in $C(W)$, something about a particular property function $f \in C(W)$ has implications that are distributed throughout W since f can be applied to every point in W. As a result, learning directly about features of W might be seen as a bottom up type of learning, where we start with small items and put them together to give us an understanding of larger items. And learning directly about features of $C(W)$ might instead be seen as a kind of top down learning where we learn by finding increasing detail in W as improved perceptual abilites provide us with that increasing detail.

Learning is the aquiring of knowledge. When it comes to learning, we tend to think of that as gaining a little bit at a time. However, the important question is whether we are picking up little bits of knowledge about W, or little bits of knowledge about $C(W)$. No doubt we do some of each. However, I shall be suggesting that learning is primarily about $C(W)$. That is, our learning is mainly in the form of constructing new perceptual functions in $P_f(W)$, as we attempt to obtain a more and more complete knowledge of $C(W)$. The idea that we learn by acquiring perceptual capacities is consistent with the perceptual learning that was described by Gibson [2].

If learning were in the form of picking up bits of information about W, then we might expect that a child's experience would be one of a world that looked a bit like a jigsaw puzzle with many missing pieced. And the child would learn about the world by trying to fill in those pieces. We do see something like that in the experience of a Sherlock Holmes or a Hercule Poirot when attempting to solve a crime. However, there is no indication that a child's experience is of this kind.

If, instead, learning mostly takes place in $C(W)$, then we would expect a rather different experience for the learning child. Incremental learning in $C(W)$ would correspond to enriching of the world W of experience. Thus we might expect the young child to perceive of a fairly complete world, yet one lacking in the amount of detail seen by an adult. The child might first see all small animals as similar, and perhaps use ”doggie” to refer to them all. But, as the childs perceptual ability grows he would begin be able to distinguish between different kinds of animals. Such a learning pattern seems more consistent with the behavior of young children, suggesting that indeed acquiring more details in $C(W)$ (i.e. acquiring more perceptual functions) is the dominant form of learning.

It is much the same in the growth of science. When scientists investigated electrical phenomena, and made important discoveries in that investigation, it had the effect of adding an electrical dimension to how people saw the world. When Marie Curie discovered radioactivity, that added a radioactive dimension to how the world was seen.

7 Paradigms and incommensurability

In his thesis on scientific revolution, Kuhn [6, 7] argued that some scientific changes are revolutionary, and can involve a shift from the use of one paradigm to another. Furthermore, this paradigm shift may lead to changes such that meanings of terms used under the old paradigm are incommensurable with those used under the newer paradigm.

> The normal–scientific tradition that emerges from a scientific revolution is not only incompatible but often actually incommensurable with that which has gone before. [7, p. 103]

Kuhn's incommensurability thesis was seen as contentious. Indeed, it was seen by some as obviously wrong. It was taken by some as a challenge to our ideas of reality, in that it was seen as suggesting that the scientific paradigm creates reality, thus making reality as a human creation rather than as something that exists external to us.

What I want to discuss here is how incommensurability can be possible, even assuming a fixed reality.

To illustrate what is happening, let us take U to be the unit interval $[0,1]$. Or, said otherwise, U is the set of real numbers x such that $0 \leq x \leq 1$. We might imagine that there are some creatures for whom U is their world. In order to perceive their world, they will need some perceptual functions. They build their perceptual functions based on two functions f_a and f_b, where $f_a(x) = 1$ and $f_b(x) = x$ for all points $x \in U$. Let us denote by $P_f(U)$ the set of functions that can be generated using f_a and f_b as combined with simple algebraic operations of addition, subtraction, multiplication and multiplication by constants. We can see that $P_f(U)$ consists of all polynomial functions (or algebraic polynomials) over U.

Now let us further suppose that the creatures of U make a paradigm shift, wherein they start to use different concepts to describe their world. Instead of using perceptual functions f_a and f_b, they begin to use g_a and g_b where $g_a(x) = \sin x$ and $g_b(x) = \cos x$ for $x \in U$. Let us denote by $T_f(U)$ the set of functions that can be formed from g_a and g_b as combined using simple algebraic operations. It should be clear that $T_f(U)$ consists of trigonometric polynomials over U.

From the trigonometric identity $\sin^2 x + \cos^2 x = 1$ we can see that $g_a^2 + g_b^2 = f_a$. However, most functions in $T_f(U)$ are not in $P_f(U)$, and most functions in $P_f(U)$ are not in $T_f(U)$. Yet the smallest algebra of continuous functions containing $P_f(U)$, and closed under taking limits of uniformly convergent series is $C(U)$, the space of all continuous functions over U. Similarly, the smallest algebra containing $T_f(U)$, and closed under uniform convergence is also $C(U)$. So the earlier $P_f(U)$ and the post paradigm-shift $T_f(U)$ are just different sets of concepts that can be used to describe the same world U. But most of the facts about U that can be represented in terms of $P_f(U)$, cannot be represented as facts in terms of $T_f(U)$.

Using Kuhn's terminology, this difference in expressible facts can be described as saying that, as a result of the paradigm shift there are new problems that can be solved but which could not have been solved with the older paradigm. However, some problems which could be solved with the older paradigm can no longer be solved under the newer paradigm. And the fact that most of the functions in $P_f(U)$ cannot be expressed exactly in terms of functions in $T_f(U)$ demonstrates that there is an incommensurability between the concepts used before the paradigm shift and the concepts used after the paradigm shift, *even though both sets of concepts were used to describe the identical world U.*

When I comment, in agreement with Kuhn's view, that Einstein's relativity changed our concepts of time and space so that they are different from what was used by Newtonians, some people raise their eyebrows and wonder whether I am saying that properties such as time and length are not real. But that is the wrong question. The issue is not whether they are real, but whether they are canonical.[1] Our perception and our scientific measurements determine functions in $C(W)$. But the function algebra $C(W)$ is infinite. We, as finite beings, can only use finitely many of them. We attempt to find a finite system of generators, so that if possible the smallest closed function algebra containing our selected generators is the whole of $C(X)$. And, as best I can tell, there is no canonical way of selecting suitable generators. So those concepts we happen to have chosen as generators are partly reflective of our history. There could be very different and incommensurable ways of selecting such generators for the function algebra.

8 Truth and correspondence

People often talk of truth as correspondence. I wish to analyze that idea. Sometimes people talk of 'correspondence with the facts.' But that seems circular, since we usually take facts to be true statements. The alternative is to talk of 'correspondence with reality.' The meaning of that is not entirely clear, and that is what I hope to clarify in this section.

When we talk of 'truth,' we could be talking of the truth of perceptual information that we receive via our perceptual systems, or we could be talking of the truth of natural language statements we express. Where I need to distinguish, I

[1] I am using *canonical* in the sense in which mathematicians use that term.

shall use the expression *perceptual truth* for the first of those, and *linguistic truth* for the second. I shall mainly be concerned with perceptual truth. Language is a complex social practice that would take us far beyond the intended scope of this report.

As a mathematician, the term ‘correspondence’ suggests to me a mapping of the world into our space of perceptual representation. And, indeed, there does seem to be such a mapping. Specifically, the perceptual functions discussed in section 3 above provide just the kind of mapping that the term ‘correspondence’ suggests.

There is a corollary to this observation. Namely, if truth is correspondence, and if perception actually implements a correspondence, then perceptual truth is pretty much automatic. And I will suggest that this corollary accounts for the observed general reliability of perception. This does not deny the possibility of perceptual error. The perceptual functions involve carrying out procedures, and there is still the possibility of mistakes made when carrying out those procedures.

While I have suggested that truth is automatic for a cognitive agent, I should be clear that this does not apply to the kind of robots that we are likely to build. For those robots do not have their own perceptual systems and their own correspondences with reality. Rather, we build them to rely on our idea of correspondence with reality, because we want those robots to be useful to us in carrying out our tasks. So we design their input systems to provide representations that we see as true. Similarly, we design their output systems to act in ways that we see appropriate when they are acting on data that we see as true.

9 On the reliability of perception

I have suggested that a cognitive agent has the benefit that its perceptual representations are generally true. Yet I have also suggested that robots would would not have the same benefit. I want to further discuss what is involved here, since such a discussion should elucidate the requirements for a cognitive system.

I will start with a thought experiment.

I wish to install a shade on my study window. So I start by measuring the width of the window. I measure it as 79 inches. As it happens, I was quite careless with that measurement. I had picked up a meter ruler, graduated in centimeters. So the measurement of 79 inches was, by traditional thinking, untrue—it should have been 79 centimeters.

Having measured the window, I now proceed to the hardware store, where I can purchase a suitable shade. To be sure that I get a shade of the right size, I will measure it before I pay for it. And, in order to measure that shade, I bring the same ruler with me. After buying a suitable sized shade, I take it home and install it. Everything works out well.

By traditional thinking, my measuring of the window was wrong, and the 79 inches width was false. Similarly, by traditional thinking, my measuring of the shade I purchased at 79 inches was also wrong, but that error just happened to cancel out the original error.

From the point of view used in section 8 above, I was using that ruler to establish a correspondence between my representation of the window's width and the reality of the window's width. And then, provided I followed that correspondence, the one that used the meter ruler, my representations were automatically true because they were in accordance with the correspondence that was in use.

What is required for this to work, is a consistency. The same correspondence must be used throughout. If a cognitive agent can form a representation using the left eye, or can form a representation of the same thing using the right eye, then these must both yield the same representation (to within the degree of precision used). We cannot have one correspondence for the left eye, and an entirely different correspondence for the right eye. Likewise, a cognitive agent must actively percieve its own actions (behavior), and judge those actions using the same correspondence with reality that it uses for other perceptions.

In short, a cognitive agent needs to develop some sort of internal 'standard', and use that internal 'standard' throughout all parts of its perceptual system. That requires some sort of program of cross calibration between the various ways that the perceptual system can form representations of reality.[2]

For science, we do see the development of public standards, usually called *measuring conventions*, and we do see a program of calibrating measuring instruments to these standards. This standardization and calibration program is essential to having a society wide *correspondence with reality* and in turn that is fundamental to the scientific enterprise. Ironically, there is very little mention of its importance in typical textbooks on scientific epistemology.

[2]The observed neural behavior known as *Hebbian Learning* may turn out to be just such a calibration program.

10 The algebraic structure

As discussed in section 3 above, the function space $P(W)$ is an algrebra, and the algebraic properties of $P(W)$ are important if we are to infer the topological structure of W. Thus there needs to be a way to find that algebraic structure.

In this section, I want to examine the question of algebraic structure for scientific information. We actually see that structure in many scientific laws. Thus Newton's $f = ma$ and Ohm's $V = IR$ are algebraic relations which form part of the algebraic structure of our scientific information system, and in turn those become part of our extended perception.

It will be useful to first look at how we find the algebraic structure of a Euclidean vector space (such as 3-dimensional space). This is a simpler question, but illustrates some of the idea involved.

We often describe 3-dimensional space in term of coordinates. Thus we give the coordinate values (x, y, z) for a point in space. These coordinate values are themselves continuous (and linear) functions over space, thus illustrating the role of continuous functions as discussed above in section 3.

We can look at the coordinate representation in terms of unit vectors. Thus we suppose unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ in the directions of the x, y, z axes. Then the location (x, y, z) represents the vector $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

These unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are actually arbitrary choices, subject only to the restriction that they be of length 1 and be mutually orthogonal. It is usual to select the z-axis, and hence the unit vector $\mathbf{k}$ to be vertical. However, the direction of vertical, as selected by me near Chicago, will be different from the direction of vertical that might be selected by somebody in New York or in London or in Sydney. That we all choose different directions illustrates that the choice of unit coordinate vectors is non canonical, and the choice is made on pragmatic grounds. Likewise, I would normally select the x-axis, and hence the vector $\mathbf{i}$ to extend from left to right. That's again an arbitrary choice based on the direction my desk happens to be facing (i.e. on pragmatic grounds).

There is a general method to find orthogonal unit vectors. It is known as the Gram–Schmidt orthogonalization process. The idea is that you pick a nonzero vector. Then you rescale to be length 1. Next you pick another vector that is not in the space generated by the vector(s) already chosen. You subtract off the component in the direction of the unit vectors already selected, resulting in a vector that is orthogonal to those already selected. Then you rescale that to length 1, and make it your next selection. Continue until you have enough unit vectors to generate the whole space, at which point you are done.

We now proceed to the problem of finding the algebraic structure of $P(W)$. Here's a method that, in principle, should work. We start with a function in $P(W)$, and then we scale that suitably so as to make it mathematically 'nice.'

Next we look at the function algebra generated by the function(s) we have already selected. If we can find another function in $P(W)$ that is not generated by what we have already selected, then we subtract off or factor out any part of this 'new' function that is better considered to be part of the already selected functions. Next we suitably scale the resulting function, and if possible we do so in a way that puts it in a 'nice' mathematical relation with the functions previously selected. We continue in this way, until done.

If we look at the history of science, we see scientists doing something similar to what I have described. When they come upon a new discovery, they attempt to factor out or subtract off what appears to be explained by previously known science. Then they develop a measuring convention for the newly discovered property, and if possible they do so in a way that will have a 'nice' mathematical relation with previously known properties.

11 What are the neurons doing

Most of what is in this section is somewhat conjectural. I am using the principles presented in the prior sections of this report, and suggesting how they might be implemented by the neural system. I will present this as a series of questions with suggested (conjectured) answers.

Does the brain use numbers?

In one sense, the brain obviously uses numbers. We use numbers, and the brain does what is needed to make that possible. But there is a different sense, namely the question of whether the brain is doing something like arithmetic in its low level operations. I suspect that the answer is *no*. It is not clear that there is any great need to use numbers.

In section 9 above, I gave a thought experiment about using a meter ruler to find the width of my window. I measured it at 79 inches (but that should really have been 79 centimeters). I did not actually need a number at all. The rule is divided up with graduation marks. I only needed to remember which of those graduation marks corresponded to the width of the window. And then, when later measuring the shade I was purchasing, I could check that against the same graduation mark.

The suggestion is that numbers are not needed, but graduation marks are. Roughly speaking, we humans use numbers in order to publicly communicate and discuss our measurements. But for a self contained system such as the brain, just having the graduation marks should be sufficient.

It is well known that the neurons work by 'activating' when their inputs have reached a certain threshold level. I suggest that graduation marks are represented in the neural system as such threshold levels.

Does the brain do algebra?

If the brain is not using numbers, then it is unlikely that it is doing algebra. However, an algebraic relation such as $f = ma$ can also be described as a functional relation $f - ma = 0$. This is a particular case of a functional relation of the form

$$F(f_1, f_2, \ldots, f_n) = 0$$

where F is a function acting on other function. The suggestion is that the brain is implementing such functional relations in the form of piecewise linear functions. If we think of the brain as using graduation marks, and think of using linear interpolation between graduation marks, then such a piecewise linear function would be implemented by connections between neurons. The appropriate connections could be constructed as part of perceptual learning.

How does perceptual learning work

As discussed in section 9 above, there needs to be a consistency throughout the neural system. For example, information received via the left eye needs to be consistent with corresponding information received via the right eye. That requires some sort of cross calibration between the two distinct input systems. Presumably there are neural connections between the two that can compare the inputs from the two eyes, and adjust the calibration of one or the other, as needed to bring them into compatibility. Those connections would be just the kind of connection needed for the piecewise linear functional relations as discussed above. And the brain seems to have plenty of connectivity to allow for this.

As the brain attempts to cross calibrate information received by the two eyes, it will discover that there are some discrepancies that cannot be removed simply by suitable calibration. These discrepancies are ‘new information’, and presumably the brain learns to use those for visual depth perception.

The overall idea is that a program of cross calibration will look very similar to what is observed as Hebbian learning [4], and that cross calibration will discover discrepancies that cannot be eliminated. These discrepancies can then be used as the basis for new perceptual functions.

12 Summary

under preparation

References

[1] Nelson Dunford and Jacob Schwartz. *Linear Operators. Part I: General Theory*. Interscience Publishers, 1958.

[2] Eleanor J. Gibson. *Principles of Perceptual Learning and Development*. Appleton Century Crofts, 1969.

[3] Leonard Gillman and Meyer Jerison. *Rings of Continuous Functions*. D. Van Nostrand, 1960.

[4] Donald O. Hebb. *The Organization of Behaviour*. John Wiley & Sons, New York, 1949.

[5] John L. Kelley. *General Topology*. D. Van Nostrand, 1955.

[6] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. The University of Chicago Press, 1962.

[7] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. The University of Chicago Press, second edition, 1970.