THE UNIVERSITY OF CALGARY

# Information Conveyance:
# A Critique of Claude E. Shannon's
# Noise/Signal Distinction

by

Dwayne F. Raymond

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF PHILOSOPHY

CALGARY, ALBERTA

AUGUST, 1997

0-612-24612-4

Abstract

The conditions required for Claude E. Shannon's notion of information and measure of information are investigated. It is shown that both the measure and notion of information require specific conditions to be applicable to actual systemic activity. Moreover, it is argued that Shannon's account of information conveyance necessarily implies a necessary noise/signal distinction. While these conditions and the noise/signal distinction appear to exist in some physical systems, they are shown not to be present in neural systems. Hence, Shannon's information theory is limited in its scope of application. If the claim that neurons convey information is to hold then the notion of information must be other than Shannon's notion.

I am indebted and grateful to both Dr. John A. Baker and Dr. C.B. Martin. If not for Dr. Baker's insightful comments, criticisms and patience, this document would be less than it is. Dr. Martin's ontology prompted me to begin the inquiry and his faith helped me to finish.

Table of Contents

List of Figures

## List of Symbols and Abbreviations

| | |
|---|---|
| $\Sigma$ | The sum of |
| 5HT | Serotonin |
| cGMP | 3'-5' cyclic guanosine monophosphate |
| $Ca^{2+}$ | Calcium Ions |
| DA | Dopamine |
| log | logarithm |
| Oct | Octopamine |
| SRC | Strong Rigidity Requirement |
| STG | Stomatogastric ganglion |
| TDN | The Definition of Noise |
| TNI | The Notion of Information |
| TTP | The Transmittability Principle |
| TRR | The Reproducibility Requirement |
| WRC | Weak Rigidity Requirement |
| $WRC^{*}$ | Weak Rigidity Requirement Two |

It is in fact, nothing short of a miracle that the modern methods of instruction have not yet entirely strangled the holy curiosity of inquiry; for this delicate little plant, aside from stimulation, stands mainly in need of freedom; without this it goes to wrack and ruin without fail. It is a grave mistake to think that the enjoyment of seeing and searching can be prompted by means of coercion and a sense of duty.

Albert Einstein, *Autobiographical Notes*

# Chapter I

## Introduction[1]

### 1 Synopsis

Shannon's account of information cannot be made to fit all physical systems in which we would want to say that information is conveyed.[2] In the following pages, I develop this claim by first establishing the conditions required by Shannon's account of information and, later, by looking for such conditions in the actual operation of biological systems. I argue more specifically and more importantly that if these conditions must be satisfied then Shannon's account of information cannot fit biological systems, for, as I argue, the conditions are *not* satisfied in biological systems, for such systems are in a sense to be identified dynamic and Shannon's conditions cannot be satisfied by (in this sense) dynamic systems. In other words, I will argue that while Shannon's account works well for certain kinds of physical systems it does not work for dynamic biological networks. Finally, and perhaps most interestingly of all it is for this reason that Shannon's measure

[1] I am indebted and grateful to J.A. Baker for numerous comments and suggestions concerning style, clarity and content, in this as well as the other chapters. The forgoing presentation would not be what it is without Baker's dedication and many laborious hours, on this and multiple drafts.

[2] Shannon, C. E. (1948). "The Mathematical Theory of Communication." *Bell System Technical Journal 27*: 379-423. Reprinted in Shannon, C.E. and Weaver, W. (1963). *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press.

of information is not applicable to human capacities. In the inquiry to follow it will be vital to bear in mind throughout that Shannon wants to explain the phenomena of information and information transmission using building blocks which are themselves non-semantic. That is, he wants to explain information and information conveyance in terms which themselves make no reference to meaning or its cognates. My critique of his views do not in fact question *this* aspect of his approach; therefore, my critique is logically independent of this feature of his theory.

## 2 Preliminaries

Recent advances in non-invasive techniques for recording brain activity have spawned a resurgence in computational or information processing based explanations of neural activity.[3] For example, in *The Oxford Companion To The Mind* it stated that,

---

[3] Examples of non-invasive brain recording techniques *may* include, but are not limited to, magnetic resonance imagery (MRI) and electroencephalogram (EEG) recordings where the electrode is placed on the surface of the skull. Invasive techniques include, but are limited to, the placement of depth electrodes for EEG recordings; lesion studies, where the lesions are created by the investigator; and positron emission topology (PET); and, this list may include MRI. The distinction between invasive and non-invasive techniques is problematic. While surface EEG recordings is the clearest example of non-invasive, MRI involves a form alteration to neural activity and may therefore be seen as invasive or non-invasive, depending on the exact definitions employed. For example, a PET scan combines both computerized topography and radioisotope imaging. In emission topography the recorded image reflects the distribution in tissue of an injected or inhaled isotope that emits radiation. Since the activity in a nerve cell is related to the cells use of glucose, the activity of a neurons' can be mapped by recording the glucose metabolism of nerve cells. An analog of glucose which is not fully metabolized within an active neuron can be used to recorded the glucose utilization in small areas of the brain. The introduction of an isotope that emits radiation is invasive as a foreign substance is introduced to the brain. Although the MRI dose not involve the introduction of a substance, MRI does involve the introduction of a radio wave to the brain. The images generated with MRI are the result of perturbing the spin axis of an element (such as hydrogen) with a brief pulse of radio waves. "When elements with an odd atomic weight, such a hydrogen, are exposed to a strong static homogeneous magnetic field, the nuclei behave as spinning magnets and develop a net alignment of

> It is now commonplace to regard the impulses that flow along nerve fibers as 'conveying information'; but just how information about the world is represented in the brain remains an unresolved question.[4]

And, in *Principles of Neural Science,* a notion of information is also employed.[5] John H. Martin, the author of chapter 23 writes:

> Sensory systems receive information from the environment through receptors at the periphery of the body and transmit this information to the central nervous system.[6]

While I do not question the correctness of Shannon's account of information so long as

---

their spin axes along the direction of the applied field....... [w]hen the pulse is turned off, the nuclei tend to return to their original orientation, and in so doing so release energy in the form of radio waves."(Martin, John, *et al.*, p. 317). MRI could be seem as either invasive or non-invasive, depending on the criterion for the distinction. What is required for this thesis is that there exist techniques that enable imaging the living brain and not merely lesion studies which destroy the areas of the brain. I am indebted the C.G. Teskey for his comments and clarification concerning the nature of various recording techniques.
An early example of a computational explanation of neural activity is found with John von Neumann. His intended topic for the *Silliman Lectures,* at Yale in the spring of 1956, was about the logical structure of the nervous system. Professor von Neumann passed away February 8, 1957, before giving the lecture. Early work on information processing and biological feedback systems is found in an area of study known as Cybernetics. The writing of Norbert Weiner (1948) details the history and development of both cybernetics and information theory. From his writing (and others, Gardner (1985)) it is clear that information theory and computation accounts have, from the beginning, overlapped in terms of some central ideas. For example McCulloch and Pitts (1943) first developed a symbolic conception of interconnected networks as part of the cybernetic movement; now their basic model is referred to as being paradigmatic of an alternative '*style*' of computation to von Neumann or Turing. It is, in part, due to this connection that this thesis is able to broach both information theory and vector transformation models, the underling assumption of both are very similar; indeed, I argue in chapter four they are the same at critical points. Chapter four presents a clearer and fuller discussion of this. See: Gardner, H. (1985). *The Mind's New Science: A History of the Cognitive Revolution*. New York: Harper Collins; Martin, J.H., Brust, J.C.M. and Hilal, S. (1991). "Imaging the Living Brain" in Kandel, E., Schwartz, J. H. and Jessell, T. M. (eds.) (1991). *Principles of Neural Science. 3rd ed.* NewYork: Elsvier. pp. 309-324.; McCulloch, W. S. and Pitts, W. (1943). 'A logical calculus of the ideas imminent in neural nets' *Bulletin of Mathematical Biophysics 5.* von Neumann, J. (1957). *The Computer and the Brain.* New Haven: YUP; Wiener, N. (1948). *Cybernetics.* New York: MIT Press.

[4] Gregory, R.L.(ed.). (1987). *The Oxford Companion to The Mind.* Oxford: OUP. p. 370.

[5] Kandel, E., Schwartz, J. and Jessell, T. (*eds.*). (1991). *Principles of Neural Science. 3rd ed.* New York: Elsevier.

[6] Kandel, E., Schwartz, J., and Jessell, T. (eds.) (1991). *Principles of Neural Science. 3rd ed.* New York: Elsevier. p. 330.

the account is not taken as being applicable generally: I will argue indeed, as I said above, that it cannot be taken as applying to information conveyance in biological systems. Hence I also do question any suggestion that it is a *commonplace* that nerve impulses can be counted as conveying information, if information in *this* statement is taken to be explicable using Shannon's theory.

While the systemic conditions Shannon assumed may hold for many physical systems there in no reason to believe they hold for all physical systems in which the notion of 'information' is applied. Since Shannon was working on a theory of communication and his diet of examples were limited to physical systems designed -- actually or theoretically designed -- for *that* purpose, there is reason to investigate the importance particular systemic structures have for Shannon's notion of information.

In this thesis I argue that a Shannon-type approach to information, *viz.*, a reductionist approach with resources limited in the way he limits them, that is, an approach which excludes use of the notion of meaning, will fail. I do not argue that the account will fail because it cannot explain meaning nor do I claim that at some point in his theory he cannot explain meaning. Rather, I argue that there are OTHER problems facing his account. While he can explain information transmission in some systems, I argue he cannot explain information transmission when the medium is a biological system as complex in various ways, for example plasticity, adaptability and modulatory as the neuronal system. Therefore, I allow that he can give at least an interesting account of information for something simple like a serial digital computer. However, this is important because -- of course -- what we want is an account of information transmission for biological systems as well. SO, minimally, even if Shannon could claim to have a theory of information for serial digital devices, this is not an account which can be generalized for biological systems.[7]

Since understanding Shannon's quantitative account of information requires understanding his measure of information, I begin with a discussion of Shannon's measure of information. I argue that Shannon's account of the measure of information cannot be construed literally and applied to human capacities. The two possible justification for this claim are (1) we lack an ability to obtain data on all requisite aspects and hence cannot without radical rewriting apply Shannon's notion; or (2) there is a deeper problem in applying Shannon's measure. Shannon's account of information is not appropriate for neural systems. I argue, in chapter three, that the latter is the reason why Shannon's measurement of information is not applicable to human capacities.

In the fourth chapter I consider a possible extension of Shannon's ideas, the extension which involves using vector transformation models of artificial networks. It might be thought that this extension might save Shannon's account from the kinds of problems I raised against the theory without this extension. In chapter III I argued that the theory without this extension requires a noise/signal distinction: in chapter IV I argue that even

[7] I am indebted to J.A. Baker for stylistic comments and clarification of some of my ideas.

with the extension the theory still needs a noise/signal distinction and indeed the same one as was provided for the theory without the extension. I argue that this means that the theory with the extension faces very similar objections to those I brought against the earlier version of the theory and hence that the extended theory cannot stand.

As an aside (on the issue of the status of my inquiry) the following comment is prompted by the extravagant claims made by modern theorists in this area concerning the nature of the inquiries we are embarking on. The combination of modern scientific techniques and advances in computational theory and resulting theories have prompted *Nobel Laureate* Francis Crick and philosopher Patricia Churchland to independently proclaim that few philosophical puzzles remain in this area of research (brain research) fast becoming the sole providence of scientific investigation.[8] Crick has gone so far as to write a non-scientific book titled, *The Astonishing Hypothesis: The Scientific Search For The Soul.*[9] Nonetheless, in spite of technological innovations -- moreover because of them -- there are philosophical problems that remain. I discuss one of them here: the problem of noise/signal discrimination in Shannon's account of 'information'. I will show that the problems of noise and signal when applying Shannon's account to biological system's result from conceptual difficulties in how Shannon develops the notions and not merely technical problems of actually distinguishing between them.

---

[8] See: Churchland, P.S. (1990). *Neurophilosophy: Toward a Unified Science of the Mind/Brain.* Cambridge: MIT Press; Crick, F. (1994). *The Astonishing Hypothesis: The Scientific Search for the Soul.* New York: Scribner.

[9] Crick, F. (1994) *The Astonishing Hypothesis: The Scientific Search For the Soul.* New York: Scribner.

# Chapter II

## The Measure of Information and the Case of Adaptation in Phototrasduction

### 1 Introduction

In this chapter I will outline the theory offered by Claude E. Shannon (1948) concerning the nature of information.[1] Shannon's theory is an exceedingly, indeed heroically, economical theory. It is economical both ontologically and conceptually. He analyses information, amazingly, using in the end only a notion of cause and effect and probability, or at least unexpectedness. He does not directly employ the terms 'cause' and 'effect'; instead, he discusses such activity in the context of a communication system consisting of five basic elements: an information source, a transmitter, a channel, a receiver and a destination. Even the briefest examination of these notions as used by Shannon reveals, however, that they are each to be understood solely in causal terms: there is, for example, in these notions no component notion which would count as a semantic notion. These points are worth making but again it should be remembered that I will not be objecting to Shannon's theory on the ground that semantic notions are not found in its building

[1] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal.27.* pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press.

blocks. Hence, these comments are at most meant to orient the reader to the nature of Shannon's theory.

It is my view that his theory is simply too economical: he does not have the tools for an adequate theory of information. But it is worth looking closely at his theory because such an examination will reveal, in stark clarity, the points where his theory needs supplementation. We will in other words see what needs to be added for an adequate theory.

My primary aim then in this and the next chapter is to use Shannon as a peg on which to hang a list of prerequisites for any adequate theory of information. The list of problems which I will argue Shannon faces and which he cannot deal with becomes a list of desiderata for any adequate theory.

In general my objective is to show that Shannon's theory simply cannot be supplemented in the way I will have argued it needs to be supplemented and this reveals my deeper aim which is to show that although his theory does seem initially attractive, given a bit of supplementation, in the end NO supplementation will do, for the supplementations in the end are incompatible with his basis idea. In other words, my thesis will be negative - an extended argument in favour of giving up on this basis approach and approaches like it.

To this end, I have two immediate aims: the first aim is to show that Shannon's measure of information -in terms of binary digits or *bits* for short- cannot without radical change be applied to all physical systems. Particular emphasis will be given to the kinds of activity involved in networks of neurons (the networks I refer to are biological networks and not artificial networks). The second aim is to show that the reason for this is part of a deep problem with Shannon's basic conception of information. Hence my contention ultimately is that the problem I will discuss limits the application of Shannon's notion of information (and hence his measurement) to specific kinds of physical systems i.e. physical systems which do not operate in the manner neural systems operate. The first aim will constitute much of this chapter; whereas the latter aim is achieved through discussion in this and the following chapter. I begin with a brief outline of Shannon's notion of information and follow this with an account of his measure of information.

## 2 *Information: The Theory and The Measurement*

### 2.1 *Shannon's Notion of Information.*

Shannon's notion of information as well as the conditions for the transmission of information is not, strictly speaking, defined in Shannon's work; rather, his notion of information and his notion of noise are implicit in his definitions of a communication system. In essence the idea is very straightforward. It could be formulated thus:

> information transmission (or information conveyance) occurs when a cause [the input] produces an effect [the output]. Therefore to specify the information one needs five terms, i.e., we have a quintuple, viz. <an information source, a

transmitter, the channel, the receiver, the destination>[2]

Taken together, these five terms can be used to define Shannon's communication system and the conditions under which information is transmitted or conveyed.

1. An *information source* [is that] which produces a message or sequence of messages to be communicated to the receiving terminal;
2. A *transmitter* [is that] which operates on the message in some way to produce a signal suitable for transmission over the channel.
3. The *channel* is merely the medium used to transmit the signal from transmitter to receiver.
4. The *receiver* ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.
5. The *destination* is the person (or thing) for whom the message is intended.[3]

Using these five ideas, Shannon suggests the following account of information. He says, firstly, that the conditions under which information is transmitted set the boundaries for what counts as information. Specifically, information is something capable of being (i) transmitted and (ii) received in the same form that it was transmitted. His idea seems to be that the *amount* of information transmitted is a function of how surprising it is that what was received was received, given the form in which it was transmitted: the more surprising it is that the information *was* received, given the form of transmission, the more information has been transmitted. That is, amount of information is an inverse function of the likelihood of transmission. By "form of transmission" in the above,

[2] This formulation is mine not Shannon's.

[3] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal27*. pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. pp. 34-35.

Shannon means "order of transmission of the elements of what was transmitted." Accordingly, what will be *surprising* in a transmission is that the elements were transmitted in some specific order. This point can be summarized thus,

> the AMOUNT of information received in a given transmission is a function of the unexpectedness of what is received. If what is transmitted is sequential, the sequence has order and this order decreases the unexpectedness and hence decreases the amount of information. This is despite the fact that understanding maybe achieved.

Shannon's believes that his account of a communication system with its implicit notion of information has the resources to provide an account of content and how content is preserved in a physical system AND this account is reductionist: that is, to tell the story he does NOT need any prior notion of meaning or its cognates. Shannon tries to achieve by fulfilling the need which is normally filled by an appeal to content by an appeal to the internal structure of messages. It is in this way and for this reason that Shannon tries to avoid having to provide an account of content.[4]

Understanding what Shannon means by *the order within a message* requires an understanding of Shannon's measure of information. In short, the order (the internal structure) is a statistical dependency that successive elements have on proceeding

---

[4] C.B. Martin has drawn my attention to a distinction between kinds of perception. The first kind of a perceiving is the perceiving of an object *as* something (propositional perceiving). The second kind of perceiving involves the perception of an object but not as anything (non-propositional perceiving). If one accepts that not all 'seeing' are seeing as then it is not clear how Shannon's account can allow for the same input to be both kinds of perceivings. This distinction is lost to Shannon's notion of information. While important, this discussion is not directly related to the discussion at hand. However, I do mention it as an important point; if the distinction holds, the distinction will establish a second and independent line of objection to Shannon's information theory. See: Martin, C.B. (unpublished) Lectures Notes, Philosophy 695, Winter, 1994.

elements within a single sequence.

## *2.2 Shannon's Measure of Information*

Shannon's measure of information is a logarithmic measure of the weighted probability attached to *all* possible alternatives, individually or sequentially, which could be manifested. In other words, it is not merely the alternatives that *are* manifested which matter here: it is *all* of the alternatives that *could be* manifested.

Shannon's measure of information takes into consideration what is fundamental to his notion of information: *viz.*, that there is a statistical structure among the elements of any message. According to this theory of information, each member in the sequence is dependent upon the member which proceeds it. Shannon's measure of information is therefore a weighted measure of the probability attached to each possible alternative. It is expressed by the general equation.

$$H = -K \sum_{i=1}^{n} p_i \log p_i$$

where K is a positive constant, $H$ is the amount of information and $p$ is probability of an outcome. Since the probabilities $p$ are fractions of 1, the negative sign is required to keep the amount of information positive, as the log of a fraction is a negative.[5]

What follows is a detailed account of what this equation comes to and why Shannon chooses a logarithmic measure of probability for his measure of information.

Shannon's measure of information is said to be based on information theorists' intuition that

> an essential notion for quantifying information is that the less likely the symbol or event, the more information it conveys. Information rate is not simply defined in terms of the number of symbols that can be transmitted, but also in terms of the probability of their occurrence.[6]

The simplest case of alternatives is the case where two alternatives exist - binary cases. A case where there is either a cup on the table or no cup on the table is an example of a binary case. How does Shannon quantify the information contained in this situation and what is the quantity of information in this situation?

---

[5] The presence of the negative sign has caused some terminological confusion on the behalf of two theorists. Norbert Wiener (1961) wrote "Just as the amount of information in a system is a measure of its degree of organization, so the entropy of a system is a measure of its degree of disorganization; and the one is simply the negative of the other." On the other hand Warren Weaver (1963) writes as if information and entropy are equivalent when we writes "This situation is highly organized, it is not characterized by a large degree of randomness of choice - that is to say, the information (or the entropy) is low." While this may be seen as a conceptual difference between the views, it is not. Nothing turns on it; the two writers present the same theory without further variation, in the end. See: Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p.13; Wiener, N. (1961). *Cybernetics.* New York: MIT Press. p. 11.

[6] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal.27.* pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p.9.

How does Shannon measure information? Shannon, who followed Hartely's (1928) lead, adopted a logarithmic measure of information.[7] Shannon claimed that a logarithmic measure is appropriate for three reasons:

> 1. It is practically more useful. Parameters of engineering importance such as time, bandwidth, number of relays, etc., tend to vary linearly with the logarithm of the number of possibilities. For example, adding one relay to a group doubles the number of possible states of the relays. It adds 1 to the base 2 logarithm of this number. Doubling the time roughly squares the number of possible messages, or doubles the logarithm, etc.
>
> 2. It is nearer to our intuitive feeling as to the proper measure. This is closely related to (1) since we intuitively measure entities by linear comparison with common standards. One feels, for example, that two punch cards should have twice the capacity of one for information storage, and two channels twice the capacity of one for transmitting information.
>
> 3. It is mathematically more suitable. Many of the limiting operations are simple in terms of the logarithm but would require clumsy restatement in terms of the number of possibilities.[8]

The measure of information involves two components, first the probability of an occurrence and second the logarithmic measure of the probability. To see how this works recall the case of a cup on a table; further, assume that the probability of there being a cup on the table is $p$ and the probability of there not being a cup on the table is $q$ where, $q = 1 - p$ and $p$ = or < 1. The quantity of information in this situation (where there is one instance of two alternatives) is expressed by Shannon's equation:

---

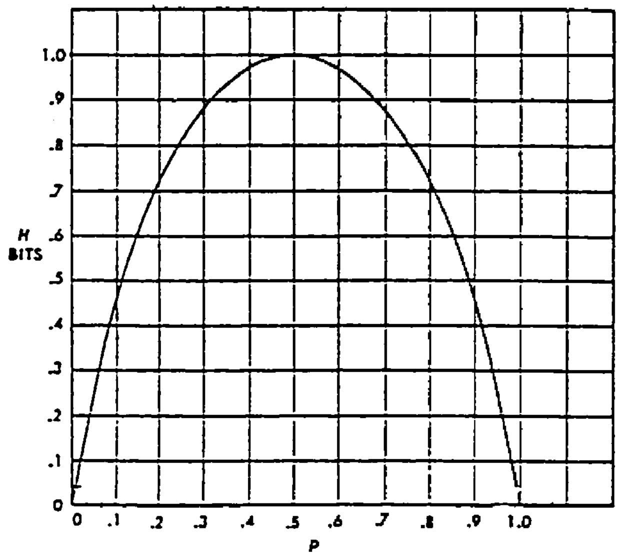
[7] Hartely, R.V.L. (1928). 'Transmission of Information' *Bell System Technical Journal.* p.535.

[8] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal 27.* pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p. 32

$$H = -(p \log p + q \log q)$$

where H is the amount of information and $p$ and $q$ are the respective probabilities.

To measure the actual quantity of information in the situation we need to know the values of $p$ and $q$. All potential values of $H$ (quantity of information) can be expressed graphically. Shannon does just this when he plots the amount of information (in single binary cases) in his figure, presented below.




— Entropy in the case of two possibilities with probabilities p and (1—p).

(Shannon, C. (1948), p. 50)

Shannon's graphical representation demonstrates how his measure of information captures (in part) his intuition, *essential to the notion of quantifying information is the idea that the less likely the symbol or event, the more information it conveys.* $H$ is at its lowest value when the presence or lack of presence (of the cup) is certain and *H* is at its highest value when, the probabilities *p* and *q* are equal at 0.5. That is, where there is complete uncertainty about which of the two alternatives will happen there is the most information in the situation.[9]

The limiting case, where there is no alternative, captures this intuition when a logarithmic measure of information is used. In such situations the probability for one of the two alternatives is unity and hence the amount of information in such a situation is zero. In general then, the amount of information increases as the uncertainty between the two alternative increases; it is at a maximum when the probability between each is equivalent, that maximum for single case of two alternatives is equal to one. The unit of measurement when binary cases are looked at is called the binary digit or *bit* (for short).

The importance of this is to show just how sensitive Shannon's actual measure of information IS to changes. In fact, as is clear from the graph, the measure of a single bit of information will itself change as the probability of either outcome is altered. This is an often overlooked aspect of Shannon's measure of information, though nonetheless it is

[9] Strictly speaking the measure of information, Shannon's provides, does not fully capture the intuition. For example, the situation where p = .9 and q = .1 given the formula is said to have little information in it. However, if the alternative with the lowest possibility were to happen surely, in keeping with the intuition, more information would be expressed then if the probabilities p and q were equal at 0.5. While this kind of consequence might prompt one to suggest that a caveat is required, namely the measure assumes that the event with the greatest possibility manifests, the information theorist would point to fact that while this is true of the single event it is not true of the measure of all possible events taken together. Hence, it is in this way and for this reason that information theory is said to be a measure of the situation as a whole and not the actual individual choice which manifests.

part of his measure of information. The reasons for drawing this out will be made clear in the discussion that follows; the measure of information will be applied to the responses of a biological system. The kind of change which alters the probability of an outcome will be argued to be basic to the operation of the system under investigation. Before turning to this discussion a more complete account of Shannon's measure of information will presented. I continue with the reason Shannon's equation is in the form it is.

Shannon claims that his general equation (see page. 12 of this thesis) is the only equation that satisfies the requirement that the probability of each successive process depends upon the probability of the previous events (generally such a process is known as a Markoff process or a Markoff chain). Shannon's measure of information is designed to exploit this point. It is important to be clear that the dependence of probabilities upon what is gone before is NOT an ontological dependency: it is a dependency existing as required by the structure of a code. ***It is a code which is fully structured at the point of transmission***.

> As the successive symbols are chosen, these choices are, at least from the point of view of the communication system, governed by the probabilities; and in fact by probabilities which are not independent, but which, at any state of the process, depend upon the preceding choices.[10]

The structural dependency is built into Shannon's account of information and of how it is

[10] Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. pp.10-11.

to be measured. The case cited by Weaver (1961) concerns the English language. He notes that in English there is no word in which the letter j is followed by the letters b,c,d,f,g,j,k,l,q,r,t,v,w,x, or z; the probability therefore that the letter j immediately precedes one of these letters is zero.[11] In this manner, the letter which follows j is not independent of the letter j; hence the probabilities for the letter which follows j is not independent. In general, Shannon's measure of information is based on the assumption that in sequences of binary elements there is structure and this structure implies, according to Shannon, that there is less freedom of choice among the elements; hence the probabilities of what can follow a given element are dependent upon what the element is. The freedom of choice is not independent; hence the measure (which measures degree of unexpectedness) reflects this fact.

As an aside, the foregoing discussion (and discussion which follows) has been structured to separate Shannon's measure of information and his notion of information from information transmission (conveyance). The purpose of doing this is allow Shannon to modify his measure to account for systemic activity which is not linear and sequential in the way Shannon's measure requires. More specifically, it is not clear that the structure of Shannon's measure allows for distributed activity. This follows from the fact that the notion of a Markoff chain is a notion of order in which there are statistical relations between successive elements; whereas, in an architecture whose activity is distributed the

[11] Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. pp. 11.

magnitude of all inputs combined is important and not the order the input is received. While an argument could be made that Shannon's measure is problematic given what has been stated (and not argued), such an argument is not the argument presented here. This kind of an argument would allow Shannon a response that his formula need only be altered to accommodate for such. While the notion of a Markoff chain is important to Shannon's current measure, it is not central to his entire account of transmission or conveyance. The criticism that is being sought in this thesis is of a deeper nature. It is a criticism which looks at the assumptions and foundations which Shannon built his measure from and not just different variations which could result with the same set of assumptions. What I take to be at the foundation of Shannon's theory is his account of information conveyance. Full discussion of this appears in chapter three. Again this is made clear over the entire thesis. Before beginning with the critique of Shannon's measure, a brief account of a common and simplified formula for measuring information will be presented, finishing off the exposition of Shannon's measure.

In situations where there are many alternatives that could be manifested (32, for example) and ALL of these alternatives are equally likely to be manifested, the received wisdom is that the amount of information can be simply calculated by using a simple logarithmic formula (which excludes the measure of probability). Consider the expression

$$m^x = y$$

'$x$' is said to be the logarithm of $y$ to the base of $m$. Since a bit is a measure of binary situations the base we are considering is two. Consider a case where there are two alternative situations, each of which is equally probable. In this situation the 'short' equation can be used where, $2^1 = 2$. The logarithm of 2 to the base 2 is 1; hence there is one bit of information. If there are 32 equally probable alternatives, the contained amount of information is measured as $2^5 = 32$ or 5 bits of information.

What is important is the requirement for the use of this shortcut in measuring information. The requirement is that all the alternatives be equally probable. Clearly, the use of the short cut is warranted when and only when the conditions of the situation are in accordance with the assumption of equally probable alternatives. The importance of this obvious point will be made clear in discussion which follows.

In what follows, an inquiry into the assumptions required for Shannon's measure of information will be undertaken. It will be my claim that Shannon's assumptions do not hold for neural systems. In particular, my claim will be that Shannon's measure of information and, moreover, his notion of information assume that the number of *alternatives* exist in a specific way in physical systems. It is true that neural systems are structured in such a way that we could imagine the sequential states of a neural system as unfolding in various alternative ways. Nevertheless, such 'alternative sequential states' (i) are because of Shannon's assumption excluded for consideration, and (ii) are not compatible with Shannon's basic conception of information.

## 3 *Objections to Shannon's Measure of Information*

The claim that there are problems in rigorously applying Shannon's measure of information to the brain and human capacity is not unique. In 1986 Richard L. Gregory argued that Shannon's measure of information is seldom rigorously applied to neural systems because we seldom if ever have the necessary information required for its application. He noted that:

> Unlike normal causes, it [the quantity of information] depends not only on what has been, and what is, but also *on alternatives of what might be*. ....... In order to apply information theory rigorously it is necessary to know the number of alternatives from which selections are made. Unfortunately we seldom if ever know just what these are for humans, so information theory can seldom be rigorously applied outside engineering situations where we have full knowledge of the system and especially its range of alternatives.[12]

Gregory's argument challenges not Shannon's theory, but our capacity to employ it. Since Gregory accepts Shannon's basic assumptions (concerning systemic activity), if these assumptions do not hold, Gregory's argument is moot. However his argument is worth starting with, as it serves as a stepping stone into the discussion of systemic change in which both the number of alternatives and the probability of the alternatives can change. Gregory's comments, as will be clear do not address what was just stated.

---

[12] Gregory, R. L. (1986). 'Whatever Happened to Information Theory?' in Gregory, R. L. (1986). *Odd Perceptions*. New York: Methuen. p.191. I have added the phrase 'the quantity of information'.

Gregory's comments are appropriate because Shannon's measure of information is a measure of both the number of elements and the weighted probability of each element; That is, Shannon's measure of information is a measure of the situation as a whole and not of individual occurrences.[13] In fact, it is for this reason that the measure of information,

> applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which it is convenient to regard as a standard or unit of measure.[14]

Given this, it is easy to see that to properly apply the measure of information one requires full knowledge of the system and especially its range of alternatives. While this is possible with systems we design (insofar as we can build in a range of alternatives) the same, Gregory correctly points out, cannot be said of neural systems.[15]

What should be clear is that his criticism is not a deep criticism about Shannon's measure of information. It is rather the acknowledgment of our limited knowledge (or lack of knowledge) of one aspect (number of alternatives) needed to employ the measure in a

---

[13] For a discussion of this point see note 9 in this chapter.

[14] Gregory, R. L. (1986). 'Whatever Happened to Information Theory?' in Gregory, R. L. (1986). *Odd Perceptions*. New York: Methuen. p.190.

[15] We may in the end even be wrong about the number of alternatives that we have built into the system. If we are wrong about the number of alternatives in a system we design then, Shannon's measure of information will be problematic in such situations too. This point does not provide difficulty for the argument I am presenting rather it provides for a stronger version of it. See the following note, note 16, for a brief discussion.

rigorous fashion.[16]

Gregory does not take account of the fact that we require knowledge not only of the number of alternatives but also knowledge of the probability that each alternative may manifest. Information depends not only on what has been and what is, but also *on alternatives of what might be* insofar as information as a measure measures using logs the probabilities of what might be. When elements are strung together, the stringing together of the elements limits the freedom of which elements can be strung together hence, information is constrained also by what has gone before in that string.

As noted, Gregory's claim that information theory as a measure cannot be rigorously applied is a claim concerning our lack of knowledge to be able to apply the measure. Specifically he claimed that

> In order to apply information theory rigorously it is necessary to know the number of alternatives from which selections are made.[17]

The kind of selection that is spoken of here is not the kind of selection made when, for example: one selects among various types of Irish Cream Ale. Rather, the choice (or the

---

[16] C.B. Martin has pointed out that while Gregory does not himself give a deep criticism, his comments are the beginnings of a deep criticism. In particular, Gregory's comments could be developed into Martin's distinction between function and disposition. According to Martin, a function is single manifestation bound abstraction; whereas, a disposition has alternative manifestations with alternative dispositional partners. As Martin points out, his distinction holds for all physical systems not just biological systems. Martin's point is not problematic for my argument. His point serves to extends the scope of the argument and not provide a problem case. Martin's comments were made during the defense of this thesis.

[17] Gregory, R. L. (1986). 'Whatever Happened to Information Theory?' in Gregory, R. L. (1986). *Odd Perceptions*. New York: Methuen. p.190.

number of alternatives which may be manifested) is a function of the range of alternatives which the material in the system is capable of being causally interactive with. It is a relationship between, for example, the different wave length, on the one hand, and the capacity for the cones to be affected by that particular wave length, on the other. Or, in the case of neural transmission, the transmission of electrical signals, or neural transmitters, or neuropeptides or whatever else neurons transmit and receive locally or at some distance.

When information theory was applied to humans' capacities (in an attempt to measure our so called maximum rate of information), researchers Edmund Hicks (1952) and R. Hyman (1953) applied the measure to situations where there were well defined choices.[18] Hicks showed that in situations with clearly defined choices or (established number of alternatives -- such as pressing response keys in response to a light) the choice time was shown to be proportional to the number of binary choices. The maximum amount of information (a bit rate) was claimed to be established at around 22 bits per second for an expert pianist.

Gregory's objection (to this and other experiments like it) concerns the fact that we do not know the number of alternatives that humans attend to and select from. That is, we may be attending more than is assumed in the experiments: hence, the assumed number of

[18] Hicks, E (1952). 'On the rate of gain of information', *Quarterly Journal of Psychology*. **4**:11-26; Hyman, R. (1953). 'Stimulus information as a determinant of reaction time' *Journal of Experimental Psychology*. **45**: 188-196.

alternatives is not accurate. Or, as an even stronger version than Gregory claims, it may be claimed the human system may be receiving even more information then it can attend to: hence, the capacity of a nervous system exceeds even the alternatives which we attend to.[19] However, (picking up on the point made earlier) the measure of information (as we have seen) does not merely involve the number of alternatives it involves the probability that the alternatives will be manifested: the shortcut technique for counting bits of information is not always warranted.[20] It is with this point (that the probability is also an important component) my claim finds a toe hold.

What is known is that neurons, unlike a copper wire, are not constant in their ability to respond to an input. The probability that the various alternatives will or will not respond (since a function of the properties of the system) cannot itself be constant in a system where the system's properties are changing. Therefore, the ability to measure information, in Shannon's manner, is even further hampered than is recognized in

---

[19] This statement is stronger than Gregory allows however, Gregory's is in need of supplementation if we are to allow for blind sight, a case first reported by Larry Weiskrantz *et al* (1974). In the paper, Weiskrantz *et al* reported on a patient DB. DB was said to have lost the primary visual cortex and, as expected, DB gave verbal reports of being totally blind. Yet, DB could perform visually guided tasks such as (a) grabbing a moving object, (b) differentiating the orientation of vertical and horizontal lines; (c) differentiating between the letters 'X' and 'O'. One possible theory (accounting for this phenomena) is found in Rafal *et al* (1990) and Schneider (1969). This theory traces a visual pathway from the retina to superior colliculi (mid-brain) to the pulivnar nuclei (thalamus) and finally to the secondary visual cortex. Rafal (following Schneider) suggest that this path plays a role in the detection and localization of objects in space. For the purpose of my claim that Gregory requires a strong version than he allows all that is required is that the phenomena known as blind sight exist. Martin has drawn my attention to the fact that this case shows a split of experiential and cognitive aspects in brain activity. See: Rafal, R, *et al* (1990). 'Extrageniculate Vision In Hemianopic Humans: Saccade Inhibition By Signals In The Blind Field.' *Science 250*. pp. 118-121; Schneider, G.E. (1969). 'Brain Mechanisms For Localization And Discrimination Are Dissociated By Tectal And Cortical Lesions.' *Science 163* pp. 895-902; Weiskrantz, L, Warrington, E.K., Samders, M.D., Marshal, J. (1974). 'Visual Capacity In The Hemianopicfield Following Restricted Occipital Ablation.' *Brain 97*. pp. 709-728.

[20] See page 19 for the formula.

Gregory's discussion.

The discussion which follows provides a detailed account of various neuronal responses which are problems for Shannon's account of information. As noted in the introduction to the chapter, it is my contention that the kinds of choices available for manifestation in neurons are *unlike* the kind of alternatives which Shannon assumed in his notion of information and the kind which inform the measure of information. In essence, my claim is simple. There is an important difference between how electrons travel down a copper wire to a receiver (as it is with transmissions of Morse code) and the way neurons respond to input and generate responses. Copper wires and the receivers do not alter their response to inputs, whereas neurons can and do. I will argue that this difference is a problem for Shannon's measure of information (in this chapter) and I will argue that it is a problem for Shannon's account of information (in the next chapter). What will be argued is that neural changes mean that the size of the set of alternatives (that set required for Shannon's measure to be applicable) is (i) not a value constant through time and (ii) not, for any given time, not a specific number but (iii) for his measure to work it would have to be both.

Neurons and other matter in the brain are highly interactive; unlike the electrons 'traveling' down a copper wire or pulses of laser 'traveling' down a glass strand which lead to a receiver where the receiver has a unique response to the input, neurons are not constant in their responses to input and hence not constant in the probability that the cell will respond at all, or in the same way to the same input on successive occasions. For example, long-term potentiation (LTP) which is defined as:

> a long lasting (10 minutes to days, depending on the conditions) enhancement of synaptic transmission.[21]

or long term depression (LTD) which is defined as:

> a decrease in the efficacy of synapse.[22]

are but two examples of types of changes altering the efficacy of transmission and hence, alter the number of alternatives which could be manifested across a period of time.

Of particular interest, and with application to the aforementioned work of Edmund Hicks (1952) and R. Hyman (1953) (where a subject pressed a key in response to the prompting of a light, purportedly measuring the rate of information), is a chapter in *Principles of Neural Science*, aptly titled 'Phototransduction and Information Processing in the Retina'.[23] My purpose in discussing this is twofold: first to provide a concrete example in support of the claim that the number of alternatives Gregory writes of are a function of range of alternatives which the material in the system is capable of being causally

---

[21] Cooper, J.R., Bloom, F.E. Roth R. H. (1991). *The Biochemical Basis of Neuropharmacology, Sixth Edition*. Oxford: OUP. p. 437.
[22] Churchland P.S., Sejnowski, T.J. (1992). *The Computational Brain*. London: MIT Press. p. 467.
[23] Tessier-Lavigne, M. (1991). 'Phototransduction and Information Processing in the Retina' in Kandel, E., Schwartz, J. and Jessell, T. (eds.) (1991) *Principles of Neural Science. 3rd Edition*. New York: Elsevier. pp. 400-417.

interactive with, second, to suggest that the adaptive responses in neurons (or in this case the adaptive response of rods and cones) alter not only the number of alternatives (even if only temporally) but also the ability of the input to elicit a response in the rod or cone. While the following discussion serves the initial purpose of addressing Gregory's claims, the discussion at hand serves a greater purpose in this thesis.

In the visual system, phototransduction takes place in the outer segments of rods and cones. The segments of all rods are densely packed with a visual pigment called Rhodopsin, cones have three different kinds of visual pigments, each of which is sensitive to different wave lengths. It is the absorption of light by visual pigments in the rods and cones that leads to the changes in membrane potential. The significance of this is that, unlike most neurons which respond with an all or none response by firing an action potential, rods and cones respond to light with graded changes in the membrane potential. In short, the adaptive response in cones and rods has an effect on how the cone or rod responds to light thereby, effecting what is transmitted by the cell. This is especially important both in cells which respond with graded changes and when the cell has adapted to a level of light intensity and then is suddenly placed in the opposite environment. The momentary 'blindness' we all experience when emerging from total darkness into bright sunlight is an example of such.

Phototransduction of light occurs in three stages:

(1) light activates visual pigments;
(2) these activated molecules cause the stimulation of cGMP phoshodiesterrase,

an enzyme that reduces the cytoplasmic concentration of cGMP;

(3) the reduction in cGMP concentration closes the cGMP-gated channels, thus hyperolizing the photoreceptor.[24]

The process of adaptation which allows continued phototransduction at new levels of light intensity is important to the measure of information in that the changes cause momentary reduction in the quantity of information received. In addition, other important factors such as diet can effect responses and alter the rate of information capacity. For example, the very beginning of phototransduction requires the *all-trans* retinol or vitamin A. Since vitamin A is not synthesized by humans,

> [a] nutritional deficiency in vitamin A can lead to night blindness and, if left untreated, to the deterioration of receptor outer segments and to total blindness.[25]

The ability of the rod or cone to respond to light places severe limitations on any claim concerning the amount of information available. Clearly, such conditions are relevant for an experiment which purports to quantify human information rates, especially when these rates are determined by our response to a light source, as was the case with Hicks and Hyman.

While a vitamin A deficiency may be seen as the break down of a system, the process of

---

[24] Tessier-Lavigne, M. (1991). 'Phototransduction and Information Processing in the Retina' in Kandel, E., Schwartz, J. and Jessell, T. (eds.) (1991) *Principles of Neural Science. 3rd Edition*. New York: Elsevier. p. 404.

[25] Op. Cit. p. 404.

adaptation cannot be seen as -- merely -- such. The process of adaptation, the rate of adaptation and all the various components involved in the process which alter the rate of adaptation are relevant for measuring information.

Above it was noted that the concentration of cGMP was modulated by light: the reduction in cGMP concentrations closed certain channels, enabling the hyperolarization of the photoreceptor. The concentration of cGMP is also modulated by calcium ions ($Ca^{2+}$). In fact,

> the modulatory effect of $Ca^{2+}$ on cGMP is important for mediating light adaptation.[26]

In particular the relationship between $Ca^{2+}$ and cGMP is an inverse relationship: as the level of $Ca^{2+}$ increase the level of cGMP decreases. This is so as, $Ca^{2+}$ has an inhibitory effect on an enzyme which synthesize cGMP.

The level of $Ca^{2+}$ inside the segment (called the outer segment) of cone containing a visual pigment is controlled by two process: the flow of $Ca^{2+}$ into the cone across a channel gated by cGMP; and, the $Ca^{2+}$ is carried out of the outer segment by special $Ca^{2+}$ carriers.

[26] Op. Cit. p. 408.

In darkness, the level of $Ca^{2+}$ concentration is constant as the amount which enters the cell is thought to be equal to the amount carried out. During a period of prolonged light exposure, the cGMP channels remain closed while the carrier which extrudes $Ca^{2+}$ remains active hence, the level of $Ca^{2+}$ decreases; the inhibition of the enzyme which synthesizes cGMP ceases; the level of cGMP rises. With the rise in cGMP, the initial hyperpolarization which resulted from the activation of the visual pigment is reversed and the cone slowly depolarizes, adapting to the new level of light.

> A very bright light closes all the cGMP-gated channels, making the cones hyperpolarize from their resting potential (-40 mV) to -70 mV, the potential determined by the non-gated K+ channels. In this state the cones cannot respond to further increases in light intensity. However, if this background illumination is maintained, the cones slowly depolarize ...... and are once again capable of responding to further increases in light intensity - ***the bright light is no longer blinding***.[27]

Apart from being details, the above account provides insight into what is common with biological systems: their ability to adapt, the cumbersome nature (indeed the Heath Robinsonian nature)[28] of the process, and subtle variations which accompany complex processes causing subtle variations in the level of responsiveness of a cell to an input.

The role of vitamin A in the operation of Rhodopsin is one example where the lack of a substance or decreased quantity greatly alters the ability of a cell to respond to an input;

---

[27] Op. Cit. p. 408.

[28] John A. Baker has drawn my attention to the fact that a ludicrously impractical or elaborate machine came to be called "a Heath Robinson contraption."

the role of $Ca^{2+}$ is a second. Since the level of $Ca^{2+}$ has been shown to be important in adaptation, should something prevent the expulsion of $Ca^{2+}$ the level of cGMP would not increase and the rod would not adapt to the new level of light. Surely in either situation, the quantity of information would be greatly effected. What is clear is that biological, adaptive elements are very unlike the copper wires and fixed response systems which work so well with Shannon's measure of information.

So far at issue is the fact that adaptive responses serve to alter the probability that an input will or will not manifest. Unlike information transmission which involves electrons down a copper wire, or pulses of light transversing down a glass fiber, the very medium of neural transmission is one of change.[29] This requires more knowledge than Gregory acknowledged to rigorously employ Shannon's measure of information.

Adaptivity of neuronal activity means that the size of the set of alternatives is (i) not a value constant through time and (ii) not, for any given time, not a specific number. Hence there is no determinate ratio of something to the number of possible alternatives. Hence the probability of the occurrence of a specific sequence of elements is not determinate. Hence there is no possibility of identifying either the measure of information in a

[29] The contrast I intend here does not concern merely the change or lack of change in a physical system, surly all physical systems can change. Rather, when a change happens to the copper wire system (for example) we tend to say that the system is no longer working or that it has malfunctioned. This kind of a response is not appropriate to adaptation in the visual system; rather the visual system is able to exploit such adaptive capacity. The distinction, therefore, is not that some systems can change and others not, the distinction is that one system can exploit such alternatives whereas another cannot (While all physical systems can change not all physical systems can exploit the changes) This systemic 'ability' is not one that Shannon's measure of information can deal with.

sequence or the information in that sequence.

Thus stated, my argument is not much more than a stronger version of Gregory's argument. While it is stronger than Gregory's claim in that I claim the neurons, unlike copper wire, are not constant (or as constant) in terms of their probability to respond, my claim is nonetheless a claim concerning the lack of knowledge we have as to the probability and what that probability is at any given time. Like Gregory's claim, this claim does not affect the notion of information which Shannon uses nor his measure for, in principle, if we knew what the number of alternatives were and what their respective probabilities were at every instant, we could in fact measure information. The possibility that God could know and hence apply Shannon's measure remains. In order for my claim to amount to more than this, I must first show that the kind of activity in neurons is part of activity which is somehow fundamentally different then the activity assumed by Shannon to exist and second (the stronger claim) that such activity cannot be accounted for by Shannon given his basic notion of information, without Shannon giving up what is fundamental to information theory.

The presented objection is modest insofar as it limits the scope of application of Shannon's measure to specific kinds of systems. This is achieved by showing that Shannon's notion of information is limited to a specific class of systems and, more importantly, the conceptual resources of his system are not rich enough to enable generalization or extrapolation from his cases. Clearly, physical systems do exist which operate in accordance with Shannon's basic notion of information; this is obvious from the benefits and advances in technology resulting from Shannon's ideas. What is at issue is not the significance of Shannon's work rather, it is the scope of its application to all physical systems, with special attention given to networks of neurons.

## 4 Chapter Summary

The proceeding discussion could be summarized as follows:

1. The amount of information in a signal is a function of how surprising it is that the elements of that signal are as they are, that is, are in the order that they are.

2. The resources provided by Shannon *et al* for spelling out what "surprisingness" comes to require, as they stand, the possibility of fixing the cardinality of the possible alternative orderings for that signal.

But

3. Though this may be fine for something like a digital computer or telegraph system, it won't do for a set of neurons [see lemma L below]

4. Therefore, Shannon's theory is too limited in range of possible applications to be useful

LEMMA L:

1. There are two features [relevant here] of neuronal activity which are salient:
(a) plasticity and
(b) adaptivity.[30]

2. The plasticity and adaptivity of neuronal activity mean that the size of the set of alternatives is (i) not a value constant through time and (ii) not, for any given time, not a specific number.

---

[30] While both plasticity and adaptivity involves changes to in a neuron's responses to input, the distinction is left to provide for the possibility that different classes of changes may be identified. Nothing in my argument turns on the relationship between plasticity and adaptivity. The possibility that plasticity and adaptvity may (in the end) be the same or different is of no consequence to what is argued here.

3. Hence there is no determinate ratio of something to the number of possible alternatives.

4. Hence the probability of the occurrence of a specific sequence of elements is not determinate.

5. Hence there is no possibility of identifying either the measure of information in a sequence or the information in that sequence.

# Chapter III

## Signal/Noise Discrimination In Information Theory And Neuromodulation

### 1 Introduction

This chapter is a continuation of the proceeding chapter, providing an extended argument in favor of abandoning information based explanations of neural activity, at least where the information based explanation is, or is a derivative of, Claude E. Shannon's (1949) account.[1] In particular, this chapter more thoroughly reviews neuronal responses and the problems that such responses have for Shannon's account of information. This chapter presents a different argument to abandon Shannon's account of information, different from those reviewed in the last chapter. Where the previous chapter's argument was based on Shannon's measure of information this chapter's argument turns on Shannon's conception of information conveyance.

---

[1] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal. 27*. pp. 379-423. Reprinted in Shannon, C. and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press.

This chapter is a more thorough discussion of Shannon's theory of information. Emphasis is given to information conveyance (and not just his notion of information) and the relationship between Shannon's account of conveyance and the activity of various physical systems. While this chapter presents additional discussion of Shannon's account of information, no reference to probability or logs is made. Instead, Shannon's notion of information is looked at as separate from his exact measure of information. This is done to allow for the possibility that a different measure could be formulated while still adhering to Shannon's notion of information and his account of information conveyance.

My primary aim then in this chapter is to continue and uncover a list of prerequisites for any adequate theory of information which purports NOT to be restricted in its domain of application. That is, if the account of information is to be applicable to all kinds of physical systems (including neural systems) then the list of problems which I argue Shannon faces and which he cannot deal with become a list of desiderata for any adequate theory.

To this end I have three immediate aims: first, to arrive at a statement of what the core ideas of Shannon's account of information are and what his noise/signal distinction is; second, to state the conditions that must be present in physical systems if Shannon's account is an account of the conveying of information in the system (if his noise signal distinction is going to hold); and, finally, to inquire into the activity of known biological systems, searching for the presence of these noted constraints.

Indirectly, it is an investigation into the supposition that Shannon's measure of information cannot be applied to human capacity (not merely because we lack knowledge of all the required variables to employ the measure but) because the systemic conditions required for Shannon's account do not hold in biological neural systems. To show that Shannon's account of information and information conveyance is not compatible with neural systems is to show that the measure cannot be applied for at least this reason.

## 2 Shannon's Account of Information

In general Shannon's account of information deals with the technical problems of communication.[2] It is this purported ability (the ability which Shannon's account provides theorist for dealing with the technical problems of communication in all physical communication systems) that is the general interest of this chapter. In particular, concern lies with the physical conditions that Shannon's model requires for its instantiation in

---

[2] Weaver makes note of this in his fact too. See: Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p. 4.

physical systems. To this end, this chapter provides a rather fuller exposition of Shannon's account of information than was provided in the last chapter, which was as was merely a preliminary investigation of the issues presented here.

As an *a priori* condition, physical systems where communication occurs between entities through a form of direct or indirect inter-action (i.e. various kinds of change) and where not all inter-action between the entities nor all changes to the entity (or entities) within the system are part of the information transmission process there must (within the theory) be a means for discriminating between physical interactions (or changes to the entities) which are part of the process from those which are not. Shannon's account of information tries to achieve this by providing a distinction between noise and signal, where the structure of a signal is limited by Shannon's notion of information (the notion of information already has been argued to be a notion of weighted order).

While Shannon's account provides for a notion of signal which is defined in terms of and only of the function of a transmitter, his account of noise is defined in reference to both what a receiver does and the transmitter does. I will argue (in the following section) that Shannon necessarily has to have a notion of noise and the notion of noise he has is necessary given what he takes information transmission to be. The purpose of establishing the necessity of Shannon's notion of noise (and his noise/signal distinction) is to eliminate, for Shannon, any possible alternative account of noise and signal. I will argue that Shannon's account of noise and signal is not rich enough to account for the activity seen in biological systems; moreover since it is a necessary consequence of what he takes information conveyance to be Shannon's account (of information conveyance) necessary cannot account for information conveyance in biological systems.

## 2.1 The Notion Of Information

As shown in chapter two the notion of information provides the internal structure of messages. It is able to perform this role in Shannon's account as Shannon's notion of information essentially depends on the notion of order.

To see that this is so one need only consider that information theory, in general, concerns the structure of transmitted messages and the degree of order among the message's elements. In his article *Information theory* Robert G. Gallager, Fujitsu Professor of Electrical Engineering, MIT, Codirector of the Laboratory for Information and Decision Systems reported that:

> Information theory has nothing to do with inherent meaning in a message; it is rather a degree of order, or non-randomness, that can be measured and treated

mathematically.[3]

Specifically (as reported in chapter two) the order of each element in a system is governed by the element(s) which proceed it. This fact is clear from discussion of the measurement of information in chapter two. Information theorist Weaver (1961) makes this point when he writes:

> As the successive symbols are chosen, these choices are, at least from the point of view of the communication system, governed by probabilities; and in fact by probabilities which are not independent, but which, at any state of the process, depend upon the preceding choices.[4]

In short, Shannon's notion of information essentially depends on the notion of order. The order of a sequence is a function of the statistical dependency which successive elements have on proceeding elements in the sequence. It could be stated as follows:

> The Notion of Information (TNI): Information occurs in statistically ordered sequence[s] of elements.

---

[3] "Information Theory" from Britannica On-line, Copyright (c) 1995 by Encyclopedia Britannica, inc. Downloaded on May 10, 1997.

[4] Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. pp. 10-11.

## 2.2 The Transmittability Principle, The Definition Of Noise and The Reproducibility Requirement

According to Shannon, a signal is that which is transmitted in a system. The signal is established and transmitted across a channel (the medium over which a transmitted signal travels) to a receiver. Shannon's notion of a signal and his notion of information are interrelated. The relationship can be captured by a principle called The Transmittability Principle. It can be stated thus,

> The Transmittability Principle (TTP): Information sequences are capable of being transmitted as signals. When transmitted the sequence (within a signal) is *fully and completely fixed at the point of transmission.*

While Shannon does not explicitly state the TTP, this principle captures what is basic to Shannon's account of information.[5] It captures the idea that information can be transmitted as a signal and, when information is transmitted, the signal used to convey such is structured in accordance with the notion of information and this is achieved, by the transmitter, at the point of transmission.

The claim that the structure of the signal is fully formed at the point of transmission can

---

[5] David Chalmers also recognizes the implicit principle in Shannon's work. Following Chalmers I have retained his name of the transmittability principle. See: Chalmers, D. (1996). *The Conscious Mind*. New York: OUP. p. 283.

be found in the following quote from Weaver (1961) concerning the transmission of information.

> Information is, we must steadily remember, a measure of one's freedom of choice in selecting a message. The greater the freedom of choice, and hence the greater the information, the greater the uncertainty that the message actually selected is some particular one. Thus greater freedom of choice, greater uncertainty, greater information go hand in hand. If noise is introduced, then the received message contains certain distortions, certain errors, certain extraneous material that would certainly lead one to say that the received message exhibits, because of the effects of noise, an increased uncertainty......... Uncertainty which arises by virtue of freedom of choice on the part of the sender is desirable uncertainty. Uncertainty which arises because of the influence of noise is undesirable uncertainty.[6]

According to Weaver, information has one main characteristic. That characteristic is the statistical dependency which the members of a sequence have upon each other for the order of the sequence. The order of a message's elements (which are to be transmitted as a signal) is established at the point of transmission. Any distortions, omissions or alterations which occur after the point of transmission are said to arise as a result of noise. What is important is that Weaver describes noise (increased uncertainty) as anything which alters what is transmitted. Shannon also describes noise this way. He writes that noise exists when:

> the received signal is not necessarily the same as that sent by the transmitter.[7]

---

[6] Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. pp. 18-19.

[7] Shannon, C. (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal.*

TTP captures the requirements that information be both transmittable as a signal and when it is transmitted, the signal's structure be fully established at the point of transmission in accordance with Shannon's notion of information. TTP reflects this fact that is found both in Shannon and Weaver.

Consider the following as an expression of Shannon's definition of noise:

> The Definition of Noise (TDN): Any addition to the sequence which happens *after* the point of transmission, or to the receiver and its capacity to receive the sequence as first transmitted is noise.[8]

This definition serves to reinforce the role of the transmitter and what has been expressed by Shannon concerning the occasion of noise.

Since, for Shannon, the received signal may be altered as a result of a change to the signal or a change to the receiver's ability to receive the signal, both possibilities are included in

---

Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p. 65.

[8] These two conditions are implicit in what Shannon counts as noise. Shannon's account of noise is found on page 41 of this thesis; it states that noise exists when "the received signal is not necessarily the same as that sent by the transmitter." See note 7 of this chapter. John A. Baker has pointed out that Shannon *et al* need the principle of transmittability. It is because they accept that principle that they can adopt the very simple account of noise which is given in the definition, for that definition works only if signal is what is there at point/time of transmission and noise is whatever appears AFTER that point/time. While the objection to this is that the presented definition is not explicitly Shannon's, the discussion of noise in Shannon *et al* implies the definition presented. See for example the quote by Weaver appearing on page 41.

the definition of noise.

In short, TTP and TDN pick out what is basic to Shannon's account of information: the preservation of a message (a statistically structured sequence) through transmission. Evidence that Shannon's account is the preservation of messages across transmission comes from what Shannon describes as the fundamental technical problem he is dealing with. He writes:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.[9]

Shannon's account of information is an account of the required conditions to reproduce a message. So far as we have seen that list includes noise and signal. However, what is important is not that Shannon has an account of noise nor that he necessarily has to have an account of noise but that the account he gives (or one with at least the same elements to it) is a necessary consequence of what Shannon takes information and the transmission of information to be.

This claim requires that we bracket TDN and for the moment look at Shannon's process

---

[9] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal.27*. pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p. 31.

of information transmission and, in so doing, show that TDN is necessary given Shannon's account of information transmission.[10] To employ something other than TDN is to give up what is basic to Shannon's account.

Shannon's account of information transmission (as noted in chapter two) is fully and completely specified by five terms. Above these were noted as:

1. An *information source* [is that] which produces a message or sequence of messages to be communicated to the receiving terminal;
2. A *transmitter* [is that] which operates on the message in some way to produce a signal suitable for transmission over the channel.
3. The *channel* is merely the medium used to transmit the signal from transmitter to receiver.
4. The *receiver* ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.
5. The *destination* is the person (or thing) for whom the message is intended.[11]

The necessity for the definition of noise as stated above is a necessary consequence of Shannon's the strict relationship that exists between a transmitter and a receiver. *For*

---

[10] Required for my the argument is not that only TDN be the definition of noise rather any account of noise which has at least the same elements to it as TDN. That is, for an account of noise to be acceptable it must include the following as instances of noise: additions or alterations to the signal after the point of transmission and changes to the receiver which alter how the receiver responds to the signal. I will refer to such as TDN.

[11] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal 27*. pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. pp. 34-35.

*Shannon, a signal is defined with reference to and only to the activity in the transmitter; and, noise is that which prevents the signal from being received in the form it was transmitted at the receiver.*[12]

Since the receiver is defined as ordinarily performing the inverse operation of what the transmitter does, the definition is open to differing readings. For example, one may argue that the presence of the term 'ordinary' in Shannon's definition of a receivers allows the receiver to respond in different way to the same input. This possibility is very real when one looks at and only at the definition. For a *reductio* consider what it would be for there to be different responses in a receiver to the same input. Consider a telegraph with only two possible responses (call then long and short) to two kinds of signals (call them long and short also). The signals may be sent in various sequences. Assume that there is no fixed response to a signal. For example a long signal can be received as a long or a short signal and a short signal can be received as either a long or a short signal. If a series of signals is transmitted in the form of three long, three short and three long signals, the signal which is received may be any combination of three groupings of three. If the signal described is Morse Code, then the characteristic international signal for help may or may not be received as such. Rather than S-O-S, any of the following combination of

---

[12] During the defense C.B. Martin made note of the fact that Shannon's definition also fails to account for those occasions when a transmitted signal is distorted and then undistorted prior to being received. This point may serve to underscore the methodology and subsequent ontological commitments present in Shannon's account. I will not here argue for nor make a claim concerning Shannon's particular ontology.

three letters, symbols or short expressions may be received: d, g, k, o, r, s, u, w, 'semi colon,' 'understand.'[13] To count this as information transmission is absurd given that the definition becomes too broad to be of use.

Even if one rejects this as a reduction, embracing the very broad definition of what it is to transmit information, this is not the kind of information transmission Shannon was concerned with, nor is it the kind of precise information transmission that a system requires when that system processes inputs, transmits the result for further processing and repeats this until the 'solution' is obtained. John L. Tienson (1987) makes a similar point when he writes that the causal activity in a computer must be "exceptionless."[14] Hence, the relationship between the receiver and the transmitter is such that what is transmitted has (in the absence of noise) a single unique response in the receiver. This requirement (while not explicit in Shannon's writing) could be summarized thus,

> The Reproducibility Requirement (TRR): A receiver must have a unique respond to an input $C_1$ such that the receiver reproduces the signal as it was transmitted, save noise.

---

[13] *The World Book Encyclopedia* (1991). Chicago: World Book Inc. pp. 880-881.

[14] Tienson was writing about connectionist networks and how to construct a computer. When building a von Neumann machine he wrote that "the trick of making a computer is to get the causal process of the mechanism to mirror the syntactic processes specified in a programmes, for this to be done the rules must be precise and exceptionless." If in building a computer one needs to align causal process such that the activity mirrors syntactic processes, and the syntactic processes are precise and exceptionless so too (given this position) must be the causal processes: precise and exceptionless. Exceptionless requires that the same input produce the same response without fail. Tienson, J. L. (1987). 'An Introduction to Connectionism' *The Southern Journal of Philosophy* ***26:*** Supplement, p.1.

The term 'ordinary' in Shannon's definition of a receiver allows for the ontological possibility that noise exists; hence Shannon's definition of a receiver can be seen as an implicit acknowledgment of the *a priori* noise/signal distinction requirement.

A second argument for the claim that a receiver requires a unique response to an input is found within the transmittability principle, and the fact that order replaces the role that semantics usually plays in an account of information. Since Shannon's account of information eschews invocation of the notion of meaning and relies instead solely on the non-semantic notion of well orderedness of elements, it is essential that he have a notion of noise clearly defined. For, it is the order and preservation of the order across a transmission which replaces the notion of meaning in Shannon's account.[15] To give up the notion of order and to give up the preservation of order is for Shannon to give what is fundamental to is account and to his measurement of information. In so far as this is true, Shannon has no choice but to employ a notion of signal which is based on the operation of a transmitter and a notion of noise which is based on (a) additions or alterations to the signal which leaves the transmitter and (b) alterations in the receiver which prevent the receiver form receiving the transmitted signal as it was transmitted.

---

[15] It is possible that what is required is that order be preserved merely in a statistically significant number of cases. This possibility is allowed for and discussed in section 3 of this chapter. Nonetheless this is not a possibility Shannon himself considers. Given what he takes information conveyance to be, Shannon is committed to the Reproducibility Requirement. This point is argued in what follows.

What has been argued is consistent with what Shannon states as the problem of establishing a physical system in which information transmission occurs. He writes that the problem is that of

> reproducing at one point either exactly or approximately a message selected at another point.[16]

Shannon could have said other than he did and thereby not be committed to the noise/signal distinction he is committed to. For example, Shannon might have claimed that the signal is a function of both the transmitter and the receiver. Rather than being fully formed by the transmitter and receiver, Shannon might have offered an account of systemic activity which looks at signals and noise ***not*** in terms of what is transmitted, preserved across transmission and received in the form it was transmitted, but in terms of the activity in the system and what that activity enable or prevents in terms of further activity.[17] Shannon's account of noise/signal need not have been based on what was fully formed at the point of transmission. While such a position would not be committed to a noise/signal distinction which is defined by what the transmitter does (as this notion of a

---

[16] Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal.27.* pp. 379-423. Reprinted in Shannon, C.and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press. p. 31.

[17] C.B. Martin provides an account like this, though his account is more sophisticated than the account presented here. See: Martin, C.B. (forthcoming). On The Need for Properties: The Road to Pythagoreanism and Back. *Synthese.* forthcoming.

signal is not fully described by what the transmitter does: moreover it is very different notion of signal) this is not Shannon's account. Moreover, it is the NEED that possibilities like this must be considered that motivates this thesis.

The central difference between Shannon's noise/signal distinction and the account of signal presented above is that for Shannon a signal is defined by what a transmitter does and noise is defined with reference to both the receiver and the transmitter; whereas on the above account, the signal was defined with reference to both the transmitter and the receiver. The latter allows for the possibility that the same thing transmitted can be two different signals in the same systems as the actual signal is a function of both the transmitter and the receiver and not merely the transmitter. The latter is not Shannon's account.

## 2.3 Section Summary

On Shannon's account, signal and noise are defined by some reference to the function of the transmitter. That is, the transmitter performs the function of imposing the required structure (an order which is established by statistical relations among the members of the sequence) onto whatever is going to serve as the signal and then transmits the signal.

Noise occurs whenever the signal received is not the same as the signal transmitted. Both the signal and noise are defined by the function of the transmitter. ***For Shannon, signals are defined by what the transmitter does and noise is defined by the reproducibility requirement, with reference to what the transmitter does.***

What is basic to Shannon's account of information could be summarized thus,

> 1. The Notion of Information: Information occurs in statistically ordered sequence[s] of elements;
> 2. The Transmittability Principle: Information sequences are capable of being transmitted as signals. When transmitted the sequences (within the signals) are fully and completely fixed at the point of transmission.
> 3. The Reproducibility Requirement: A receiver must have a unique respond to an input $C_t$ such that the receiver reproduces the signal as it was transmitted, save noise.
> 4. The Definition of Noise: Any addition to the sequence which happens *after* the point of transmission, or to the receiver and its capacity to receive the sequence as first transmitted is noise.

Whereas the notion of information limits the structure of the signal, the transmittability principle, and the reproducibility requirement provide the necessary conditions for the definition of noise. While information, transmittability, reproducibility and noise provide a criterion for the identification of systemic changes which are information conveying (those changes brought about by the signal) from those which are not (those changes brought about by noise) the middle two are the central point for the transmission of information, the first point sets the limits on what counts as a signal and the fourth is a necessary consequence of the second and third. Hence the noise/signal distinction is necessarily necessary given what Shannon takes information transmission to be. To provide a different distinction is to provide a different account of information transmission.

## 3 Possible Conditions For The Realization Of Shannon's Account

Under what condition or conditions will a physical system be able to transmit and receive information, given what noise and signals are for Shannon? Do all physical systems which involve communication operate in accordance with such conditions? The first question is addressed in this section. An answer to the second forms the content of remaining part of this chapter.

The condition or conditions which must be satisfied for a physical system to be able to transmit and receive information, for Shannon, will of course be that condition or conditions which will enable the realization of the four essential elements of Shannon's theory of information, *viz*., the transmittability principle, the reproducibility requirement, the notion of information and the definition of noise. In other words, we need to know what a physical system would need to be like for those four elements to be realizable. I will suggest that in fact there is just one such condition, at least that there is just one such central condition, all other being variations on it. I will call this the rigidity condition. This condition can however be formulated in at least two main ways, the first I will call the strong rigidity condition or sometimes the rigidity condition proper, and the second I will call the weak rigidity condition.

The rigidity condition proper is a strong version of the rigidity condition. For that reason I will indeed call it the Strong Rigidity Condition (SRC); it is very restrictive in terms of which systemic activity can fulfill it.

> SRC: The Strong Rigidity Condition: for any given input $C_i$ there is ONE and only ONE fixed response / one and only one piece of information conveyed. $C_i$ is a well defined order of elements, the exact same order transmitted.

The second version of the rigidity condition, the Weak Rigidity Condition (WRC), is of course weaker. Whereas SRC assumes that $C_i$ can only be fulfilled by a very rigid notion of order, the second version WRC is tempered in its strength by allowing for a weaker condition on $C_i$.

> WRC: The Weak Rigidity Condition: for any given input $C_i$ there is ONE and only ONE fixed response / one and only one piece of information conveyed. The role of $C_i$ can be fulfilled by the entire sequence which is transmitted or an incomplete or altered version of the transmitted sequence. The degree of incompleteness, or the amount of alteration, which is may occur and not effect the 'correct - $C_i$ ' response in the receiver is dependent upon TNI.

What is common to each is the condition that the receiver have one and only one response to $C_t$. The requirement that there be one and only one response is necessary given TRR. That $C_t$ is a signal which is transmitted is defined by the action of a transmitter is a requirement established by TTP. Accordingly, noise is defined as that which prevents $C_t$ from being received. That the rigidity condition entail a necessary noise/signal distinction is a result of the fact that the rigidity condition is defined by TRR, TTP, TNI. An argument for what Shannon is committed to might be very briefly stated as follows:

> 1. TTP requires that information be transmittable as a signal $C_I$ and a signal $C_I$ be fully formed at the point of transmission.
>
> 2. TRR requires that a receiver have a unique response to an input $C_I$, such that the $C_I$ receiver is the same $C_I$ transmitted.
>
> 3. Hence, the first condition is that a receiver have one and only one response to an input $C_I$.
>
> 4. TNI states that information occurs in statistically structured sequences of elements.
>
> 5. How specific $C_I$ has to be is a function how well structured the code is that informs the statistical structure of information.
>
> 6. Hence the second condition is that $C_I$ depends upon the structure of the code.
>
> 7. Conditions one and two specify the rigidity condition.

The difference between SRC and WRC lies in what can fulfill the role of $C_i$, the input which is efficacious for one and only one response in the receiver. Whereas the common condition between SRC and WRC arises from the Shannon's account of information transmission, deciding among the two requires additional discussion as both SRC and WRC state the conditions for the realization of the Shannon's account of information.

## 3.1 Deciding Between SRC And WRC

While SRC has an *a priori* appeal to (what is received must be what is sent), there are many *a posterior* reasons to adopt WRC as being both more reflective of Shannon's account and more respectful of empirical data.

Three reasons to look at WRC are:

1. SRC imposes VERY strict constraints, constraints which would not be realizable in real world physical systems.
2. Moreover, SRC is probably not Shannon's because he clearly allows for relations weaker than the kind of *strong* rigidity. For example, see his discussions of redundancy and his emphasis on the statistical nature of information: the former allows, he thinks, for the latter.
3. Thirdly, SRC will not fit (does not appear to fit) certain kinds of neuronal interaction. For example, see graceful degradation and cases like Parkinsonian disease where patients seem able to function, despite there being reason to believe that they don't have rigid or even close to rigid sequences from inputs to outputs. Whereas SRC does not allow for such WRC does allows for systemic response which are empirically recorded.

Each reason will be looked at in turn.

### 3.1.1 Shannon's Intent

Shannon's intent was to engineer a system which would actually work in 'real world' conditions. Given his intent, and given that noise is a 'real world' problem, there is reason to expect that Shannon's actual account would be flexible enough allow for both variation in the input received and the transmission of the message, despite the variations. Since WRC allows for variation and SRC does not, there is *prima facie* reason to think that Shannon would opt for WRC. Hence WRC and not SRC is more likely Shannon's condition.

### 3.1.2 Redundancy

The case of redundancy poses a problem for SRC but not WRC. The redundancy of a message is a function of the statistical structure of the message. The more structure to the message, the less choice one has as to which element can be in a given place. Because of this, when an element is missing or altered, it is easily picked up and the correct element filled in. For example,

> [t]he redundancy of English is exhibited by the fact that a great many letters can be deleted from a sentence without making it impossible for a reader to fill in the

gaps and determine the original message. [18]

Clearly, the more redundancy in the language (the language in which the message is cast), the greater the number of element that can be omitted from the message without altering the message. Consider the following passage from William Blake's collection entitled Eternity[19]:

He wh bnds t hmslf a jy
Ds ts wngd lfe dstry
bt h wh ksss th joy as it fls
livs in etrnit's snrse

Even through vowels have been removed the passage is still largely readable. It reads,

He who binds to himself a joy
Does its winged life destroy
But he who kisses the joy as it flies
lives in eternity's sunrise.

Redundancy (which increases with the decrease in the un-expectedness of each subsequent element) affords the receiver of the message an ability to receive and respond properly to incomplete messages or messages with errors in them. Not only is this clearly an advantage in an environment where what is transmitted is more likely than not to be distorted, it is a problem case for SRC. Redundancy allows incomplete or altered

---

[18] "Information Theory" from Britannica On-line, Copyright (c) 1995 by Encyclopedia Britannica, inc. Downloaded on May 10, 1997

[19] Washington, P (eds.) (1994). *William Blake: Poetry*. Everyman's Library. p. 94.

messages to evoke the "correct" response in a receiver. Of the two version of the rigidity condition, WRC is most appropriate as WRC allows the role of $C_i$ to be fulfilled by (a) the entire sequence which is transmitted or (b) an incomplete or altered version of the transmitted sequence. The degree of incompleteness (or the amount of alteration) which may occur without altering the receiver's response from the 'correct' response is a function of the notion of information: in particular, the amount of redundancy. Given this, WRC may be recast in terms of redundancy as follows,

> WRC*:The Weak Rigidity Condition: for any given input $C_i$ there is ONE and only ONE fixed response / one and only one piece of information conveyed. The role of Ci can be fulfilled by the entire sequence which is transmitted or an incomplete or altered version of the transmitted sequence. The degree of incompleteness, or the amount of alteration, which is may redundancy in the code.

### 3.1.3 The Case Of Graceful Degradation

Graceful degradation also poses a problem case for SRC and provides support for WRC and WRC*. Graceful degradation is said to exist when:

> a system is capable of sustaining some hardware damage without being totally incapacitated [and second, when] ... a system is capable of 'behaving sensibly' on the basis of data that is partial or includes errors.[20]

---

[20] Clark, A. (1991). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*. London: MIT Press. pp. 89-90

Of relevance for the discussion at hand is that, given that a system can respond correctly with only partial inputs or incomplete inputs or inputs containing errors, SRC is too restrictive, as it excludes such possibilities. The case of graceful degradation illustrates the fact that if Shannon's notion of information is to be appropriate then WRC or WRC* as opposed to SRC will have to be the condition.

This kind of response is seen in Parkinson's Disease where a reduction of between 70% to 80% of dopamine-producing neurons may occur before the afflicted individual exhibits the characteristic symptoms of the disorder.[21]

This section was a review of cases which posed a problem for SRC. WRC was argued to be more appropriate given the considerations noted in the cases presented above. Despite this there are cases which pose a problem for SRC, WRC and WRC*. While some of the

---

[21] The reduction in DA producing neurons is between 70% to an 80% reduction. A *Scientific American* article reports a 70% reduction. Cooper *et al.* differ from the *Scientific American* article noting between a 70-80% reduction in DA and Cote, *et al.* report a straight 80% reduction and not a range. In addition to a reduction in DA Parkinson patients also record decreases in Norepinephrine and Serotonin level. However, the decrease in DA is by far the greatest reduction. The symptoms noted include: a) rhymical tremor; b) increased muscle tone or rigidity, sometimes called cog-wheel rigidity; c) difficulty in initiating movement, leading to a characteristic gate; d) slowness in execution. The cell bodies of DA generating neurons originate in the brain stem in a structure called the *substantia nigra* (black substance). The axons project to the frontal cortex where they are thought to be involved with the initiation of motor function. For a more through review of Parkinson disease see: Cooper, J.R., Bloom, F.E. and Roth, R.H. (1991). *The Biochemical Basis of Neuropharmology, Sixth Edition.* Oxford: OUP; Cote, L., Crutcher, M. (1991) 'The Basal Ganglia' in Kandel, E., Schwartz, J. and Jessell, T. (*eds.*). (1991). *Principles of Neural Science. 3rd ed.* New York: Elsevier; Youdim, M.B.H. and Riederer, P. (1997). 'Understanding Parkinson's Disease. *Scientific American.* pp. 52-59.

cases can be dealt with using Shannon's tools or by providing supplementation to Shannon's account, some of these cases (I will argue) cannot be dealt with either by applying the tools provided by Shannon's account or by supplementing Shannon's account. Rather, I will argue that accounting for such cases require that what is basic to Shannon's account be given up.

## 4 Potential problem cases for SRC, WRC And WRC*

The three cases contained in this section all involve the alteration of synaptic activity. These case are adaptation, plasticity and neuro-modulatory activity.[22] Each will be defined and discussed in subsection which follows. Each has the potential of being a problem case for SRC, WRC and WRC* in that each deals with changes in neural responses to inputs.

### 4.1 Revisiting The Case Adaptation: A Possible Objection To it Being a Problem For Shannon

In chapter II a case was presented - the case of adaptation. Adaptation was argued to be a

---

[22] Neuromodulation, adaptation and plasticity all deal with alterations in synaptic efficacy. The exact distinction between the three is not clearly defined. I employ all three terms to allow for the possibility that a clear and obvious or even a subtle distinction may be found. Nothing in my argument turns the existence or lack of existence of a distinction between all three.

problem case for the application of Shannon's measure of information. While this appears to be a problem case for the Shannon's notion of information (because it is incompatible with the principle of transmittability and the reproducibility requirement), there is a strong argument which rejects this position. In this section I look at this objection.

In essence the object is that while adaptation is, in fact, a problem for measurement it is not a problem for Shannon's notion of information as the empirical evidence is consistent with what SRC, WRC and WRC* would predict. The possible objection proceeds thus: while at first glance this case appears to be a problem case (as it appears to violate the condition of SRC, WRC or WRC*), the process of light adaptation does not itself contravene the transmittability principle nor the reproducibility requirement; rather, it seem to support them. Recall the case:

> 1) When one emerges from total darkness into the bright sunlight, one does not see.
> 2) For a brief period, the cone is not capable of responding and one sees very little if anything until;
> 3) The cone regains it ability to respond at the new level of light intensity.

Rather than being a problem case, the lack of response (or so the argument could go) is consistent with SRC, WRC and WRC*. When the cones (and rods) cannot respond, no information is transmitted and as predicted we do not see and see very little until the process is finished.

Moreover (the argument and the objection against viewing adaptation as a problem case could continue) WRC and WRC* allow the different variations of $C_1$ to cause the unique response in the receiver. One may claim that the receiver still has unique response: it now just has that response to a different input: an input at a different level of intensity.

While the first part of the objection is true (momentary blindness is compatible with no information transmission), the second part of the objection is false. Admittedly, it is true that the cone is responding to an input (possibly with a unique response) at the new level of intensity; yet the fact remains that because of the adaptation process the cone cannot *now* respond to the input (which it responded to at the previous level of intensity). Hence, the uniqueness requirement is violated.

## 4.2 Learning

> If you root yourself to the ground, you can afford to be stupid. But if you move, you must have mechanisms for moving, and mechanisms to ensure that the movement is not utterly arbitrary and independent of what is going on outside.[23]

---

[23] Churchland, P.S. (1986). *Neurophilospohy: Towards a Unified Science of the Mind/Brain*. London: MIT. p. 13

While Churchland (1989) suggests a general requirement for anything which is going to be capable of surviving in a changing environment, what is needed (among other things) is a form of memory that enables the entity to adopt the more appropriate responses, adapting to changes in its environment. Neuroscientists Ronald M. Harris-Warrick, Frederic Nagy, and Michael P. Nusbaum (1992) reinforce a similar point when they write:

> To survive, an animal must be able to continually adapt its behaviour to changes in the environment. Thus, the neural network that generates a behaviour must not only be able to generate a particular motor pattern faithfully it must be able to modify that pattern when the situation changes.[24]

Memory on a compact disk (CD) which is stored and capable of being accessed at the required time is not the kind of memory or learning which poses a problem for Shannon's model. Memory of the kind used in a CD player is consistent with the kind of memory required for Shannon's account. The kind of memory which poses a problem is memory that purports to occur in biological networks involving the alteration of cellular responses. Such alterations and the ability to alter or respond (an ability known as plasticity) poses a problem for Shannon's notion of information in that neither the conditions of SRC, WRC nor WRC* are met during period of change.

---

[24] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p. 87

Of concern is not any particular model of learning involving plasticity; rather it is that models of plasticity are models of alterations occurring among the relationships between neurons.[25] Given that biological systems undergo such alteration the requirement that an

---

[25] The origin of the idea that 'thought' is the interaction of neurons and that of neuronal change is a possible foundation for how thoughts form is generally credited to Donald O. Hebb (1949). Hebb wrote:

> When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cell firing B, is increased.
>
> (Hebb, D.O. (1949), p. 62).

While Hebb (1949) is generally credited with originating such ideas, other sources pre-date Hebb's first published statement of the idea. Included in these are William James (1890) and Warren McCulloch and Walter Pitts (1943), the founders of a school of thought which come to be known as connectionism or parallel distributed processing (PDP). For example, in his section concerning the Will, during a discussion on the tendency that a current in a cell has towards a particular path James wrote a passage which describes neural changes which approximate an early account of plasticity. James wrote:

> the result is a new-formed 'path' running from the cells which were 'rearward' to the cell which was 'forward' on that occasion; which path, if on future occasions the rearward cells are independently excited, will tend to carry off their activity in the same direction so as to excite the forward cell, and will deepen itself more and more every time if is used."
>
> (James, W. (1890), pp. 584-585)

McCulloch and Pitts provided additional support for the idea from a model showing that networks made of very simple units connected together can, in principle solve arithmetical and logical functions.

The suggestion that thought was contained in the circuits of neuron -known by Hebb as a reverberating circuits- and that learning is the alteration of such is an early statement of an idea that is now widely accepted among neural theorists as being a general intuition concerning how learning might occurs in the brain. The research into this possibility is divided by Cooper, Bloom and Roth (1991) into four large classes:

1. Cellular and molecular research which seeks to attribute the changes in cell-cell interactions to specific molecular events in their trans-synaptic operations;
2. Neuronal process research which seeks to define either the functional changes in a brain network or the set of neuronal pathways that are necessary and sufficient to account for the behavioural changes observed in an experimental learning paradigm;
3. Behavioural research in which intracerebral or parenteral drug treatments or other perturbations of brain structure and function are used to disrupt the ability of an animal to modify its behaviour in a predictable way in specific environmental settings;
4. Model systems research in which the object is to determine, from simulations of neurobiological

established relation exist between transmitters and receivers cannot be met.

When one admits into the system the capacity to learn *via* the mechanisms of plasticity, a given input will not elicit a unique response in the receiver; the response will change over time and therefore not be unique. Hence SRC, WRC and WRC* are all violated and Shannon's account (without some way of allowing for such) cannot account for this basic activity of neural systems.[26]

### 4.2.1 A Possible Reply To The Problem Of Learning And Plasticity

Worth considering is the possibility of supplementing Shannon's account in a way which might circumvent the objections just developed. One such supplementation might say

---

> events, the minimum number of hypotheses required to explain an equally abstractly defined memory or learning phenomenon; this model network of hypothetical 'neurons' is then used to predict either how brains work of how better computers ought to.
>
> (Cooper, J.R., *et al.* (1991), p. 429)

See: Cooper, J.R., Bloom, F.E. and Roth, R.H. (1991). *The Biochemical Basis of Neuropharmology, Sixth Edition*. Oxford: OUP; Hebb, Donald, O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons; James, William. (1890). *The Principles of Psychology. Volume Two*. New York: Dover; McCulloch, W. S. and Pitts, W. (1943). 'A logical calculus of the ideas imminent in neural nets' *Bulletin of Mathematical Biophysics 5*.

[26] During the defense, C.B. Martin made note of an additional case, presenting further problems for Shannon. In particular, Martin noted that interactive weather systems also exhibit adaptive responses to inputs. Martin's point is that the kinds of alterations in responses I discuss (adaptation, modulation etc.) are not limited to merely biological systems. Hence, the gravity of the problem need not be limited to and only to biological systems to any and all input/output systems which exhibit the requisite alterations.

that the periods of learning (which involve plastic changes) are periods of activity which, for some reason, are *not* required to meet neither SRC nor WRC nor WRC$^*$ In other words, the supplementation would limit the kinds of occasion when SRC, WRC and WRC$^*$ are not to be met by non-learning or un-learning periods.

> Supplementation S1: Conditions set out in either SRC, WRC or WRC$^*$ are conditions which apply to and only to physical systems which are 'trained up,' or during activity which is not learning but the basic operation of the system.

For example, it might be argued that to require that a unique relation exist when no such relation exists is an absurd requirement. The requirements of SRC, WRC and WRC$^*$ make sense (it could be claimed) when relations are established and not during the establishment (or the dis-establishment) of such relations.[27] Accordingly, the argument would be that the uniqueness requirement is only a requirement for systems which have established relations between transmitters and receivers. Hence, the required supplementation is a simple constraint, not on the uniqueness requirement, but on the application of the requirement: in other words, whichever rigidity condition is adopted does not apply to certain modes of neuronal activity is precluded from being applied to certain modes of neuronal interaction: namely periods of learning.

---

[27] During the defense Martin noted that dis-establishment should be added to the sentence.

Research on artificial networks has spawned a volume of literature which allows for the possibility of categorizing neuronal activity along the lines suggested, while not explicitly attempting to do so. In general, it is claimed that there is no unique input output relation until the network is trained-up; thus, for example:

> [o]nce the parameters are set, the network will give the output suitable to the input, and the answer to any given question is stored in the configuration of weights.[28]

On this account, learning is merely the establishment of conditions which enable a unique response to an input $C_i$.

With supplementation of the kind suggested, the case of plasticity or any case of learning involving changes in neuronal responses no longer pose a problem for Shannon's account of information. They no longer pose problems as they are excluded from consideration. They are not considered instances of information transmission in the strict sense but instances of learning where alterations are occurring such that information transmission can occur in accordance with SRC, WRC or $WRC^*$.

The problem with this attempted supplementation of Shannon's theory (the

---

[28] Churchland ,P. and Sejnowski, T. (1992). *The Computational Brain*. London: MIT. p. 137

supplementation in fact consisting merely of a limitation of the scope of application of his theory) is that this limitation is itself merely an *ad hoc* and unmotivated attempt to save the theory, for until the objection none would ever have thought of denying that in learning there is information conveying. Indeed, surely in much learning the learning is exactly by information conveyance if anything is.[29]

## 4.3 New Problems: The Case Of Dynamic Biological Systems

Research from the stomachs of crustacean provides problem cases for SRC, WRC and WRC*. Neither SRC, WRC nor WRC* can account for the kind of activity we see in dynamic biological networks. Moreover, I will argue that to account for such activity we need to give up what is basic to Shannon's account.

Like Aristotle, who studied the movements of the crustacean forgut in *Historia Animalium*, the cases presented in this section are of movements of the crustacean forgut.[30] Drawing on recent research, I argue that the kind of activity involved in the

---

[29] John A. Baker has suggest this response as a possible response. While a response is required such that learning and plasticity can be included on a list of problem cases, even if the supplementation were acceptable it would not prove to be sufficient to account for cases presented in the following section. Hence, even with an *ad hoc* attempt to salvage Shannon's account his account of information transmission cannot be salvaged.

[30] This point is noted by Bruce Johnson and Scott L. Hooper in Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p.1.

operation of biological networks is such that the basic operation of an established network requires that post-synaptic neurons have the capacity to elicit varied responses and not a unique response. This is contrary to all three conditions SRC, WRC and WRC* and contrary to Shannon's account of information. Since this is also in violation of Shannon's noise/signal distinction I will argue that Shannon's account necessarily cannot account for the activity in biological systems.

In particular, the claim put forward in this section echoes a claim by Peter A. Getting (1989) who wrote that:[31]

> No longer can neural networks be viewed as the interconnection of many like elements by simple excitatory or inhibitory synapses. Neurons not only sum synaptic inputs but are endowed with a diverse set of intrinsic properties that allow them to generate complex activity patterns. Likewise synapses are not just excitatory or inhibitory but possesses an equally diverse set of properties. The operation of a neural network must be considered as the parallel action of neurons or classes of neurons, each with potentially different input/output relations and intrinsic capabilities interconnected by synapse with a host of complex properties.[32]

---

[31] I am indebted to C.B. Martin for this case and my awareness of this body of literature; as well, I am indebted to Martin for the his insistence that alternatives which could, but may never, but could be manifested are very import for theoretical consideration when modeling physical interaction. Moreover, it is Martin's claim that to account for a particular response, all that is involved in that response are required to be considered in a theory. This includes pre-synaptic and post-synaptic aspects as well as all else which is present that could contribute to or prevent the response or provide alternative responses. It is in this way that alternatives are of central importance in Martin's theory. It is because of Martin's position, his teaching and repeating of the position that I present that position I do. See: Martin, C.B. (forthcoming). On The Need for Properties: The Road to Pythagoreanism and Back. *Synthese*. forthcoming.

[32] I am indebted to Martin for this particular quote. It will appear in a forthcoming article by Martin. See: Getting, P.A. (1989). 'Emerging principles governing the operation of neural *networks' Annual Review of Neuroscience 12*. pp. 187-188; Martin, C.B. (forthcoming). On The Need for Properties: The Road to Pythagoreanism and Back. *Synthese*. forthcoming.

In particular, I present evidence that a single neurotransmitter, such as dopamine (DA), does have many different effects or the same effect on a post-synaptic neuron.[33] These effects do not merely contribute to the firing or non-firing of a signal-all-or-none response; rather, the very different effects (that neural transmitters can have on neurons) result in the ability of a single network to manifest vastly different patterned responses. Getting writes,

> If the ability of a neural circuit to perform a function derives from the collective action of the constituent network, synaptic and cellular building blocks, then altering the properties of a building blocks can change the operation of that network. Thus, a single network could subserve several different functions. ..... By changing the properties of selected cells, or pathways, the operation of a network can be dramatically altered. A single network can be multifunctional, participating in or generating more than one behaviour. This in not to say that an auditory system could be made into a visual system, but, within the confines of an anatomical substraite, the functional organization of many neural networks appears to be under dynamic control, changing in accordance with the conditions at the moment.[34]

What we have come to know about dynamic biological systems is that activity which alters the relations between cells (and thereby preclude, the uniqueness requirement) is not limited to merely instance of learning; rather, dynamic biological systems incorporate

---

[33] Since neurotransmitters are but one of substances by which neuron interact, the neurotransmitter may be said to act as an 'input' on the post-synaptic neuron. It is in this sense that I use the term 'input'.
[34] Getting, P.A. (1989). 'Emerging principles governing the operation of neural *networks' Annual Review of Neuroscience 12*. pp. 193-194.

a rich variety of vastly different response to the same input by modulating the electrophysiological properties of neurons. Specifically, study of crustacean forguts (with small neural networks, capable of being isolated) has shown that,

> Modulatory substance can modify the activity state, number, and/or kinetic properties of these ion channels and thus change the intrinsic excitability of the cell (Kaczmarek and Levitan, 1987). Consequently, a single neuron can display a variety of different intrinsic activity patterns and can switch between them, depending on the modulatory environment.[35]

The above quote makes reference to the general point that known modulatory transmitters such as dopamine (DA), serotonin (5HT), octopine (Oct) --to name only a few -- ***alter the intrinsic membrane properties and therefore the response pattern that the neuron has for transmitters***.

> The stomatogastric ganglion (STG) is itself the target of a large number of neurons that synapse on and influence the component neurons of its neural networks. Some of these inputs may use classic transmitters with rapid actions that can change rhythmic motor patterns on a cycle-by-cycle basis. However, many of these inputs are modulatory neurons that use slowly acting transmitters, or neuromodulators, for at least part of their synaptic actions. *Their major purpose is to modify or alter the properties of the STG neurons and their synapses for periods of time varying from hundreds of milliseconds to minutes.* The result of this modulatory input is to change the activity pattern produced by the STG, thereby allowing anatomically fixed neural networks to produce a large variety of behavioral variants on a basic motor theme.[36]

---

[35] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks.* Cambridge: The MIT Press. *p. 117.*
[36] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks.* Cambridge: The MIT Press. p. 87

Moreover, the response of neuron to an input is not merely a function of an established electrophysiological property, but at times can be mediated by factors which are exogenous to the neuron. For example,

> Bursting induced by DA is very dependent on calcium currents, and is blocked by modest reductions in extracellular $Ca^{2+}$ ions.[37]

This case is of importance as it presents a clear problem case to the uniqueness requirement that has been argued Shannon's notion of information requires if it is to be an account of neural systemic activity. It is a problem case in that the role an input has is clearly not *merely* a function of what is transmitted, in an established network, rather it is a function of what is transmitted AND the environment in which the system presently exists. In the case considered, the environment includes amount of $Ca^{2+}$ present. The claim here is simple: the kinds of activity which need to be accounted for are not found within the resources of Shannon's notion of information. What is transmitted (the input) does not play the role of being sole determinant of the output in the system. A second case further illustrates this point.

> The same transmitter can even use different ionic mechanisms to induce bursting at different temperatures (Johnson et al, (1992)), presumably reflecting changes in the current that are active or available for modulation at different temperatures.[38]

---

[37] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. pp. 117-118.
[38] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological*

Clearly, what is transmitted is not the sole determinant of output. Moreover, how a neuron responds to input can be altered as a course of action which is not noise but part of the properties that a neuron possesses. These are properties which need to be accounted for in the theory in such a way that how a system exploits and constrains the properties (allowing for alternative responses to the same input) are made full and relevant use of in the theory. With Shannon's information theory this is not done; nor do we find the tools to do such in Shannon's account of information or information transmission.

Recall the claim presented above that,

> [o]nce the parameters are set, the network will give the output suitable to the input, and the answer to any given question is stored in the configuration of weights,[39]

while this claim is accurate for artificial networks it is not accurate for networks which alter their properties in such a way that the same input elicits a different response, generating a different pattern of activity. With biological networks which alter the response properties of its constituents even when the parameters are set, the network will *not always* give the output suitable to the input.

---

*Networks*. Cambridge: The MIT Press. pp.117-118.
[39] Churchland, P. and Sejnowski, T. (1992). *The Computational Brain*. London: MIT. p. 137

Shannon must not only distinguish between signal (information) from non-signal (noise): *Now* the requirement is that he distinguish (a) transmissions which are noise, (b) transmissions which are information and (c) transmissions which "enable" (b).[40]

Call this the richer signal/noise requirement.

If the richer signal/noise requirement can be fully established as a requirement then, given

---

[40] While John A Baker has assisted me with the wording of this point, the point itself has its origin in a distinction made by C.B. Martin. The distinction I have drawn between kinds of signals is not Martin's distinction, although it does have similarties to Martin's distinction. Martin once differentiated between systemic activity by distinguishing between, what he called kinds of use in a physical system. The first kind, Martin labeled 'Use' and the second he labeled 'Representational Use.' Martin writes:

> 'Use' occurs in a system's reactivities to input and reactivities for the production and continuation or alteration of output.
>
> 'Representation Use' requires a system of sufficient complexity and adaptive variability of response to be capable of integrative, adjustive, combinatorial, projective negative and positive feedback and feedforward reactivities to the system's input and output for which there is a disposition base involving patterns of interrelated dispositions for:
>
> (a) alternative kinds of potential processing reactivities to the reception of alternative kinds of potential input that provides a capacity background of capacities (largely unmanifested on the occasion but still essential) of the system that provides a degree of specificity and richness for the actual results (the system's parallel to interpretation) of the processing reactivities that determine the specific directednesses and selectivities (content) of the actual input; and for
>
> (b) alternative kind of potential directive and selective reactivities for potential formation and continuance or alteration of potential kinds of output that provide a capacity background within the system (largely unmanifested on the occasion but still essential) that provide a degree of specificity and richness for the actual directive and selective reactivities (content) for the formation and continuance or alteration of the actual output, that is a projective activity parallel to stating. (Martin,C.B., (unpublished (a))

See: Martin, C.B. (forthcoming). On The Need for Properties: The Road to Pythagoreanism and Back. *Synthese*. forthcoming.; Martin, C.B. (unpublished (a)). "Toward A Model For Mind And Brain: Some Prepatory Notes." unpublished; Martin, C. B. (unpublished (b)). "What is Imagistic About Verbal Imagery and Why Does it Matter." unpublished.

this it is in not merely Shannon's noise/signal distinction AND given that Shannon's noise/signal distinction is a necessary distinction, Shannon's account is inadequate as an account of systemic activity for it requires a richer signal/noise account and subsequently a richer account of information conveyance than the account allows.

The neurons possessing a property known as a plateau potential provide additional support for the richer signal/noise requirement. A neuron which is capable of a plateau potential displays bi-stable membrane potentials. That is,

> a neuron capable of plateau potentials has two quasistable [sic] membrane potentials (a more hypolarized "rest" potential and a depolarized "plateau" potential). The neuron can make transition between the two states in response to brief synaptic input, postinhibitory [sic] rebound, or current injection. The transitions themselves are regenerative, i.e. a depolarization above a certain threshold voltage from the rest state will activate voltage dependent depolarizing conductance that then drive the neuron to the fully depolarized plateau, and relatively small hyperpolarizations from the plateau will induce an active reploarization to the rest state (Russell and Hartline, 1978, 1982).[41]

Harris-Warrick, Nagy and Nusbaum's write:

> When such neurons [neurons with plateau potentials] are part of a network, they can be switched between two activity states in a manner that is relatively insensitive to the intensity or duration of synaptic inputs, once a particular threshold is excited.[42]

---

[41] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p.22.
[42] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p. 119.

When a neuron is relatively insensitive to an input (and the activity of a neuron is being generated by its bi-stable membrane property), the network will *not always* respond with the output (which it is trained up to respond to) when the input is present in the previously required way. Hence, the uniqueness requirement is not met during a period of network operation.

A dynamic biological network exploits the benefits which accompany having the capacity to response to the same input in many different ways (as opposed to merely having a unique response). This exploitation is part of the basic operation of a biological network; the mechanisms for variability in response are not merely the mechanisms of learning.

While the evidence presented comes from invertebrates, Llinas (1988) showed that vertebrate neurons have similar amount of variability in terms of intrinsic responses, thereby[43]

> dispelling the myth that complex properties are important only in numerically smaller invertebrate networks.[44]

---

[43] Llinas, R. (1988). "The Intrinsic Electorphysiological Properties Of Mammalian Neurons: Insights Into Central Nervous System Function." *Science* **242**. pp. 1654-1664.
[44] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p. 135.

## 4.4 Discussion On The Relevance Of Neuromodulation

Networks of neurons within the crustacean's forgut elicit many different patterns. The network's ability to achieve this is a result of the fact that individual neurons respond differently to the same input. This fact prevents such neural networks from meeting any of the condition SRC, WRC and WRC* outlined as the required conditions to instantiate Shannon's account of information.

The above discussion of neuromodulation provides evidence for what was claimed above: Shannon's account requires supplementation to distinguish between (a) transmissions which are noise, (b) transmissions which are information and (c) transmissions which "enable" (b).

However, if Shannon's account were to be broadened to provide a distinction between these three, TTP, TRR and hence Shannon's current distinction between noise and signal would have to be abandoned.

The TTP and TRR (which are the basis of Shannon's account) provide for a distinction between signal and noise which is based on the function of the transmitter and the preservation of what is transmitted. Shannon's account IN FACT calls for the preservation of what is transmitted in the form it was transmitted. While this type of an account works well for systems which do not require varied response to the same input, his account fails to explain the activity of systems which DO make (as a course of action) have many different responses to the same input. This argument could be summarized thus. The plasticity, adaptivity and neuromodulatory activity of neurons mean that the same signal can have many different response in a receiver. The Shannon *et al* noise/signal distinction does not allow for this. Hence Shannon *et al* cannot provide an account of learning, neuromodulation or adaptive responses. The Shannon *et al* distinction is a necessarily necessary distinction given Shannon's account of information transmission. Hence Shannon *et al* necessarily cannot provide an account of what is basic to neural systems.

To account for neuromodulatory activity is to give up the TTP, TRR and TDN. More importantly, it is to include in the explanation and description of activity BOTH the pre-synaptic and post-synaptic elements of the system and the current and possible states of such.[45] After all, the effect that a neuromodulatory transmitter may have on a system's pattern of activity is dependent upon the previous activity of the system's constituent neurons. The response to a modulatory input varies with the state of the system and

[45] I am indebted to discussions with C. B. Martin for drawing my attention to the significance of both the pre and post-synaptic elements in the explanation of neurons and more generally, the importance of all factors during 'causal' activity or instance of change. See: Armstrong, D.M., Martin, C.B., Place, U.T. (1996). *Dispositions: A Debate*. New York: Routledge.

hence, its neurons. Therefore, noise and signal cannot be accounted for by merely restricting attention to and only the function of a transmitter. To do this is to give up what is basic to Shannon's account.

The suggested supplementation which dealt with the problem of learning (by limiting when SRC or WRC need be met) does not account for the evidence presented. The alternative responses are not limited to periods of learning but are part of the basic operation of neural systems.

Hence, the tools found within Shannon's account of information are not capable of doing the job as the Shannon's basic notion of information is inconsistent with empirical data concerning the operation of dynamic biological networks.

If it is needed, additional evidence concerning the richness of the operation of a biological network comes from modulators which act as circulating neurohormones. The relevance of this comes in terms of the fact that even less direct control is exercisable in such situations. That is, the circulating transmitter while it does have results down the system it need not always. This further complicates the activity in the systems as well as accounting for it. The case below merely illustrates evidence of the existence of circulating neurohormones.

When a modulatory neurotransmitter acts as a circulating neurohormones, the transmitter's effect is not localized to receptors at the site of its release; rather, the transmitters travel to non-local receptors sites. At such sites, the acting neurohormone acts as a locally released transmitter. For example, Harris-Warrick et al. cite a case where:

> 1. Serotonin is absent from terminals in the STG of spiny lobsters;
> 2. Serotonin is known to be released at sites near the STG in lobster;
> 3. Serotonin has modulatory effects on a network (the pyloric network) whose activity is influenced by the STG:
> 4. The kind of effect serotonin has is lower by several magnitudes of order than networks where serotonin is in local terminals.[46]

They write,

> This is consistent with serotonin acting as a circulating neurohormone to modulate the STG, where it acts as a locally released transmitter in other species.[47]

Notions are required to account for this kind of activity, activity in which individual properties are fully exploited as well as controlled and these notion are not found in Shannon's account nor can they be added to Shannon's account, for reasons noted above.

---

[46] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p. 94.

[47] Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.). (1992). *Dynamic Biological Networks*. Cambridge: The MIT Press. p. 94.

This is the challenge to be met and the challenge I take C.B. Martin (1997) to be partly embracing and I believe largely succeeding, though I will not argue this point here.[48]

## 5 Chapter Summary

The proceeding discussion could be summarized as follows:

1. Shannon's account of information conveyance is a function of three things: the transmittibility principle (TTP), the reproducibility requirement (TRR), the notion of information (TNI). A fourth thing, the definition of noise (TDN) is a necessary consequence of TTP and TRR.

2. The resources provided by Shannon *et al* for spelling out the condition required for the conveyance of information (and hence the realization of these four items) in physical systems is limited to a specific condition. [see lemma L2a below].

But

3. Though the condition may be fine for something like a silicon system, it won't do for neuronal interaction [see lemma L2b below].

4. Therefore, Shannon's theory is too limited in range of possible applications to be useful.

LEMMA L2a:

1. TTP requires that information be transmittable as a signal $C_t$ and a signal $C_t$ be fully formed at the point of transmission.

2. TRR requires that a receiver have a unique response to an input $C_t$, such that the $C_t$ received is the same $C_t$ transmitted.

---

[48] Martin, C.B. (forthcoming). On The Need for Properties: The Road to Pythagoreanism and Back. Synthese. forthcoming. Heil, John and Martin, C.B. (forthcoming). 'Rules and Powers.' Philosophical Perspectives.(forthcoming); Martin, C.B. (1993). The Need for Ontology: Some Choices. *Philosophy 65:* 505-522.

3. Hence, the first condition is that a receiver has one and only one response to an input $C_1$.

4. TNI states that information occurs in statistically structured sequences of elements.

5. How specific $C_1$ has to be is a function how well structured the code is that informs the statistical structure of information.

6. Hence the second condition is that $C_1$ depends upon the structure of the code.

7. Conditions one and two specify the rigidity condition.

LEMMA L2b:

1. There are three features [relevant here] of neuronal activity which are salient:
(a) plasticity;
(b) adaptivity and
(c) neuromodulation

2. The plasticity, adaptivity and neuromodulatory activity of neurons mean that the same signal can have many different responses in a receiver.

3. The Shannon *et al* noise/signal distinction does not allow for this.

4. Hence Shannon *et al* cannot provide an account of learning, neuromodulation or adaptive responses.

5. The Shannon *et al* distinction is a necessarily necessary distinction given Shannon's account of information transmission.

6. Hence Shannon *et al* necessarily cannot provide an account of what is basic to neural systems.[49]

---

[49] The structure of the argument here has resulted from discussion with Dr. Baker. Its present form has been heavily influenced by these discussions.

# Chapter IV

## Noise/Signal Distinction Within Vector Transformation Models

### 1 Introduction

In this chapter an investigation concerning the distinction between noise and signal in vector transformation models is presented. In particular, the relationship between the noise/signal distinction in vector transformation models of a network's activity and the noise/signal distinction in Shannon's account of information is investigated.

One possibility of salvaging Shannon's account (or mitigating the problem that were argued to exist for it) is to appeal to vector transformation models. I will argue that such models face exactly the same problems as the cases examined by Shannon, IF they work with his notion of information, and certainly most of them do.

The aim of this chapter is to inquire into vector transformation models as a possible retreat for Shannon *et al.* I will show that if Shannon *et al* retain the transmittibility principle (TTP), the reproducibility requirement (TRR) and the notion of information

(TNI) [the basic elements] then he is not helped by moving to vector transformations. This chapter is largely an application of the arguments presented in chapter three.

Since vector transformation models are models of artificial networks I will draw a distinction between artificial networks and biological networks. This will allow for the possibility that vector transformation models can be descriptions of artificial networks without also being descriptions of biological networks.

## 2 Vectors And Vector Transformations: A Linear Case

This section is purely expository. The basic mathematical notions of a vector and the mathematical tools which underlie the notion of vector transformation as well as the theoretical relationship these notions have to artificial networks is outlined.[1]

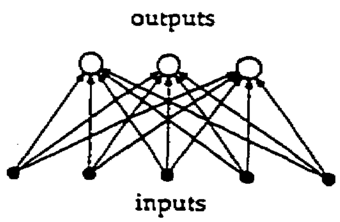
A vector is a mathematical notion. It is an ordered set of values such as the set <3,2,1>. What is labeled as a 'vector transformation' by computational theorists (such as Churchland and Sejnowski) has at its foundation linear matrix algebra. In particular, vector transformation is the multiplication of matrices. For example, when matrix $t$ and $x$ are multiplied together a new matrix results $y$.

$$t = \begin{matrix} 1 & -3 & 8 \\ 2 & 0 & 1 \\ 0 & 1 & 5 \end{matrix} \qquad x = [3,2,1] \qquad y = [?,?,?]$$

[1] J.A. Baker has suggested this way of introducing the topic of this chapter.

The input vector $x = (3,2,1)$ will yield an output vector $y = (5,7,7)$. The solution $y$ is arrived at by employing matrix multiplication. The first component of the vector $x$ is multiplied by the first item in the top row of matrix w (3 X 1) , next the second component by the second item in the top row (2 X -3), and finally the third component by the third item in the top row (1X8). The resulting three products are added together (3+-6+8) resulting in 5. When this is repeated for each row of the matrix, the full result $y$ is obtained (5,7,7).[2]

According to theory, the above account is the mathematical description of activity occurring in artificial networks labeled *feedforward, linear association networks (or linear associators, for short).* The diagram below depicts a feedforward network consisting of one layer of weights connecting input units to output units. It is called a feedforward network due to the single forward direction input travels. And, it is called an associator due to the manner (vector transformation) different inputs (such as input vector $x$) are associated (with a weight matrix such as $t$) to produce an output (such as output vector $y$).




(Diagram taken from Churchland and Sejnowski (1992).p. 76)

[2] This can be expressed by the equation: $y_i = \Sigma\ w_{ij}\ x_i$ Interested parties can refer to a basic text on linear algebra for a more complete account. For example, see: Nicholson, W.K. (1986). *Elementary Linear Algebra With Applications*. Toronto: Wadsworth Publishers of Canada.

Within such a network the matrix $t$ (the weight matrix) represent the weights connecting the inputs and outputs. The inputs are vectors. In the above example, the vector $x$ (3,2,1) represented the input vector. The output units are where the output vector $y$ results. In short, the product (called an inner product) is the sum of the weighted inputs to the outputs. In the above description,

> the output is a linear transformation of the input, and the network is called a linear associator.[3]

In matrix algebra the relationship between inputs and outputs is well defined. The relationship between variables servings as inputs and the outputs which results when the input is applied (in a manner specified by a formula to other variables) is a well established and well defined relationship: it is defined by the formulas and the rules for performing the operations. This is true for both linear and non-linear transformation. [4]

---

[3] Churchland, P. S., Sejnowski, T. (1992). *The Computational Brain*. London: MIT Press. p. 78.

[4] A note concerning non-linearity in vector transformation accounts, vector transformations in non-linear systems occurs when each processing unit performs a nonlinear operation on the linearly weighted sum of its inputs. The primary difference between the two is the nature of the operation (linear operation vs. non-linear operation). This is not a difference which makes a difference in the argument presented.

The demand for a non-linear model arises from a simple *a priori* requirement. Namely, if physical interaction is to be fully modeled then non-linearity must be included in the model. Churchland and Sejnowski noted this requirement when they write "the regularities of computational interaction and regularities of physical interaction should be even formally similar." (p. 85) This requirement was met, for a certain class of systems, in 1982 by John Hopfield. Hopfield (a physicist turned molecular biologist turned brain theorist) employed the notions of thermal dynamics to prove that the dynamics of the system would converge to a single solution and not wander endlessly or oscillate. His supposition which prove to be true was that "the interactions between representations in networks that end up delivering an answer to a question [are] described by the same laws that describe the behavior of certain systems in physics." (p. 85)

When an error occurs, in matrix multiplication, a correct solution is not achieved for a given set of values when applied to a given model. One suggestion (concerning what error might amount to when vector transformations are said to be instantiated in physical systems) worth considering is that in a trained up network (a network which has an established input-output response) noise represents one source of error and malfunction represents another.

---

Hopfield genius resides in the fact that he saw that:

> the local energy minimum in spin glasses could be a physical embodiment of prototypes in associative memory.(p. 86)
>
> Spin glass is characterized by particles with spin, either up or down, in mixture of attractive and repulsive interactions. (p. 86)

Clearly if, Hopfield was successful in his task, computational theorists would no longer be limited to rules based in logic or linguistics rather, connectionist theorists would be given the very powerful theoretical tools of physics.

There is a property of spin glass called 'frustration'. Frustration occurs when spin glass is cooled, the particles are not all aligned up or down; rather the particles reflect the 'inability' of the system to attain a unity of 'direction'. Hopfield postulated that this particular property of spin glass closely approximated the dynamics of a network of inter-connected units. He argued that in a like manner to the cooling of spin glass which results in a non-uniform configuration, the same stable minimum energy level would manifest in a theoretical network. Hopfield was able to show that the minimum energy state is achieved by assigning each connection either a 1 or a 0 indicative of an up or a down spin. Like the effect of cooling spin glass - which brings about the property of frustration in which a minimal energy state is attained- Hopfield applied this technique to networks. The units of the network, likewise, reached a minimal energy state. The units were flipped one at a time. If the unit initially had a value of 1 it was 'flipped' to a value of 0 and *vice versa*. At each flip the total energy of the network was checked. If the energy decreases as a result of the flip the change is acceptable. The flipping of units was continued until any given flip increased the overall energy in the network. Since there is a lower energy boundary, the flipping will not continue forever and the network will settle into a minimal energy state. It is because of the spin glass analogy that a stable state is (at times) referred to as an 'energy function'. All quotes are taken from Churchland, P. S., Sejnowski, T. (1992). For a clear account of Hopfield see Crick, Francis (1994). *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Charles Scribner's Sons; Churchland, P. S., Sejnowski, T. (1992). *The Computational Brain*. London: MIT Press.; Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings National Academy of Science*. *79* pp. 2554-2558.

## 3 Noise/Signal Distinction In Vector Transformation Accounts: The Linear Associative Network.

Is there a distinction between noise/signal in an associative network? In artificial networks, information conveyance of a sort occurs between the elements of the network. If the information conveyed is information in the sense developed by Shannon, and given the purpose of this chapter it must be Shannon's account of information if the discussion of this chapter is to serve the function outlined for it, then these networks are subject to the same *a prior* condition that Shannon's information conveying system is subject to. Thus, the constraint was stated above as:

> As an a priori condition, physical systems (where communication occurs between entities through a form of direct or indirect inter-action i.e. various kinds of change) and where not all inter-action between the entities nor all changes to the entity (or entities) within the system are part of the communication process there must (within the theory) be a means for discriminating between physical interactions (or changes to the entities) which are part of the communication process from those which are not.

Shannon's noise/signal distinction was fulfilled in a necessary way due to the principle of transmittibility and the reproducibility requirement. Does the same hold true for vector transformation models? To investigate this consider what noise and what signal are in a vector account, in particular in a linear associative network.

The linear associative network is one example of a vector account. The difference between the linear associative network and other networks lies in the complexity of the mathematical task that is claimed to be performed. Since my discussion hinges not on the linearity nor non-linearity of mathematical operation but upon the basic requirements to perform matrix multiplication what is argued holds for linear associative net will hold in general for a vector account.

The argument presented here is the following. The signal/noise distinction in a vector transformation account must be the same signal/noise distinction which Shannon employed if we are to say that the notion of information used in vector transformation systems is the same as that discussed by Shannon. Hence implicit within a vector transformation account is Shannon's TTP, TRR and hence so too is TDN. These were stated as,

The Transmittability Principle (TTP):

> Information sequences are capable of being transmitted as signals. When transmitted the sequences (within the signals) are fully and completely fixed at the point of transmission.

The Reproducibility Requirement (TRR)

> A receiver must have a unique respond to an input $C_1$ such that the receiver reproduces the signal as it was transmitted, save noise.

The Definition of Noise (TDN):

> Any addition to the sequence which happens after the point of transmission, or to the receiver and its capacity to receive the sequence as first transmitted is noise.

To show that TDN is implied in the vector transformation account is to show that in a vector transformation account both TTP and TRR are implicit. To show the necessity of TTP and TRR is to show that without TTP or TRR the notion of vector transformation cannot be viewed as working with the same notion of information as that developed by Shannon.

The argument which follows draws heavily form what was argued in chapter three. In particular, I will argue that in a vector transformation models the transmittibility principle (TTP) and the reproducibility requirement (TRR) are present and necessarily present. Hence, given chapter three, so too is the definition of noise (TDN). Given TTP, a vector model has the same notion of a signal, though not necessarily Shannon's notion of information. But, given chapter three, Shannon's noise/ signal distinction may be fine for something like a silicon system, Shannon's distinction won't do for neuronal interaction. Therefore, even though vector transformation model may fit artificial networks they do not fit biological networks. TTP is all that is required to have Shannon's account of signal given that TTP states what it is for there to be a signal.[5]

[5] The argument which cannot be borrowed from chapter three is the argument for the rigidity condition. While the vector transformation account has the first two elements for the SRC or WRC, the final requirement was Shannon's notion of information (TNI). This has not been shown to exist in a vector transformation account. Therefore, an argument that showed that WRC or SRC also holds for a vector account would be a different argument than stated in chapter three. However, since Shannon's noise/signal distinction is a necessary consequence of TTP and TRR it follows that vector models because they necessarily (as I will argue) have TTP and TRR also necessarily have Shannon's noise/signal distinction. This claim is sufficient to show the problems with vector transformation models. It is important to note that the relevant aspect of the rigidity requirement (for the discussion at hand) is the condition which results from TTP and TRR hence, it would not take much to establish that a vector account employ a version of the rigidity condition. The version would retain the unique response in a receiver to $C_1$ however, what would count as $C_1$ would be an input vector.

## 3.1 Vector Transformation Accounts, And Linear Associated Networks Have The Elements For TTP And TRR.

There are two elements to TTP. These are:

(a) information sequence;
(b) the transmutability of (a);

### 3.1.1 The Information Sequence

The notion of vector coding is a notion that identifies that which performs the role of the information sequence. In such a model it is referred to as a symbol or representation in the model.[6] While the notion of vector coding is contrast with local coding (coding in single elements) [the distinction is between coding in a serial digital computer (local coding) and coding in a network where activity is distributed across may co-active units] the contrast concerns the nature of representations in systems and not system and systemic activity involving representation and systems and systemic activity without representation. This distinction is summarized thus,

[6] While there is more than a terminological difference between Shannon's account of an information sequence and the notion of a symbol or representation, the relationship between the two and Shannon's notion, is such that if you have a symbol or representation you have an information sequence. Whereas an information sequence (in Shannon's sense) assumes a lack of semantic content, the same is not said of a symbol nor (at times) a representation. Minimally, what is important is that the notion of a symbol or representation fulfills the required conditions for my argument. Additionally a point raised by J.A Baker is worth discussing. According to Baker, given that Shannon avoids semantics it is not clear that Shannon could accept the notion of a symbol or representation. While this is true on the surface, it is also true that Shannon's account is not precluded from, at some point, incorporating a notion of semantics. Should this be done, it may very well turn out that Shannon (or someone on behalf of Shannon) adopt some notion of symbols or representations. This is a line of inquiry pursued by Donald MacKay (1969) where he discusses the relationship between semantics and information theory. See: MacKay, D. (1969). ***Information, Mechanism and Meaning***. London: MIT Press.

> vector coding uses many cells to code a representation; local coding uses one.[7]

Hence, with vector coding

> distinct representations will be coded as the overall set of activity levels across the neurons in the relevant population.[8]

The first element of TTP is present in the vector transformation account. The presence of this element may be shown a second way. Since the basic notion of computation requires a notion of representation or symbol and since neural networks are taken to be a 'style' of computation, the existence of a notion of symbol or representation is implied.[9] Either way it is looked at, the first element of TTP is present in the vector transformation account.

### 3.1.2 The Transmittability Of The Signal

The very nature of a network's operation involves the transmission between unit. This is explicit in the definition of a network: a network is set of individual but interconnected unit. Notwithstanding this, the requirement which must be met is not merely that something be transmitted, the requirement is that a particular piece of information (namely the information which meets the first element of TTP) be transmitted. For the

---

[7] Churchland, P. S., Sejnowski, T. (1992). *The Computational Brain*. London: MIT Press. p. 163.
[8] Churchland, P. S., Sejnowski, T. (1992). *The Computational Brain*. London: MIT Press. p. 164.
[9] See, for example: Boden, M.A. (1988). *Computer Models of Mind*. New York: CUP.; Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*. London: MIT Press; Gardner, H. (1985). *The Mind's New Science: A History of the Cognitive Revolution*. New York: Harper Collins.

case at hand, the inputs and the outputs must be vectors. Evidence for this is found not only in the theory of what vector transformation (discussed in section two of this chapter) but in the following direct quote concerning what is transmitted in networks.

> The in going and out going representations are vectors, meaning ordered sets of values.[10]

Hence, the second of the two elements are present in vector transformation accounts. The final requirement is really the requirement which is at issue here. Is there (and is there necessarily) a relationship between input and responses in a receiver such that there is a unique response to the same input each time it is presented?

### 3.1.3 The Reproducibility Requirement

Recall that TRR states that

> a receiver must have a unique respond to an input $C_1$ such that the receiver reproduces the signal as it was transmitted, save noise.

In the case for a linear associator network where the task which is claimed to result is a mathematical task, the necessity of single well established response arises from the claim that a mathematical function is being performed. This claim is present in such account as the below quote indicates.

> The key mathematical task that networks can perform is computing inner products, that is taking two vectors multiplying them component by component

[10] Churchland, P. S., Sejnowski, T. (1992). *The Computational Brain*. London: MIT Press. p. 77.

and then adding up the products.[11]

Given that in mathematics relationships are well defined, in a trained-up network the input output relationship must also be well defined and not transient. The same input *i1* at time *t1* must produce the same output *o1*, at times *t2.....tn*, barring noise. This was shown in section two of this chapter to be the operation of an artificial network (one which is fully trained up). In such networks, vector transformation was said to occur when a received input (in combination with the weightings at the hidden units) produces an output. And, barring error (noise or malfunction), the system will always produce the correct output (vector) given an input (vector) and a set of weightings (the weighted matrix).

The claim that a mathematical task is being performed not only provides the evidence that the requirement is met, it also provides the evidence that it necessarily has to be met. For if it is not met, either vector transformations are not happening or a vector transformation is only part of the story and the account is incomplete.

An argument from chapter three applies here. In chapter three a case of transmission involving Morse code was considered. It was concluded that, if a system is to process inputs, transmit the result and repeat the activity until a "solution" is secured, then that system requires the kind of precise information transmission which is secured by TTP and TRR. Without TTP and TRR the result of applying a vector transformation would not be

[11] Churchland, P. S., Sejnowski, T. (1992). *The Computational Brain*. London: MIT Press. p. 78.

preserved across transmission for further processing.

Given that TTP and TRR are necessarily required for vector transformation models, the noise/signal distinction which is a necessary consequence of TTP and TRR is necessarily necessary for vector transformation models. Hence, vector transformation models are committed to the same noise signal distinction as Shannon's account.

## 4 Discussion

The two elements of the transmutability principle and the reproducibility requirement are found in vector transformation accounts. Hence, TTP and TRR are implicit in vector account transformation accounts and hence (given what has been established in chapter three), TDN is necessary within vector transformation accounts. Moreover, from chapter three the condition WRC or WRC* is also a consequence of TTP and TRR.

While artificial networks may be designed to operate in a manner that adheres to TTP, TRR and the resulting noise/signal distinction, actual biological networks (of the type discussed in chapter three, in which modulatory activity is manifested) do not operate in such a manner. Therefore, while vector transformation accounts might well be an account of information conveyance in artificial networks, the same cannot be said of actual biological networks, at least not in the same sense of the term "information" and not in the sense discussed by Shannon. In chapter three biological system were argued to be not described fully by Shannon's noise/signal distinction nor to operate in accordance with any of the rigidity conditions.

Said another way, the claim is that the rigorous application of the notion of vector transformations precludes what was argued (in chapter three) to be required in any explanation of neural systemic activity. That claim was that the models must somehow account for and distinguish between (a) inputs which are noise, (b) inputs which are 'information' and (c) inputs which "enable" (b). Since mathematical models do not include distinctions between these, pure mathematical models (like vector transformation accounts) cannot be considered as complete explanations.[12]

---

[12] In a forthcoming article C.B. Martin offers both an argument against pure vector accounts of systemic activity and an alternative to these accounts. In arguing against pure vector accounts Martin notes that the consideration of causal interaction and potential interaction in terms of numbers is, at most, "the partial consideration of what is more that number, yet expressible by number". Martin's concern is that leaving the account in purely numerical terms may amount to the **denial** of that which the numbers are measures of (namely the qualities). If a theory includes the denial of qualities then, according to Martin, that theory is committed to Pythagoreanism. According to Martin, Pythagoreanism is a form of "unfettered deontologizing which results in a world of pure number [that] seems as clear a *reductio* as any in philosophy."

Working in back of Martin's criticism is a distinction he draws between qualities and quantities. A quantity is always a quantity of something, namely a quality, either actual or potential. Martin's point is that while quantities are a needed part of a theory so too are the qualities. Footnote 34 of chapter three made reference to essential elements in Martin's ontological account of systemic activity. One element was the material of use. A full account of the material of use is an account of both the qualitative and quantitative aspects of the material. To do this is to include "the different research emphases of each [qualities and quantities] made necessary to the other." This, Martin claims, is not "suggesting the addition of 'an excess of work from nowhere,' but of making salient and evident to the mind what is intrinsic to the properties in virtue of which the work is done."

Martin's position posses a problem for my discussion on vector transformation models insofar as my discussion could be claimed to be incomplete. While I do believe that Martin's distinction holds, the investigation in this thesis does not explicitly embrace Martin's ontology. Rather, the investigation is an inquiry into the problems that information theory has given its limited tools. Martin's ontology would find additional problems with a pure vector account in terms of its inability to distinguish between signals which are quantitatively similar but qualitatively distinct. If Shannon's information theory were to be placed into a

## 5 Possible Reply and Supplementation

As a possible reply to the claim that vector transformation models do not provide tools rich enough to account for the three kinds of input one may argue that the neuromodulatory activity is limited to a class of neurotransmitters and not all neurotransmitters. Hence, a possible reply may be that there are two classes of transmitters one in which vector transformations are occurring and another class that switches a neuron's state such that thhese neurons can perform a different vector transformation. While this account still fails to deal with learning, it does allow for the same input to have different responses at the receiver, provided that a modulatory input alters so-called vector transformational properties of the receiving neuron.

The strategy of the reply is to acknowledge that the vector transformation account requires supplementation and to provide supplementation which adheres to the three kinds of signals shown to be required in chapter three.

While this is a plausible account, the evidence indicates that such an account is (like the initial account) deficient. It is true that different neuromodulatory transmitters acting on different modulatory neurons evoke changes in either the intrinsic membrane properties

---

pure vector transformation form, it too would be open to Martin's objection.

A clearer understanding of Martin's motivation and theory requires a more complete account of his ontology. Interested readers can find a clear account of Martin's views in his debate with D.M. Armstrong and U.T. Place. See: Armstrong, D.M., Martin, C.B., Place, U.T. (1996). *Dispositions: A Debate*. New York: Routledge; Martin, C.B. (forthcoming). 'On the Need for Properties: The Road to Pythagoraenism and Back.' *Synthese*. (forthcoming).

of neurons or synaptic efficacy. Writing about modulatory priorities Harris-Warrick *et al* write:

> These properties radically change the activity of the neuron within the network as well as changing its responses to synaptic input.

However this is not the whole story: even a neuron's response to a modulatory input will vary. Harris-Warrick *et al* note that,

> [t]he response to a modulatory input varies with the state of the system.

This point was made in chapter three when it was noted that temperature and $Ca^{2+}$ levels can determined a neuron's response to modulation. Hence the modulator input does not simply switch between states; rather, it is one of a number of required elements required. The kind of control which is needed is much richer than assumed for the case where neurotransmitters simply modulate activity.

Theoretical attention which is merely focused on transmission and not on the state of the system provides a gross misrepresentation of what is happening in the system.[13] Included in this systemic activity are subtle variations in response which may be manifested but also may never be manifested; nonetheless such activity is the activity which enables further systemic operation to be carried out. It is this activity which is

---

[13] As noted above, this point is central to both C.B. Martin's ontology and his ontological theory of mind. I am indebted to Martin for the point.

absent from a vector transformation account, and this activity which has been shown to be of central importance for the systemic capacities and the activation of such capacities in biological networks.

What is shown, however, is the incompleteness of the condition initially put forward. That claim was that the models must somehow account for and distinguish between (a) inputs which are noise, (b) inputs which are 'information' and (c) inputs which "enable" (b). However, in accounting for such, the inputs cannot be taken as the whole story. The system's existing state is seminal for any input, including modulatory neurotransmitters and local neurotransmitters. Hence, while the three part distinction for inputs is part of the required story, so too is an account of existing neural states such that the input may or may not respond or respond alternatively.

Of interest is the fact that central to any adequate theory is the need to account for the variability in responses and to account for it in such a way that the system both limits and exploits such activity. Accounts which do not do this are and will be deficient and incomplete. Part of what is being argued for here is the need to account for systemic activity (not in terms of pure input driven models but) with a theory that makes full and relevant use of benefits which are afforded to a system by controlling and exploiting properties of the material in the system and ultimately the activity in the system. As noted at the end of chapter three, C.B. Martin provides such an account.[14]

---

[14] See note 36 chapter three.

## 6 Chapter Summary

The proceeding discussion could be summarized as follows:

1. In vector transformation models the transmittibility principle (TTP) and the reproducibility requirement (TRR) are present and necessarily present if Shannon's notion of information is to be retained in this new setting.

2. Hence, given chapter three, so too is the definition of noise (TDN).

3. Given TTP, a vector model has the same notion of a signal, though not necessarily Shannon's notion of information.

But

3. Given what was argued in chapter three, Shannon's noise/ signal distinction may be fine for something like a silicon system but Shannon's distinction won't do for neuronal interaction.

4. Therefore, vector transformation models may fit artificial networks but] they do not (and cannot without giving up any attempt to retain Shannon's notion of information) fit biological networks.

5. IF such models retain the notion of information developed by Shannon with only such modifications as are forced by the move to vector transformation settings, then vector transformation models will be too limited to be useful in trying to make sense of biological systems.

# Chapter V

# Objections and Conclusions

## 1 Review

This thesis has been an inquiry into the suggestion (a suggestion which some have claimed is a *commonplace)* that nerve impulses convey information. To this end Claude E. Shannon's measure of information and his accounts of information and information conveyance were investigated. It was concluded that Shannon's theory cannot be taken as an account of information conveyance in biological systems. Two different arguments were provided for this conclusion: the first appears in chapter two and the second, in chapter three. Chapter four was an application of the second argument to vector transformation models. These arguments are summarized thus:

(a) Argument One (Chapter Two):

1. The amount of information in a signal is a function of how surprising it is that the elements of that signal are as they are, that is, are in the order that they are.
2. The resources provided by Shannon *et al* for spelling out what "surprisingness" comes to require, as they stand, the possibility of fixing the cardinality of the possible alternative orderings for that signal.

But
3. Though this may be fine for something like a digital computer or telegraph system, it won't do for a neural system [see lemma L1 below]
4. Therefore, Shannon's theory is too limited in range of possible applications to be useful as a complete account of conveyance.[1]
LEMMA L1:

1. There are two features [relevant here] of neuronal activity which are salient:
(a) plasticity and
(b) adaptivity
2. The plasticity and adaptivity of neuronal activity mean that the size of the set of alternatives is (i) not a value constant through time and (ii) not, for any given time, not a specific number.
3. Hence there is no determinate ratio of something to the number of possible alternatives.
4. Hence the probability of the occurrence of a specific sequence of elements is not determinate.
5. Hence there is no possibility of identifying either the measure of information in a sequence or the information in that sequence.

(b) Argument Two (Chapter Three):

1. Shannon's account of information conveyance is a function of three things: the transmittability principle (TTP), the reproducability requirement (TRR), the notion of information (TNI). A fourth thing, the definition of noise (TDN) is a necessary consequence of TTP and TRR.
2. The resources provided by Shannon *et al* for spelling out the condition required for the conveyance of information (and hence the realization of these four items) in physical systems is limited to a specific condition. [see lemma L2a below].
But
3. Though the condition may be fine for something like a silicon system, it won't do for neuronal interaction [see lemma L2b below].
4. Therefore, Shannon's theory is too limited in range of possible applications to be useful.

---

[1] G.C. Teskey has noted that Shannon's model may be useful may be useful in terms of approximation probability theory however, the ontological commitments of such a claim are antirealist.

LEMMA L2a:

1. TTP requires that information be transmittable as a signal $C_I$ and a signal $C_I$ be fully formed at the point of transmission.
2. TRR requires that a receiver have a unique response to an input $C_I$, such that the $C_I$ received is the same $C_I$ transmitted.
3. Hence, the first condition is that a receiver has one and only one response to an input $C_I$.
4. TNI states that information occurs in statistically structured sequences of elements.
5. How specific $C_I$ has to be is a function how well structured the code is that informs the statistical structure of information.
6. Hence the second condition is that $C_I$ depends upon the structure of the code.
7. Conditions one and two specify the rigidity condition.

LEMMA L2b:

1. There are three features [relevant here] of neuronal activity which are salient:
(a) plasticity;
(b) adaptivity and
(c) neuromodulation.
2. The plasticity, adaptivity and neuromodulatory activity of neurons mean that the same signal can have many different responses in a receiver.
3. The Shannon *et al* noise/signal distinction does not allow for this.
4. Hence Shannon *et al* cannot provide an account of learning, neuromodulation or adaptive responses.
5. The Shannon *et al* distinction is a necessarily necessary distinction given Shannon's account of information transmission.
6. Hence Shannon *et al* necessarily cannot provide an account of what is basic to neural systems.

(c) An Application of Argument Two (Chapter Four):

1. In vector transformation models the transmittability principle (TTP) and the reproducibility requirement (TRR) are present and necessarily present if Shannon's notion of information is to be retained in this new setting.
2. Hence, given chapter three, so too is the definition of noise (TDN).
3. Given TTP, a vector model has the same notion of a signal, though not necessarily Shannon's notion of information.
But
3. Given what was argued in chapter three, Shannon's noise/signal distinction may be fine

for something like a silicon system but Shannon's distinction won't do for neuronal interaction.
4. Therefore, vector transformation models may fit artificial networks but they do not (and cannot without giving up any attempt to retain Shannon's notion of information) fit biological networks.
5. IF such models retain the notion of information developed by Shannon with only such modifications as are forced by the move to vector transformation settings, then vector transformation models will be too limited to be useful in trying to make sense of biological systems.

Given argument two of chapter three, the measure of information is hampered by something other than simply our inability to obtain all variables required to perform the measure. Shannon's measure of information does not apply to human capacities because the conditions necessary for Shannon's account of information conveyance do not hold in biological systems, let alone humans.

There are, as expected, objections against what has been argued. What follows is an account of, and reply to at least some of these objections.

## 2 Objections And Replies:

**Objection 1: The Objection to TTP**

Surely Shannon could work with something less than the transmittability principle? If he could work with some other account, then he could have a richer notion of noise, one compatible with plasticity and adaptivity thereby avoiding the problems raised in this thesis.

**Reply**

It is possible to have something other then TTP and thereby have a different account of noise/signal. In fact, the purpose of this thesis was to show the need for something other than TTP and Shannon's account of information conveyance. But, Shannon's theory requires TTP. Given what he takes information conveyance to be, Shannon is committed to TTP. Recall that on Shannon's account information conveyance involves: (a) a signal which is fully formed at the point of transmission, (b) preserved across transmission and then (c) reproduced at the receiver. TTP is nothing more than an expression of Shannon's very conditions. For Shannon to give up TTP (so that he can account for information conveyance in biological systems) is for Shannon to abandon the very core of his account of information.[2] The purpose of this thesis was to show the need for just that.

**Objection 2: The Objection To TRR**

If TTP cannot be done away with, surely Shannon could work with something less than TRR? If he could work with some other account, then he could have a richer notion of noise, one compatible with plasticity and adaptivity, thereby avoiding the problems raised in this thesis.

---

[2] The strength of this claim has been increased as a result of a comment by J.A. Baker.

**Reply**

Yes, it is possible to do way with TRR in a theory of information conveyance; the question is, if you gave up TRR, would you still have Shannon's theory? While the purpose of this was to show the need for something other than TRR, it was also argued that Shannon cannot do away with TRR. Given what he takes information and information conveyance to be, Shannon is committed to TRR. Recall the *reductio* presented in chapter three, where this was argued. It proceeded thus,

> For a *reductio* consider what it would be for there to be different responses in a receiver to the same input. Consider a telegraph with only two possible responses (call then long and short) to two kinds of signals (call them long and short also). The signals may be sent in various sequences. Assume that there is no fixed response to a signal. For example a long signal can be received as a long or a short signal and a short signal can be received as either a long or a short signal. If a series of signals is transmitted in the form of three long, three short and three long signals, the signal that is received may be any combination of three groupings of three. If the signal described is Morse Code, then the characteristic international signal for help may or may not be received as such. Rather than S-O-S, any of the following combination of three letters, symbols or short expressions may be received: d, g, k, o, r, s, u, w, 'semi colon,' 'understand.'[3] To count this as information transmission is absurd given that the definition becomes to broad to be of use. Even if one rejects this as a reduction (embracing the very broad definition of what it is to transmit information), this is not the kind of information transmission Shannon was concerned with. Nor is it the kind of precise information transmission that a system requires when that system processes inputs, transmits the result for further processing and repeats this until the 'solution' is obtained. John L. Tienson (1987) makes a similar point when he writes that the causal activity in a computer must be "exceptionless."[4] Hence, the

---

[3] *The World Book Encyclopedia* (1991). Chicago: World Book Inc. pp. 880-881.

[4] Tienson was writing about connectionist networks and how to construct a computer. When building a von Neumann machine he wrote that "the trick of making a computer is to get the causal process of the mechanism to mirror the syntactic processes specified in a programmes, for this to be done the rules must be precise and exceptionless." If in building a computer one needs to align causal process such that they mirror syntactic processes, and the syntactic processes are precise and exceptionless so too are the causal processes. Exceptionless requires that the same input produce the same response without fail. Tienson, J. L. (1987). 'An Introduction to Connectionism' *The Southern Journal of Philosophy 26:* Supplement, p.1.

> relationship between the receiver and the transmitter is such that what is transmitted has (in the absence of noise) a single unique response in the receiver.

**Objection 3: The Objection To The Necessity Of Shannon's Account Of Noise**

Surely, Shannon's account could do with some other account of noise then TDN? If he is not committed to TDN, then the second argument is too strong and the problems raised for vector transformation models are too strongly stated.

**Reply**

Shannon's account of information conveyance necessarily has an account of noise; moreover, the account Shannon gave is necessarily in the form it is. This was argued in chapter three to be a function of TTP and TRR. Shannon's account of noise was argued to be a necessary consequence of TTP and TRR. Since TTP and TRR are necessary for Shannon's account of information conveyance, to opt for a different notion of noise is to opt for an account of information conveyance that is not Shannon's account. The purpose of this thesis was to motivate the need to find a different account of information transmission than Shannon offers.

To see again that TDN follows from TTR and TTP recall what TTP and TRR are.

> The Reproducibility Requirement (TRR): A receiver must have a unique response to an input $C_1$ such that the receiver reproduces the signal as it was transmitted, save noise.

> The Transmittability Principle (TTP): Information sequences are capable of being transmitted as signals $C_1$. When transmitted, the sequence (within a signal $C_1$) is *fully and completely fixed at the point of transmission.*

If you accept TTP and TRR, then a signal is based on the operation of a transmitter and the notion of noise is based on: (a) additions or alterations to the signal which leave the transmitter and (b) alterations in the receiver which prevent the receiver from receiving the transmitted signal as it was transmitted. If you have this you have Shannon's notion of noise.

Turning to the second part of the objection, even if it could be shown that the Shannon notion of noise is not a consequence of TTP and TRR, the problems raised for vector transition models still hold. Vector transition models work because they make use of a specific set of systemic assumptions; these assumptions are not consistent with the activity seen in biological systems. For example, neuromodulation poses a problem for vector transformation models in that an input (given what modulation does to the post-synaptic neuron) can have many different effects on that neuron such that the neuron responds differently to the same input.[5]

What is not claimed here is that connectionist models will never be able to model biological systems, for surely one day such a model may incorporate the required (what

[5] See for example chapter three section four.

ever they are) properties and model, within limits, a certain systemic activity. Rather, what is being claimed is that ***the theory*** (vector transformation) which is suppose to inform us about what the model is teaching us about what the system is doing ***is not rich enough to fulfill the task at hand.*** Modeling is one thing but what that model is supposed to teach us is another thing. A model informs us of at least two thir.gs: firstly a model can mirror and be used to predict the dynamics of a system and second, a model can be used to better understand the activity in the system and supply a theory for what is going on, possibly in terms of information conveyance. When theory (as Churchland proclaims it is) is being 'bootstrapped' up, and the models which are model of systemic activity are broadened to incorporate different variables in them (variable whose importance have come about from other empirical research), the inclusion of a new variable may have negative, positive or no consequence(s) for the theory which is taken to inform the model. (The separation of the theory from the model is required to allow for the possibility that (for example) connectionist models may exist which do not merely perform vector transformations in the strict sense required to merely perform a vector transformations but yet are connectionist models). While a model may mirror a system's dynamic behaviour, our theory about what is going on may not be rich enough to account for the activity in the biological system nor what is observed in the model. At this point, the theory is in need of revision; but this SHOULD be expected as this is exactly what it is to bootstrap a theory.

As an aside, it is important to note that so long as it is possible for a notion of noise (any notion of noise) to be present in Shannon's account, then the objections raised in the first argument hold. This follows from the fact that the first argument does not turn on any particular notion of noise; it turns on what is required for Shannon's measure of information to be applicable (that there be a determinate ratio of the number of things in a 'message' to the number of alternatives). This is not obtained in biological systems, given both plasticity and adaptivity. Hence, even if the objection against TDN holds, more is required to show that Shannon's measure of information can be applied to biological systems.

**Objection 4: An Objection To The Discussion Of Learning and Plasticity**

Surely, given the relationship between information theory, cybernetics and neuroscience, someone has either seen the problem of plasticity or discussed systems that can modify their internal relations? What has been said by theorists about plastic systems and information theory? Does their writing present a solution to either the problem of learning or noise?

**Reply**

In 1953, Donald MacKay wrote a paper 'Generators of Information' dealing with systems that alter their transmitter/receiver relations.[6] In his paper MacKay does provide for a

[6] MacKay, D. (1953). 'Generators of Information' in *Communication Theory*. Jackson, W. (ed.), New York:

distinction between two kinds of systems, those in which code is imposed (at the time of design) and those in which the code results from changes to transmitter and receiver relations in response to input. The latter are called code-generating artifacts. While MacKay discusses such systems, his paper contains no discussion on how such systems affect Shannon's basic notion of information and information conveyance. Unless MacKay wants to admit that no information is conveyed during the period the code is being altered (a period of learning), he must choose some other option. However, to claim that Shannon's strict conditions are not required to be fulfilled is to be committed to (what was argued in chapter three as being) an *ad hoc* claim. Apart from MacKay very little has been said in the literature on the relationship between plasticity and information theory. Given that vector transformation theories are relatively new and given that learning is also a problem for them, it is not clear that the relationship between information theory and plasticity or adaptivity or neuromodulation for that matter has been investigated. I fail to see what could be said apart from saying that the account needs to be changed if it is to account for such. To do so is to require a different noise/signal distinction that Shannon has. A different noise/signal distinction requires that either or both of TRR and TTP be given up. To do so IS to give up Shannon's account.

---

Academic Press. pp. 475-485. Reprinted in MacKay, D. (1969). *Information, Mechanism and Meaning*. London: MIT Press. pp. 132-145.

**Objection 5: The Objections Raised In This Thesis Are Too Simple Minded To Be Beyond Reply**

The line of objection in this thesis is simple minded, surely there exists a reply to what has been argued?

**Reply:**

True, the line of objection is simple minded; in fact, what is argued is really quite obvious. However, despite this I am unable to think of a reply. Since the mere possibility of a reply is not itself a reply to the objections raised in this thesis, the problems will remain as problems, for the time being.

Rather than giving rise to an objection, the simple mindedness of the argument provides strength. The argument turns on two point TTP and TRR. TTP provides an account of what a signal; $C_l$ is (something, which is transmittable and formed - fully formed at the point of transmission). TRR, with reference to $C_l$, provides for an account of noise. Noise is what $C_l$ is not when $C_l$ is received or noise is what prevents $C_l$ from being received in the form $C_l$ was transmitted.

With TTP and TRR a necessary noise signal distinction results. Signal is what is transmitted and noise is what is not the signal, and what prevents the signal from being received in the form it was transmitted. $C_l$ is kept undefined to allow for a receiver to have various degrees of sensitivity to inputs: different inputs can elicit the 'correct' response in the receiver.

While Shannon's exact account of information requires TNI, a vector account can also play the role of $C_I$ and since it is not clear that a vector account is statistically structured in the same way as Shannon's account, $C_I$ was left a little vague, to allow for such various possibilities. This allowed the weakest possible account which would still be Shannon's account but consistent with much of today's findings.

While Shannon did not state TTP nor TRR they are nonetheless a combination of both his definitions of a transmitter and a receiver and his assumption of what it is for information transmission to occur in a system.

Given that TTP and TRR have a necessary noise/signal distinction (the signal being explicitly defined by TTP and the notion of noise being implicit) the central objection to my argument arises with the claim that Shannon could not do with other than TTP and TRR. These objections were dealt with above.

In short, TTP and TRR are Shannon's account of information conveyance, if, that is, you allow on Shannon's behalf the possibility that he may want to change his exact measure of information. If one requires that for an account to be Shannon's account, necessarily his actual measure of measure of information must be included then TNI, along with TTP and TRR are required to account for Shannon's account. However, rather than assisting Shannon, this is a stronger requirement. If TTP and TRR alone provide for too rigid a notion of noise and signal the addition of TNI will not help.

The cumbersome and often tedious presentation of details in both chapters two and three provide a rich and complete account of what is basic to Shannon's account. In this, I hope, the foundation for this very objection to my position lies: that the arguments presented are simple minded. When Shannon is clearly explained it is (I hope) clear that TTP and TRR (TNI is optional, as noted above) ARE Shannon's account. [As an aside, although relevant, the details are not what the arguments turn on. For example, neither Shannon's actual formula nor the details of adaption are actually required to establish the points. The same points could have been made with thought experiments and hypothetical cases.]

Not only are the arguments presented simple minded they are obvious. Despite this the obviousness has not been addressed nor faced by many of today's theorists (at least those who embrace an information processing or vector transformation model). Rather, as was shown with the discussion of vector transformation models, many of Shannon's notions are used, even today. It is because of this that the above presentation may find some justification for the amount of exposition of vintage theory.

Consider the question: How can something open to such a simple minded refutation be accepted by (and be implicit in the theories of) so many as being so obviously right as to escape any serious inquiry? The first obvious response to this question is that what has been presented as Shannon's Theory is not, in fact, Shannon's Theory. Rather (or so the objection could continue) it is a straw man version of Shannon's powerful theory. In part the large amount of exposition is intended to address this question, by replying with thoroughness. A fairly complete account of Shannon prevents (or limits the likelihood) of this kind of an objection.

In passing, my supervisor, Dr. Baker, asked whether what I was doing as anything more than philosophical archeology. While this is probably an accurate assessment, what is also true is that much of today's theory turns on Shannon's very notions or at least employs them. And, while too much time was spent digging for and trying to tease out (and get clear about) what lies in back of today's conceptual work on neural computational theory (by looking at it and its foundations), there may be some value in the holes I dug (and almost failed to climb back out). The apparently anti-climatic realization that TTP and TRR are what lies in back of many of today's theories (at least Shannon's account of information conveyance and vector transformation models) resulted from a form of philosophical archeology however, because of the digging insight into the puzzling question (noted above) may be found. While I do not claim to have the full answer to the question, my suggested reply lies in the fact that TTP and TRR taken together provide for an account of communication and information conveyance, which is intuitively appealing. The account is intuitively appealing insofar as it fits with what is required for a vector transformation account and possibly a symbolic account and even a linguistic account, where the symbols, words or vectors are formed and then transmitted. While this last comment requires much fleshing out, I cannot think of a single linguistic account, symbolic account or vector transformation account that does not require that what be transmitted be fully and completely formed at the point of transmission. This is the requirement of TTP. Given the reduction for TRR, it not hard to see that TRR also fits the intuitive appealing notion of information conveyance. Despite the nice fit with an intuitive account of information conveyance, TTP and TRR require conditions and have a noise/signal distinction which are not found in neural interaction. It is THIS point that my thesis is mostly about.

This last comment I will leave as a comment and not an argument, despite the fact that I cannot think of a language or symbolic account, or vector account that avoids the requirement that what is transmitted be fully formed at the point of transmission.[7]

[7] This thesis would have been finished a lot sooner, if I know both where to look and what I was looking for. When I began reading computational accounts of systemic activity all I knew was that there was something I did not like, something that seemed fundamentally wrong about the approach.

### Objection 6: The Objection That The Argument Shows Too Much

Given that Shannon type accounts are widely used, this thesis may show too much. Much has been accomplished with the basic tools this thesis questions. For example, connectionist modeling has allowed for an interesting account of learning and memory within distributed activity. Surely, Shannon's account can be said to hold in some form or another.

### Reply

How widely used a theory is, is no defense for the problems raised in this thesis. Hence, this is not an objection. If, on the other hand, the objection is that a theoretical vacuum is created then this is an objection, of sorts. However, the claim required by this objection is too strong. No theoretical vacuum is created. For example, the work of C.B. Martin which has been referred to (largely in the footnotes) numerous times in this thesis offers at least one alternative explanation to Shannon's account of information conveyance. The intent of this thesis was to take a serious look at what has received little attention and ask if what we have come to learn about biological systems has any consequences. Shannon's account of information, an account which dates back to 1948 and it should not be surprising that recent findings pose a problem for (a) a theory which is purported to address the 'technical' problem of information conveyance and (b) a theory that is approaching 50 years in age. Surely our technical understanding has changed over some fifty years.

While ***some*** benefits have come from modeling (such as connectionist modeling), greater understanding may be found elsewhere. It is for this reason that (in the above reply to objection three) a separation between a modeling technique and a theory about a particular class of models in that technique was made.

The lessons learned about distributed network activity may have been learned without a vector transformation account and without Shannon's account of information or information conveyance. One can easily imagine a model that is distributed in activity but not a pure vector transformation model nor one which employs Shannon's account of information conveyance. Hence, the fact that some theoretical benefits have arisen does not mean that Shannon's account is right nor does it mean that the theory which has arisen is correct. There is always the possibility that much needs to be rethought, although this is only a possibility.

3. Glimpses Beyond

What has not been established in this thesis is the extent to which Shannon's form of information conveyance is present in theory of brain function. This must be investigated. Included on a list of candidates for investigation are Fodor's computational theories of mind and Turing's actual description of the Turing machine.[8] The goal is simply to show that the activity in biological systems does not reduce down to a Turing machine without remainder. This comment require more fleshing out than can be established here. However, theses comments find preliminary justification in the fact that even Turing seems to have employed a form of Shannon's information conveyance.

Martin's work is deeply appealing in that it provides a framework for theory which is not founded on or implicitly constrained by a Shannon type account. It is for this reason that many of Martin's theoretical tools are made note of in the footnotes.

I have not in this thesis made an attempt to discuss Martin; nor have I made full and relevant use of his ontology. The intent of this investigation was not to merely accept an ontology (Martin's ontology) and see what follows from it but to see whether there are deep conceptual difficulties which can be shown to exist for a widely and largely un-investigated theory of conveyance using the simple tools supplied by that theory. The deeper questions addressed by Martin concerning ontology and the need for a rich enough foundation upon which to base a theory at least finds support within the arguments presented here: support that comes not from within Martin's particular ontology but from discussion about the limitations of an account which employs notions of cause and effect

---

[8] Fodor, J.A. (1975). *The Language of Thought*. Cambridge, Mass.: HUP.

that are apparently impoverished, at least theoretically impoverished.[9]

If nothing else, the forgoing (momentary) indulgence with philosophical archeology has served to bring to the surface what must have been suspected by many. The few trinket brought to the surface hopefully increase understanding of what has been said in the past and what is still present today.

[9] See: See: Armstrong, D.M., Martin, C.B., Place, U.T. (1996). *Dispositions: A Debate*. New York: Routledge; Martin, C.B. (forthcoming). 'On the Need for Properties: The Road to Pythagoraenism and Back.' *Synthese*. (forthcoming).

Bibliography Of Work Cited

Armstrong, D.M., Martin, C.B. and Place, U.T. (1996). *Dispositions: A Debate*. New York: Routledge.

Boden, M.A. (1988). *Computer Models of Mind.* New York: CUP.; Clark, A. (1989). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing*. London: MIT Press.

Chalmers, D. (1996). *The Conscious Mind.* New York: OUP.

Churchland, P.S. (1986). *Neurophilosophy: Toward a Unified Science of the Mind/Brain.* Cambridge, Mass.: MIT Press.

Churchland, P. S. and Sejnowski, T. (1992). *The Computational Brain.* London: MIT Press.

Clark, A. (1991). *Microcognition: Philosophy, Cognitive Science and Parallel Distributed Processing.* London: MIT Press.

Cooper, J.R., Bloom, F.E. and Roth, R.H. (1991). *The Biochemical Basis of Neuropharmology, Sixth Edition*. Oxford: OUP.

Cote, L., Crutcher, M. (1991). 'The Basal Ganglia' in Kandel, E., Schwartz, J. and Jessell, T. (*eds.*) (1991). *Principles of Neural Science. 3rd ed.* New York: Elsevier.

Crick, Francis (1994). *The Astonishing Hypothesis: The Scientific Search for the Soul.* New York: Charles Scribner's Sons.

Gardner, H. (1985). *The Mind's New Science: A History of the Cognitive Revolution.* New York: Harper Collins.

Getting, P.A. (1989). 'Emerging principles governing the operation of neural *networks'* *Annual Review of Neuroscience 12*:185-204.

Gregory, R. L. (1986). 'Whatever Happened to Information Theory?' in Gregory, R. L. (1986). *Odd Perceptions*. New York: Methuen.

Gregory, R.L. (ed.) (1987). *The Oxford Companion to The Mind.* Oxford: OUP.

Harris-Warrick, R.M., Marder, E., Selverston, A.I. and Moulins, M. (eds.) (1992). *Dynamic Biological Networks.* Cambridge: The MIT Press.

Hartely, R.V.L. (1928). 'Transmission of Information' *Bell System Technical Journal.*

Hebb, Donald, O. (1949). *The Organization of Behavior: A Neuropsychological Theory.* New York: John Wiley & Sons.

Heil, John and Martin, C.B. (forthcoming). 'Rules and Powers.' *Philosophical Perspectives.* (forthcoming).

Heil, J. (ed.) (1989). *Cause, Mind and Reality: Essays Honoring C.B. Martin.* Dordrecht: Kluwer.

Hicks, E. (1952). 'On the rate of gain of information', *Quarterly Journal of Psychology.* **4**:11-26.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings National Academy of Science. 79:* 2554-2558.

Hyman, R. (1953). 'Stimulus information as a determinant of reaction time' *Journal of Experimental Psychology.* **45**: 188-196.

"Information Theory" from Britannica On-line, Copyright (c) 1995 by Encyclopedia Britannica, inc. Downloaded on May 10, 1997.

James, William. (1890). *The Principles of Psychology. Volume Two.* New York: Dover.

Kandel, E., Schwartz, J. and Jessell, T. (*eds.*) (1991). *Principles of Neural Science. 3rd ed.* New York: Elsevier.

Llinas, R. (1988). "The Intrinsic Electorphysiological Properties Of Mammalian Neurons: Insights Into Central Nervous System Function." *Science 242:* 1654-1664.

MacKay, D. (1953). 'Generators of Information' in *Communication Theory.* Jackson, W. (ed.) New York: Academic Press. pp. 475-485. Reprinted in MacKay, D. (1969). *Information, Mechanism and Meaning.* London: MIT Press.

MacKay, D. (1969). *Information, Mechanism and Meaning.* London: MIT Press.

Martin, C.B. (1993). The Need for Ontology: Some Choices. *Philosophy 65:* 505-522.

Martin, C.B. (forthcoming). On The Need for Properties: The Road to Pythagoreanism and Back. *Synthese.* forthcoming.

Martin, C.B. (unpublished a). 'Toward A Model For Mind And Brain: Some Prepatory Notes.' unpublished.

Martin, C. B. (unpublished b). 'What is Imagistic About Verbal Imagery and Why Does it Matter.' unpublished.

Martin, C.B. (unpublished c). 'Lectures Notes' Philosophy 695. Winter, 1994.

McCulloch, W. S. and Pitts, W. (1943). 'A logical calculus of the ideas imminent in neural nets' *Bulletin of Mathematical Biophysics 5:* 115-133.

Nicholson, W.K. (1986). *Elementary Linear Algebra With Applications.* Toronto: Wadsworth Publishers of Canada.

Rafal, R, *et al* (1990). 'Extrageniculate Vision In Hemianopic Humans: Saccade Inhibition By Signals In The Blind Field.' *Science* ***250:*** 118-121.

Schneider, G.E. (1969). 'Brain Mechanisms For Localization And Discrimination Are Dissociated By Tectal And Cortical Lesions.' *Science* ***163:*** 895-902

Shannon, C (1948). 'The Mathematical Theory of Communication.' *Bell Systems Technical Journal. 27:* 379-423. Reprinted in Shannon, C.E. and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press.

Tessier-Lavigne, M. (1991). 'Phototransduction and Information Processing in the Retina' in Kandel, E., Schwartz, J. and Jessell, T. (eds.) (1991). *Principles of Neural Science. 3rd Edition.* New York: Elsevier.

*The World Book Encyclopedia* (1991). Chicago: World Book Inc. pp. 880-881.

Tienson, J.L. (1987). 'An Introduction to Connectionism' *The Southern Journal of Philosophy* ***26:*** Supplement.

Washington, P. (eds.) (1994). *William Blake: Poetry.* Everyman's Library. p. 94.

Weaver, W. (1963). 'Recent Contributions to The Mathematical Theory of Communication' in Shannon, C. and Weaver, W. (1963). *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press.

Weiskrantz, L., Warrington, E.K., Samders, M.D., Marshal, J. (1974). 'Visual Capacity In The Hemianopicfield Following Restricted Occipital Ablation.' *Brain* ***97:*** 709-728.

Wiener, N. (1961). *Cybernetics.* New York: MIT Press.

von Neumann, J. (1957). *The Computer and the Brain.* New Haven: YUP.

Youdim, M.B.H. and Riederer, P. (1997). 'Understanding Parkinson's Disease. *Scientific American:* 52-59.

# IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989