A Discipline Independent Definition of Information*

Robert M. Losee
Manning Hall, Room 100,
U. of North Carolina-Chapel Hill
Chapel Hill, NC 27599-3360

Published in
J. of the American Society for Information Science
48 (3) 1997, 254–269.

November 15, 1998

Abstract

Information may be defined as the characteristics of the output of a process, these being informative about the process and the input. This discipline independent definition may be applied to all domains, from physics to epistemology. Hierarchies of processes, linked together, provide a communication channel between each of the corresponding functions and layers in the hierarchies. Models of communication (Shannon), perception, observation, belief, and knowledge are suggested that are consistent with this conceptual framework of information as the value of the output of any process in a hierarchy of processes. Misinformation and errors are considered.

1 Introduction

The term "information" is used differently by individuals in different walks of life, from specialists working in information based professions, such as communication media and information management, to those in the computing and cognitive sciences, as well as by people involved in less scholarly pursuits. For example, communication

---

*The author wishes to express appreciation to Diane Sonnenwald, Lee Ann Roush, Jordan T. Davis, and to two anonymous referees for their suggestions.

scholars express concern about information overload due to the plethora of mass media sources. Electrical engineers strive to improve modem technology to increase the amount and speed of information transmitted, i.e., the number of bits per second transmitted over coaxial or fiber optic circuits. An interest in understanding information as well as the ability to measure it allows these specialists and others to work on their chosen problems, make reasonable decisions, and to communicate effectively among themselves.

Many have developed interesting and useful definitions of information for specific disciplines and classes of problems. Some definitions or measures may be consistent with ideas held by several disciplines. These interdisciplinary definitions are superior in some senses to discipline specific definitions useful in only one domain when the more general definition encompasses all the phenomena of interest to the field that is covered by the field-specific definition and is consistent with the field-specific definition. A more general definition allows frameworks, theories, and results to be transferred across disciplinary boundaries, and provides for dialogue across these boundaries, while at the same time allowing individual disciplines to focus on the specific information phenomena of their discipline. Unfortunately, people in different fields and professions differ on what information *is* or how to evaluate the different definitions that are assumed explicitly or implicitly by different fields or social groups.

We suggest that there are phenomenon common to what most definitions of "information" refer and that this phenomenon *is information* and that most definitions of information refer only to the subset of information as studied in that particular discipline. Information may be understood in a domain-independent way as *the values within the outcome of any process* [Los90]. By "value" we refer to a variable's attribute or characteristic, and not to economic value unless economics is explicitly mentioned [HR92]. Following Russell [Rus37], we view a variable as a component in a system whose "value" may be replaced by another value with the system remaining the same type of system as before.

This definition of information has the fortunate consequence of allowing different kinds of processes to be information carrying or information producing phenomenon. These processes may be viewed as ranging from the simple, mathematical function, to processes of such complexity, often involving humans, that some may believe they can't be studied or accurately characterized. The ability or inability to study the nature of a process will not affect our study of the value in the output of the process, and our work may be applied to both processes that are formally defined and to those that are viewed as non-formal in nature.

Viewing the processes together, in an organized fashion, facilitates the engineer, the linguist, and the television producer speaking in the same terms about how what they produce (the output of a process) "carries" information or is informative. A common language allows for theories and data to be transferred from discipline to discipline without confusion; term and concept migration are common as academic fields change and grow [LW86, Los95, Mul74, NCB72]. Considering how information and informativeness can be viewed consistently across different environments is the focus of this work. We provide examples of ways that different disciplines might wish to describe some of their information-phenomena of interest in terms of processes, making them compatible with our approach to information; these are suggestions, and other models also may be useful in integrating other discipline based information-phenomena into the proposed general definition of information. This general definition has the capability to model information as understood by physicists and those studying human behavior. It also has the advantage of being able to formally capture the subjective and individual processing and culture-specific nature of human processes and knowledge that affect information.

We propose that data from a variety of disciplines suggest and is consistent with this field-independent conceptual framework for information. Our definition thus is a generalization of the work of discipline-specific concepts of information provided by scholars such as Buckland [Buc91], Dretske [Dre81], and Shannon [SW49]. While we believe that our definition is consistent with the problems and the data found in a wide range of fields in both the humanities, social sciences, and "hard" sciences, our definition will conflict with other definitions that limit the domain of a field's study of information; this conflict arises between any general definition and more specific and limiting definitions. We are not saying that these field-specific definitions are wrong for the field in which they are defined; instead, we suggest a more general definition that can be applied to a broad range of fields and can facilitate communication about information-related phenomena.

Note that we do not consider "errors" or "misinformation" in our discussion of information until the penultimate section. The reader should assume processes are error-free until that point, unless otherwise stated.

2 Definitions of Information

One of the most common ways to define information is to describe it as one or more statements or facts that are received by a human and that have some form of worth to the recipient. For example, the Sesame Street character "Cookie Monster" describes information as "news or facts about something," or, as the first definition in the *Random House College Dictionary* suggests for information, "knowledge communicated or received concerning a particular fact or circumstance; news." Cookie Monster's definition is consistent with the common notions that information must:

1. be something, although the exact nature (substance, energy, or abstract concept) isn't clear;
2. provide "new" information: a repetition of previously received messages isn't informative;

3. be “true:” a lie or false or counterfactual information is *mis-information*, not information itself;
4. be “about” something.

This approach to information, like most human-centered approaches to information, leads one to emphasize the meaning and use of message, “what the message is about?” and “what is known already?” over the information carrying messenger and the message itself [Art73, AD75, BR76, Far80, Har84, Lev77, MM83]. When the message is essentially random, or the message is of no value to the recipient, such as a repeated message previously received and understood, it is colloquially said that no information was received and no information was transmitted.

Some individuals equate information with meaning [Mil87]. Hearing a statement isn't enough to make an event an informative act; its meaning must be perceived to make the statement informative. Arguing against this approach, Bar Hillel points out that “it is psychologically almost impossible not to make the shift from the one sense of information, ...i.e. information = signal sequence, to the other sense, information = what is expressed by the signal sequences” [BH55]. As Stonier reminds us, “we must not confuse the detection and/or interpretation of information with information itself” [Sto90]. For many who have worked with quantitative models of information in the engineering disciplines, this concern with meaning lies outside the scope of the traditional mathematical theory of information; communication engineers seldom concern themselves professionally with the meaning of messages.

In an approach similar to defining information as meaning, information is often understood in terms of knowledge that is transmitted to a sentient being. For example, Peters defines information as “knowledge with the human body taken out of it” [Pet88]. Similarly, information may be understood as “that which occurs within the mind upon the absorption of a message” [Pra82].

Information and its cousin entropy have long been studied as fundamental characteristics of physical systems and structures. Systems of molecules are often studied by considering an imaginary being, Maxwell's demon, a hypothetical gatekeeper between two sections of a closed system of molecules [LR90]. Assume that molecules move in a frictionless way. If the demon opens and closes the door at just the right times, can it essentially allow particles with higher levels of energy through to one side of the doorway and particles with lower levels of energy through to the other? Is this a perpetual motion machine, with the demon expending no energy, pumping energy to wherever it is wanted? This would violate the second law of thermodynamics. If the demon can't perform as described, why is it limited in its capabilities?

Attempts to solve this paradox have centered on the relationship between information and structure of the systems and the information needed and used by the demon in decision making, as well as the energy involved in making an observation and remembering the state of the system. The description of the positions of the molecules, the structure of the system, can be described using concepts and formulae consistent with some communication models, such as Shannon's model of communications.

The form or structure of systems is viewed by some as being equivalent to information. Thus, "information is what remains after one abstracts from the material aspects of physical reality" [Res89]. The information in a structure is an immaterial ghost that co-exists with the physical object about which it informs. The word "form" comes from the same etymological roots as "information." Form or state of nature may be reflected in the set of characteristics' values in the output of a producing process. The randomness associated with an information producing process about a system's form and structure may be understood as the inverse of information. Consider a pile of coins on a counter; whether they are "heads" or "tails" is information, while if we have no such information, their heads or tails orientation is random to us. Receiving information about the orientation of a coin may result in the removal of uncertainty, decreasing the ignorance or lack of information about the structure.

Information has long been understood as a concept appropriate for discussion in the humanities and social sciences [Rit91]. Electrical engineers began using the term to describe data transmission during the first half of the twentieth century. Instead of providing a definition of information, these engineers focused on measuring information, as they attempted to maximize information transmitted or received, or minimize noise, or both. The social science literature of the 1950s and 1960s used ideas about information measurement developed by Shannon and Weaver in the late 1940s crossing back from engineering to the liberal arts. Outside of electrical engineering, Shannon's formal ideas about information are used most profitably today in the computing and cognitive sciences.

All of these ideas about information serve to facilitate discourse for those describing discipline specific concepts, each used in solving a particular set of problems. Electrical engineers wish to study the capacity of pieces of hardware and the physical connections between them. Linguists wish to understand how information is transmitted by languages and the nature of what lies at the core of communication. Mathematicians and computer scientists wish to study the processes by which software transforms input into output and the fundamental characteristics of transforming processes.

Researchers in these disciplines want tools that manipulate the phenomena of their domains. These needs have produced ideas about information having varying degrees of overlap, as well as areas where they fail to intersect. Given the number of definitions or metaphors that have been proposed for information, how does one compare them? We propose a commonality within these definitions; this underlying commonality can be defined, studied, and measured.

We suggest here a general definition of information: *information is produced by all processes and it is the values of characteristics in the processes' output that are information.* This captures most concepts of information in individual disciplines. The number of possible values in the output and their relative frequencies of occurrence may be used in measuring the amount of information present.

3 Defining Information

A definition of information should capture the essential nature of the information phenomena in a precise description. While making explicit the similarities between information phenomena and other related concepts such as meaning, certainty, or knowledge, at the same time it should bring forward the differences between these concepts. For example, in what sense does information move through time or space? How is information similar to knowledge and how do they differ?

The nature of information may be brought forward through use of a metaphor or a model. For example, Shannon modeled information in communication processes as being transferred through a channel. Scholars and their students who have been trained in this model since the 1950s now consider this self-evident and an obvious base on which to build virtually any discussion about communication, although communication scholars have gone far beyond Shannon's initial model.

Information has been defined innumerable ways. As previously mentioned, Cookie Monster referred to it as "news or facts about something," while others have defined information as the meaning of a signal and some understand it as the signal itself. It may prove useful here to examine the requirements of a satisfactory definition of information. Dretske suggests that we must "preserve enough of our common understanding of information" if we are to maintain a link with the majority of the ideas about information present in our culture [Dre83]. Further, the definition should be as precise as possible, allowing non-informative phenomena to be clearly rejected. When the definition is applied in studies of information processing in humans, information may or may not have a person-specific aspect, and the quantity of information need not be the same for all sentient beings in a given situation; however, the amount of information should vary only as other factors, parameters of a definition, vary between the individuals. These external factors may be either personal or subjective and still allow for a precise definition and the useful measurement of information.

A good definition or theory of information both describes factually what occurs or what exists, as well as provides an explanation of events. In addition, it should bear some resemblance to the natural language notion of information but need not adhere to it when the natural language definition loses its generality and explanatory power. This happens when the common language definition of information, for example, becomes conflated with the notion of *useful* information, that is, information is understood to be in all cases useful. For those accepting this concept of information, if it isn't useful, it isn't information. Requiring that all information be useful limits the domain of discussions about information to cognitive processes that can "use" something; it excludes the information carried by a subatomic particle which is not sensed by a cognitive process. We try to avoid excluding information phenomena.

Information is often defined in terms of the human mind, although it is clear that very similar phenomena can be studied in lower level beings, such as communication and information transfer among ants [RR89]. Dretske notes that "it is common among cognitive scientists to regard information as a creation of the mind, as something we conscious agents assign to, or impose on, otherwise meaningless events. Information, like beauty, is in the mind of the beholder" [Dre83]. This human centered view of information limits information to that perceived or produced by the human mind; it rules out information not perceived by a mind, such as physical events on a microscopic, non-discernible level, such as is studied by physicists. It also means that the concept "information" is definitionally intertwined with the concept "mind." There is a commonality to what physicists and epistemologists study, and this is captured by our broad use of the term *information.* Our purpose here is to capture this commonality—we do not wish to imply that particular disciplines or group of scholars, such as those interested in information collection, information organization and information use in humans, should not place limitations on what they choose to study, or to narrow the concept of information further, according to their specialized interests. We believe that it is helpful if this field-specific definition is consistent with the general definition except for limitations imposed by a field. The author believes that the general definition is as useful as the field specific definitions in all cases.

A clear statement of *what is information* and *what is informative* can lead to a strong qualitative understanding of the fundamental nature of information. A *measure* of one or more of the characteristics of informational events is inherently quantitative. Such a measurement either determines a characteristic value of the phenomena of interest or compares it to a second phenomena. Two different measurements taken of the same type of characteristic, for example, what might be obtained when measuring the amount of information transmitted by two different computer modems, may be compared to determine whether the information capacity of one channel is greater than, equal to, or less than the other. These relations compare data whose values can be ordered by using these relations. This allows statements such as

x is more informative than y

or

x is equally informative to y

to be made, based on the measured informativeness.

Given a number of possible definitions for a concept, the better definition is usually the simpler and often broader definition. This rule, referred to as Occam's razor [BEHW87], suggests that the simplest explanation has the highest ratio of signal to noise; that is, the highest ratio of helpful information to distracting or "erroneous" information. Occam's law may be expanded to a range of environments, suggesting that the simplest law or definition that describes a phenomenon over the largest set of situations, the widest set of disciplines and theories, is the superior law. Consider a definition of heart disease that describes heart disease only in Caucasian males. A better definition would capture the essence of the disease in all *homo sapiens*. This more general definition might emphasize less those factors more likely to be experienced only by one subset of the population and is more likely to capture characteristics of *homo sapiens* in general. We do not want a definition of heart disease that is so broad as to be useless, such as a definition that covers all human processes and not just heart disease; a definition of heart disease is needed that describes the phenomena found in all species with various forms of hearts and circulatory systems, capturing the generalities that do exist in all heart disease. This develops theory as suggested by data. Obviously, medical doctors specializing in human cardiology need to understand the peculiarities of that species, while also understanding the broader processes taking place.

A similarly general but precise definition of information should exist at the core of information science. Imposing one's political and cultural interests on others in an academic discipline should be minimized—this imposition is clearly seen in those who define the field of information science as limited to the human use, organization, production, and retrieval of information, excluding other information phenomenon.

We note that individual and group biases and interests do affect what one knows and how one views the world. A strength of our definition of information is that it provides a place in a hierarchy of physical and mental processes for these biases to affect information in the form of biases acting as the input to mental processes. This acceptance of human biases as a cognitive reality is separate from our rejection of the imposition of individual biases on other scholars; we continue to believe that disciplines capture generalities and that information science should not reject any valid research on information and should attempt to embrace all forms of information.

4 A Process

Information is always informative about something, being a component of the output or result of the process. This "aboutness" or representation is the result of a process or function producing the representation of the input, which might, in turn, be the output of another function and represent its input, and so forth. Consider a common process such as cooking. Baking a cake begins with ingredients and a set of instructions, either written, spoken, or in the mind of the cook. Following the instructions, the cook transforms the ingredients into a sloppy mess which, after an appropriate amount of baking, results in a cake, if one is careful and perhaps lucky.

Examining the cake provides information about both the process and the original ingredients, assuming that the cake may be examined without the act of observation changing the cake. The choice of high quality ingredients or the addition of a special flavoring will affect the outcome, ideally in a beneficial way. Varying the process, such as the amount of time in the oven or the temperature at which the cake is cooked, also changes the final product, and an examination of the final product provides information about the process used as well as about the ingredients. Note that the information will seldom allow one to fully reconstruct the producing process and its input, and any prior knowledge about the process or its input will aid in the reconstruction. The cooking process changes one set of ingredients, one set of materials, into another set of materials: the cake. The change from one set of materials to the cake provides information about the process and the original materials and the baking process. We may speak of the cooking process as carrying information about the original materials.

A cook can't move backwards from a cooked cake to regenerate the original ingredients; baking is almost always an irreversible process. Processes may be totally reversible, allowing the process to move backwards from the final state to the initial state. Reversible processes are such that no information is unrecoverable (lost) during the operation of this process; thus, given the output, one can still move back to the input. A simple reversible process is one that increments the input by 1 and returns the incremented value. One can always take the output and, knowing the nature of the process, move backward to the unique input that produced the output. On the other hand, non-reversible processes may lose information as they operate. Given the output of a non-reversible process, one can't always tell which input produced the output. The square function that produces the number 4, for example, could take $+2$ or -2 as its *argument*, what it takes as input; knowing the result does not provide all the information needed to determine the input to the function. Information about the sign of the original number is lost when squaring occurs, making the reversal of the process not possible for all cases. Similarly, if the reader is told that the sum of two numbers is 7, it is impossible to determine whether the two initial numbers were 6 and 1, 4 and 3, or some other combination. One can imagine a reversible variant of this function that produces the sum and one of the original numbers. One can always move from these two outputs back to the original values.

Other types of processes produce information. An interesting phenomena is found at the quantum level in physics. Consider two particles that are produced from a single process such that they are moving in opposite directions. Many pairs of particles produced from a single creating process will each have a characteristic which does not "take on" a value (for either particle) until this value is observed or measured by instrumentation. Because these particles (will) have opposite characteristic values, measuring the value of one of the particles causes or forces the other particle, no matter at what distance, to take on the opposite value for the characteristic. A measuring process here makes information appear or become available—we try to avoid saying that information was "created" by the measuring process. The measuring process takes the particles that are valueless in regards to the characteristic as input and produces particles that have values.

All processes produce information: making cakes and measuring characteristics of sub-atomic particles, physical processes and processes commonly understood as non-physical, describable and indescribable processes. An understanding of the information produced by processes requires some understanding of the nature of a process. Processes may be complex, or they may be simple and easily described and studied. All produce information about the input and the process. The author believes that all processes can be described, given enough time and resources. However, even if some processes can not be described it is still useful to recognize the output of the process as "about" the process itself and the input. Furthermore, the notion of information as the values in the output of a process is helpful in understanding information phenomenon.

Processes consistent with assumptions defined by mathematicians may be defined as mathematical functions, such as those obtained by pressing mathematical operator keys on a calculator. These functions take one or more *arguments* as input and *return* a single value. Each input will produce the same given output each time a deterministic function is used, acting mechanically, always giving the same output from a given input. The process of addition, being deterministic and given common mathematical assumptions, will always produce 5 from inputs 2 and 3. Consider the *increment* function, which returns the value one more than the amount assigned to the argument. This assignment is referred to as the value *instantiated* or temporarily assigned to the argument. The increment function may be formally defined as $f(x) = x + 1$. Given a valued assigned to the variable x in parentheses, the function will have as its returned value the value to the right hand side of the equal sign, that is, $x + 1$. For this function, $f(2)$ would return the value (or have the value) 3.

Other functions may have probabilistic characteristics and may be able to emulate random processes. The values returned vary depending on occurrences independent of the input. A coin toss might be emulated by a probabilistic function $f(heads, 1/2)$ which returns the value *heads* approximately one half of the time.

Processes also may be defined as algorithms, sets of rules to be followed which produce an output, often in a certain order. A function may be thought of as a computer program or mechanical device that takes the characteristics of its input and produces output with its own characteristics. Every process may be defined functionally and every process may be defined as one or more functions. This interrelationship between functions, algorithms, and processes is governed by Church's Thesis, which formally describes the way in which several different descriptive languages or paradigms (e.g., processes, functions, agents, and algorithms) are capable of describing the same processes [KMA82]. For this reason, we can use the terms function, process, and algorithm interchangeably in many circumstances.

Processes always produce an effect – some change in the world – and thus can communicate information about the process and the input. Information occurs when the process produces something. Information is

> the value currently attached or instantiated to a characteristic or variable

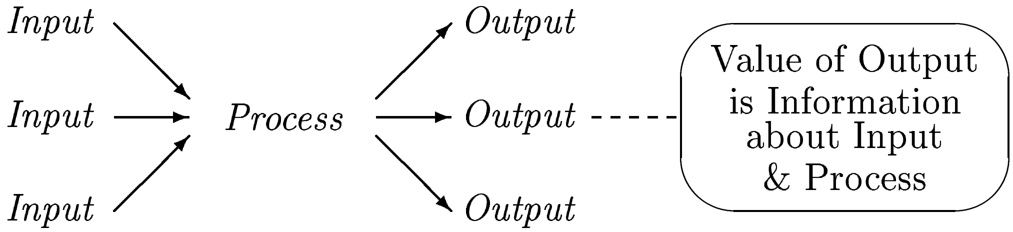

Figure 1: The value of the output of a process is informative about the process and its input.

> returned by a function, $f(x)$, or a process. The value returned by a function is *informative about* the function's argument x, or about the function $f()$, or about both.

This is graphically shown in Figure 1. We use x in our definition to represent either a single variable or set of variables. For the sake of linguistic simplicity, we treat x as a single variable below. The information is fully *contained* in $f(x)$, e.g., the received signal, but may not be fully contained in x, what is sent; the process may affect the output. A *message* is the information present at a point, with the term "message" being shorthand for this information. The received message is the value of the set of characteristics of $y = f(x)$ and not necessarily of x, the transmitted message, and $f(x)$ is informative *about* x and *about* f. The input x to the function $f(x)$ is informative about the process that produced it, that is, some other function. For example, a tree falling in a forest produces information in that the process produces an output: pressurized air waves that are perceived as noise by those with unimpaired hearing, the noise being informative about the falling of the tree. The tree itself is informative about the growth process, the original seed, and the nutrients in the soil, among other factors.

The process and its input cause the information to exist in the output, ignoring for purposes here the claimed ability of some to forsee the future. As with any causal phenomena, the cause must temporally precede the result. The existence of information thus always comes after the process that produced it has occurred.

Two factors affect the information in $y = f(x)$: the processing function itself and the initial values of variables such as x. These factors are combined by the process. The input, when processed, produce an output that is informative about both the input and the process. A function transforms or maps the input into an output with each input being "mapped" into a particular output.

The mapping f of a value is from one domain X into another domain Y. Each domain represents a set of possible values, with X being the set of possible values upon which the function operates and Y the set of possible values that can be produced by the function. For example, a small domain of male names may be represented as

$$X = \{abe, \quad bob, \quad charlie, \quad don\}.$$

A function f capitalizing the first letter in a name and mapping from X onto Y might have

$$Y = \{Abe, \quad Bob, \quad Charlie, \quad Don\}.$$

System information is the set of values produced by all possible inputs to a function, $Y = f(X)$, over the domain of X. X is the domain of possible messages that might be transmitted and Y the domain of possible received messages, with the values in Y being the information about the process and set X. When describing the actions of a function, one may refer to the operation on the domain X as $f(X)$, or one may refer to the action being applied to a specific x as $f(x)$. For example, a square root function may be defined as operating on any positive integer, while it may also be applied to specific number, e.g. 16. The particular value x is *bound* to X, taking a specific value. That is, the characteristic X has in a particular instance, the value x. The mapping is for all possible values in the domain X onto $Y = f(X)$.

Each function physically implements a *channel*, a causal mechanism that converts or transmits values from the initial domain X to domain Y. The function itself (as opposed to the value of the function) represents the information transmission process, while the value of the function provides information about the function and its input. System information is based on the relationships between the characteristics' values in the output. The quantity of information is measured by counting the relationships or a surrogate for the relationships. As these relations are produced by functions or processes, and only by functions, we may claim that information is contained in and only in the values returned by a function.

What is not information? Given our definition, information is not the process itself. The input to the process is not information about the process, although it clearly may be information about *another* process. The output is also not information by itself—the values in the output are information only in the sense that they are information *about* the process and the input, that is, information in the context of the process and its input.

5 A Hierarchical Model of Information Transmission

One strength of the function-output approach to information is its ability to describe information on a number of different levels: the electronic pulse in a wire representing a bit, sound waves representing linguistic phonemes, or thoughts being represented through written language or printed image. Consider *receiving* a speech sound x

generated by a speaker. We represent this action by

$$phoneme(x).$$

The values returned by the *phoneme* function are the information produced by the hearer.

We will first examine inverse functions. The inverse of function $f(x)$ is denoted as $f^{-1}(x) = \frac{1}{f(x)}$. Thus, $f^{-1}(f(x)) = x$ for all x and for all f that are reversible deterministic functions. For positive numbers, the mathematical square function has as its inverse the square root function. If f represents here the square function, $f(4) = 16$ and $f^{-1}(16) = 4$. Thus, $f^{-1}(f(4)) = 4$. For notational consistency below, we assume that a function $f(x)$ acts to decode the encoded message x. The inverse function $f^{-1}(x)$ acts to encode x.

By combining hierarchically arranged functions together, with one "feeding" the other in a hierarchical manner, one can capture common principles when describing complex information systems. These functions now may be "stacked" so that one function provides input to another in order of increasing complexity. Given that $f^{-1}(f(x)) = f(f^{-1}(x)) = x$ and $g^{-1}(g(x)) = g(g^{-1}(x)) = x$ for all x, and assuming that f and g are reversible, it will always be the case that

$$f(g(g^{-1}(f^{-1}(x)))) = x.$$

Note that the assumption of reversibility is an assumption that no "errors" or "loss" occurs in processing; in most systems, some loss occurs. These losses are often acceptable in situations where processes are not very sensitive to variances in their inputs. For example, slight changes in light intensity or in color usually go unnoticed in human vision, and slight variations in the type before the reader will not cause the content to be lost.

Through the combination of functions and their inverses, a variety of processes may be modeled, including high level communication behaviors such as how language is sent and received through a sonic medium. The information in the language is transmitted from sender to receiver through the encoding of the thought in a sonic form. We thus have two forms of representation: phrase and sound. This process may be represented by

$$phrase(phoneme(phoneme^{-1}(phrase^{-1}(x))))$$

where $phrase(x)$ is the phrase received by the hearer. If I desire to transmit a word to you, it is assigned to x and the $phrase^{-1}$ function encodes the word, producing the input to the $phoneme^{-1}$ function. This function places the coded sound into the atmosphere, where it is picked up and decoded by the *phoneme* function, which decodes the message into a form acceptable to the *phrase* function, which decodes its input, producing the original word if there was no interference with the process.

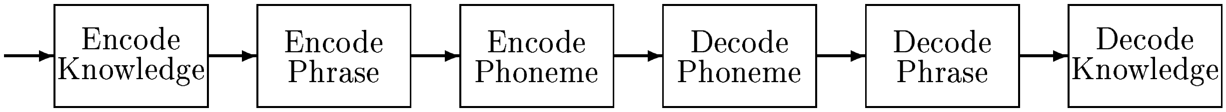

Figure 2: Encoding and decoding processes in communication.

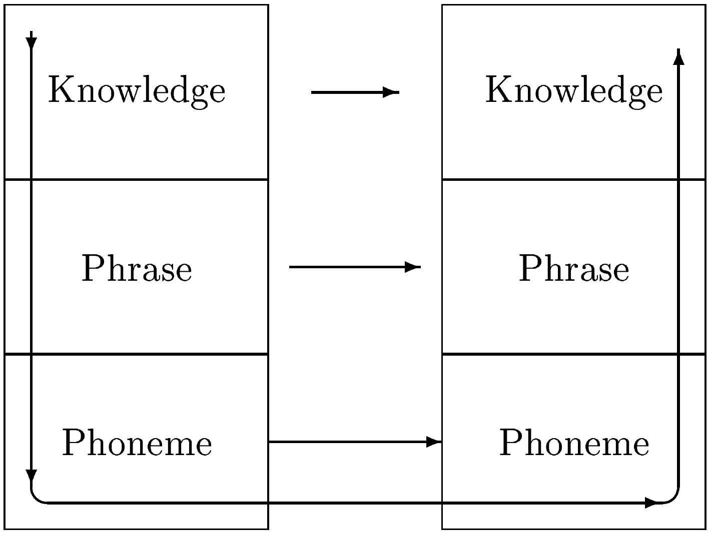

Figure 3: Hierarchical model of human communications. The “U” shaped arrow represents the passage of something being transmitted, being encoded and decoded as in Figure 2.

When one desires to add additional functionality to this hierarchical model, it becomes necessary to add both encoding and decoding functions for an added layer, e.g. $knowledge^{-1}$ and $knowledge(x)$. We may add a "knowledge" layer as follows:

$$knowledge\left(phrase\left(phoneme\left(phoneme^{-1}\left(phrase^{-1}\left(knowledge^{-1}(x)\right)\right)\right)\right)\right).$$

This assumes functions with lossless operation, and is illustrated in Figure 2.

Each function in Figure 3 can be thought of as a black box that accepts communications from above (on the left side) and "processes" the input. The output (from the bottom of these devices) indirectly feeds into the inverse of the function (flowing upwards on the same level on the right side) but goes directly into another function below it (on the left.) Lower level functions are necessary but not sufficient if communication is to occur; additional functions must be added until the function at the bottom can provide a physical linkage between the bottom layer on one hierarchy and the corresponding functional layer on another hierarchy.

One of the advantages of this form of model is that it allows for personal processes to operate differently from one individual to another, as long as they are largely reversible within the individual. This allows one to incorporate personal factors, cultural biases, etc. into higher level cognitive processes near the top of the hierarchies.

Information is always and only transmitted through a series of physical processes. In some cases, it may be helpful although not necessary to view this set of processes in a hierarchical manner because of the inherent hierarchical nature of communication and representation systems. In representational systems, one representation may represent another representation, such as when these printed words represent English language terms representing some ideas in the author's brain. The movement of a neural signal from a sensory organ to various parts of the brain is evidence of both a simple connection of processes and a hierarchical view of communication.

A given level has below (or above) it one or more hierarchies, referred to here as "legs," stretching out and touching the ends of other legs. For example, a given human mind may communicate through a number of different physical processes, such as through speech, gesture, or the written word. The same knowledge process can be working in the brain, but the transmitted message is traveling to its destination through one or more of several different physical methods. More complex networks of information transmission structures, such as those used in information diffusion [Cha86, LB75, Wil92] also may be modeled using "legs." These networks may be studied by examining the operation and transmission provided by "legs" taken as a whole, or by examining how the transmitting characteristics of each individual process or type of process within a leg.

A communication hierarchy consists of a number of layers. Our hierarchy is modeled after the telecommunications hierarchies that have been proposed by standards organizations (e.g. the ISO) and companies (e.g. IBM and SNA). What should constitute a layer, and what set of processes should be agglomerated into a single layer or process? Each layer is defined by the interface with the layer above or below as well as by the process producing the layers above, or below, or producing itself. Yet, any process can be continually broken down until the smallest possible physical process is reached. For the study of macroscopic environments, the arbitrary choice of the grouping of layers is required. They may be most beneficially selected based upon naturalness considerations, processes and layers being described so that the processes are easily understood.

Treating a specific range of possible levels as a single unit is consistent with information being viewed as a "thing" [Buc91]. Lumping together all levels below a certain point for several hierarchies results in a *channel* that can be treated as a unit. This channel is sometimes referred to as "information as thing."

The defining limits of a function are somewhat arbitrary and most hierarchies can usually be decomposed further. At the bottom of each hierarchy is a layer that contains the physical mechanism that allows communication to occur. We refer to this bottom layer as the "physical layer," no matter what size the function. The physical layer, as with all the other layers, can be defined such that it is very large, so that one layer performs a particular human sized task, or it may be very small, with several different smaller layers performing a particular task.

The value of characteristics of one level may be passed on to other levels or the values may be dropped. Loss of a value may be irrecoverable; the information will be permanently lost if the characteristic is independently valued and can't be inferred from other characteristics. For example, a lost value for the author's gender variable can be easily recovered from the author's name or knowledge about the presence of a beard, while the loss of the value for the author's first name would be a permanent loss, not being easily deduced from other characteristics, such as gender or the presence of a beard.

Most definitions and measures of information address *the transmission of these characteristics' values from one level to the corresponding level at the destination.* The use of the hierarchical model allows one to focus on the level in the hierarchy that is of greatest interest, rather than getting into a debate about whether information is of one nature or another, whether it is located at one level in the hierarchy or another. Information is produced and may be studied at *your* level of interest and the processes at your level *are* worthy of discussion, as are the information producing processes at other levels in the hierarchy.

6 Representation

A representation is the relationship existing between a set of characteristics' values in the output of one or more functions and the set of characteristics' values that were the function's input, with one or more of the characteristics' values in the input set being correlated with values in the output set. If there are no related values, we can say that no representation exists. Obviously, the nature of the set of characteristics present at a function's input and output is crucial to whether representation occurs or not. We can look at representation as the process through which information is transmitted to a neighboring level in the hierarchy or, if it is the bottom, physical layer, through the connecting media from one physical layer to another. The relationships between representations and names are examined in [Cof91, Cum91, Cum96, Gil92, Kri80, Moo93, Ros94].

Consider two layers in an information hierarchy: one containing a representation and the other what is being represented (a representandum). An apple might have a set of characteristics including *edible*, *red*, *stemmed*, and *round*. It might also have a worm hole on the backside, away from the viewer. A pictorial representation of the apple, let us say a black and white image, might show that it is round and has a stem. This representation isn't *perfect*; it does not contain all the characteristic values of the object, such as that there is a worm hole or that the apple is red.

Imagine that I am a very sophisticated biochemist and botanist and can make an apple representation in the laboratory that appears to be identical to the original; it crunches, oozes apple juice, tastes, and generally appears just like the original. If it is like the original in all characteristics being evaluated, then it is a perfect representation. The set of characteristics chosen obviously determines the degree of accuracy of a representation. A botanist might note after careful examination that the apple seeds in the representation are poor copies and consider the representation to be far from perfect, while a child with less sophistication and using a smaller set of characteristics, might consider the representation to be perfect. The choice of characteristics incorporated into the representation and observation has a great deal to do with its perceived quality.

A representation is usually thought of as being somehow less than the original. This loss (of information) often occurs as one moves from one process to another within the hierarchy. Each stage receiving information from a neighboring level has the potential to lose some of the information. If the representation does not encode all the characteristics of the neighboring level, information about the omitted characteristics is not present. This neighboring representation is *imperfect* and the process *lossy* when these characteristics of the neighboring layer are not present.

If a tree falls in a forest and nobody hears it, is information transmitted? The hierarchical approach to information suggests that the answer is "yes." While no human hears the crash of the branches hitting the ground, there have been changes brought about, although not in a human's ear or mind. A series of functional relationships transmitted the sound through physical processes, each process providing output informative about the process and about its input. None of these processes eventually resulted in a process feeding the representation of the sound to a human's mind. Information was transmitted, but none was received by a human.

In our conceptual framework, information doesn't require the presence of a human, and information science should not be viewed as a discipline with humans as its only focus. The focus of information science is information, with the discipline containing individual scholars interested primarily in information used by humans, as well as, for example, those interested in information as measured and discussed by physicists. The author believes that a diversity of interests and views about information are necessary in information science, and rejects the notion that it is acceptable to define the discipline as limited to the study of information created, organized, retrieved, and used by humans alone, excluding other ideas about information. While studying humans is obviously useful, humans are not the information universe; information predates humans and long after there are no more humans there will continue to be information.

7 The Beginnings of "Information Theory"

The hierarchical model of information may be applied to the description of any domain in which information appears, such as that used by Claude Shannon in the 1940s when he developed what is now referred to as "information theory" to study communication systems. His notions concerning information and measures of information are special cases of our functional definition of informativeness and information. Information theory is often considered to have begun with work by Harry Nyquist [Nyq24]. While new knowledge is built by individuals standing on the shoulders of those who performed earlier research, people such as Nyquist can be seen as being extraordinarily creative for putting together previous work to produce a new and unique model.

Writing in the *Bell System Technical Journal*, Nyquist suggested that two factors determine the "maximum speed of transmission of intelligence." Each telephone cable is implicitly considered to have a limit imposed on it such that there is a finite, maximum speed for transmitting "intelligence." This limit was widely understood by practicing electrical engineers of the era to be related to such factors as power, noise, and the frequency of the intelligent signal. Accepting such a limit as a given, Nyquist was able to work backwards towards the study of what was transmitted. He began referring to what was transmitted as "information."

The two fundamental factors governing the maximum speed of data transmission are the shape of a signal and the choice of code used to represent the intelligence. Responding to the earlier work of Squier and others, Nyquist argues that telegraph signals are most efficiently transmitted when the intelligence carrying waves are rectangular. Given a particular "code," use of square waves allows for intelligence to be transmitted faster than with sine waves in many practical environments.

Once the proper wave form is selected, a different problem arises: how should "intelligence" be represented? Telegraphers had long used Morse code and its variants to transmit text messages across distances. Each character was represented by a set of short or long electronic signals, the familiar dots and dashes. The letter *C*, for example, is represented in modern Morse code by a *dash dot dash dot* sequence.

Experienced telegraphers listen to messages at speeds far exceeding the ability of humans to consciously translate each individual dash or dot into a "thought representation" of the symbol; instead, Morse code is heard as a rhythm, with the rhythm for letters and common words being learned through long periods of listening.

Working backwards from the maximum telegraph speed, Nyquist considered the characteristics of an "ideal" code. Morse code is adequate for many applications, but an "adequate code' is far from being the best or optimal code available. Suggesting that the speed of intelligence transmission is proportional to the logarithm of the number of symbols which need to be represented, Nyquist was able to measure the amount of intelligence that can be transmitted using an ideal code. This is one step away from stating that there is a given amount of intelligence *in* a representation.

Four years later, another engineer, R.V.L. Hartley [Har28] expanded on ideas about *information.* Publishing in the same journal as Nyquist, the *Bell System Technical Journal*, and yet not citing Nyquist (or any one else, for that matter), Hartley developed the concept of information based on "physical as contrasted with psychological considerations" for use in studying electronic communications.

In the first section of his paper, titled *The Measurement of Information*, he noted that "information is a very elastic term." In fact, Hartley never adequately defines this core concept. Instead, he addresses the "precision of ... information" and the "amount of information." Information exists in the transmission of symbols, with symbols having "certain meanings to the parties communicating." When someone receives information, each received symbol allows the recipient to "eliminate possibilities," excluding other possible symbols and their associated meanings. "The precision of information depends upon what other symbol sequences might have been chosen;" the measure of these other sequences provides an indication of the amount of information transmitted. Nyquist then suggests that we take "as our practical measure of information the logarithm of the number of possible symbol sequences." Thus, if we received 4 different symbols occurring with equal frequency, this would represent 2 *bits* of information.

It is likely that Hartley was aware of the earlier work of Nyquist and that he assumed implicitly, as Nyquist did explicitly, that all symbol sequences were of the same length or size. The formula Hartley uses is consistent with this assumption, but serves only as an approximation of the information amount if the symbols are of different lengths [Bel73]. Symbols need not be equi-probable for Hartley's formula to be correct if symbols are of equal length. It is probable that Hartley did not make a statement concerning the probability of symbol sequences because of his (implicit) assumption of equal length symbols.

Hartley was aware of a relationship between the amount of energy in an information system and the amount of information that could be transmitted. Applying energy to an information transmitting system increases the ease with which the recipient receives or hears the transmitted signal. Energy serves as a component of the transmission process. Increasing the signal to noise ratio increases the proba-

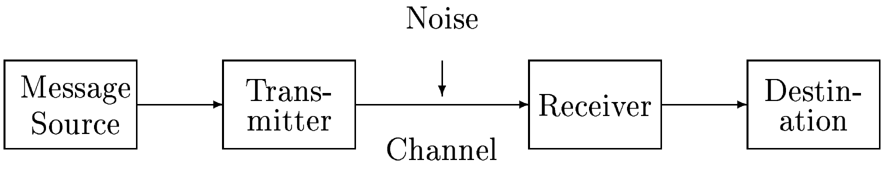

Figure 4: Shannon's channel model.

bility that the information will be received correctly. Information itself isn't energy carrying; it is energy that carries information.

During World War II, Claude Shannon developed a model of the communication process using the earlier work of Nyquist and Hartley. Published in 1947, *The Mathematical Theory of Communication* became the founding document for much of the future work in information theory. Given a number of desired properties for an information measure, the Shannon and Hartley measures of information *and only these* measures [AFN74] have properties desirable in an information measure. The importance of this work soon became apparent to scholars in a range of disciplines, resulting in its use (and abuse) from the middle of the twentieth century to the present. Shannon does not provide a definition; he is merely providing a model and the capability to measure information.

Shannon's work was intended to provide exactly what the title indicated: a theory of communication, useful in understanding telecommunications systems. In a private conversation in 1961, Shannon indicated that applications of his work to areas outside of communication theory were "suspect" [Rit86].

Shannon thought that

> the fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages [SW49].

Using this engineering perspective, the communication process may be understood as a *source* communicating to a *destination.* The source provides its message to a *transmitter* through a perfect connection. The transmitter communicates through a channel to the *receiver*, which receives the message and gives it in a lossless manner to the destination.

One of the key additions that Shannon made to the earlier work of Nyquist and Hartley was the formal integration of noise into the communication model. Noise is introduced into the channel between the transmitter and the receiver and acts to changes messages so that what is received differs from what is transmitted.

Sources may be discrete or non-discrete. A discrete source generates "the message, symbol by symbol. It will choose successive symbols according to certain probabilities depending, in general, on preceding choices as well as the particular symbols in question" [SW49]. Coding takes place at the transmitter. The source of the message does not transmit the message; the coded form of the message is what leaves the transmitting process and moves to the receiving process. The representation of the original message moves to the next process that transforms it, with the process continuing.

Between the source and the channel, the data being transmitted must be encoded, that is, it is represented in some form that can be transmitted by the medium supporting the channel. Transmitting data inherently requires that a change of medium take place, as the information moves from the source to the transmitter to the channel. When a signal moves from one medium to another, it must be physically represented somewhat differently, making an encoder necessary.

Given a source producing symbols at a rate consistent with a set of probabilities governing their frequency of occurrence, Shannon asks "how much information is 'produced' by such a process, or better, at what rate information is produced?" For Shannon, the amount of self-information that is contained in or associated with a message being transmitted, when the probability of its transmission is p, is the logarithm of the inverse of the probability, or $I = \log 1/p$ [Los90, TS95].

> The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used the resulting units may be called binary digits, or more briefly *bits*, a word suggested by J. W. Tukey. A device with two stable positions . . . can store one bit of information. N such devices can store N bits, since the total number of possible states is s^N and $\log_2 2^N = N$ [SW49].

The amount of information in the output of a process is proportional to the number of different values that the function might return. Given n different output values, the amount of information (I) may be computed as $I = \log_2 n$. The amount of information in the output of a process is related to the amount of information that is available about the input to the process combined with the information provided by the process itself. It is not just the amount of information about the input, although if the process always reproduces the input exactly at the output, there would be no difference in the amount of information present at the input to the process and at the output of the process. The information that is input to the function has measurable information, in its capacity as being the output of some other process, about which it provides information, the amount being measurable in terms of this earlier process.

The model for information transmission proposed by Shannon has been heavily abused by scholars who have applied the theory in domains distant from the electrical communication environment in which it was developed. By this, we mean that it has been frequently used to characterize situations that do not meet the assumptions and constraints of the model as proposed by Shannon. Ordering food at a restaurant might be modeled as a channel based process. The thoughts concerning food preference might be seen as the source, the vocalized order comes from the transmitting mouth, the waiter's ear is the receiver, and the chef is the destination. There is a symmetrical nature to the Shannon model that is missing from this example but, nevertheless, using the Shannon model may help an individual studying restaurant operations to be able to elucidate aspects of the operation that they hadn't considered before. For example, use of this model may suggest that noise effecting the channel might be examined. Using care in the choice of codes (names for food) might help decrease the error rate in recording customer orders.

A channel requires spatial or temporal distance between the sender and the receiver. Energy is necessary to transmit the message from the sender to the receiver. For Shannon, a channel is defined by a set of conditional probabilities that a certain message is received given what was transmitted. In cases where there is no noise, the conditional probability that a message is received given what was transmitted is simply the unconditional probability that the message is received. In noisy environments, what is transmitted is not always what is received.

The hierarchical model is, in some senses, a generalization of Shannon's model of a communication system. Both allow "information" to be encoded, "transmitted," and then decoded. Both provide a channel through which values may be passed. They differ in several respects. Shannon describes a communication system, while the more general hierarchical model can encompass communication, observation, and, in fact, any process that can produce a change in the universe. A communication channel is treated here as a process taking the input to a function in one hierarchy and producing output in another, providing information similar to what is communicated by a Shannonesque channel. Observation represents the presentation of information at a level in the hierarchy from a level in a second hierarchy, with the second level and hierarchy often being different from the first. The hierarchical model is broader than Shannon's model while retaining the ability to describe any particular communication system that Shannon's model can describe.

The communication model popularized by Shannon and Weaver may be understood in functional terms. Each channel may be understood as one function processing the input to produce characteristics in the output which takes on values related to the input. Unlike Shannon's model, the hierarchical model provides a definition of information in the system, in addition to measuring the information. The hierarchical model also makes explicit the large number of individual processes that participate in a Shannon channel. This hierarchical model may also be applied to more abstract notions, including perception, observation, belief, and knowledge. Use of a general model of information such as this allows for scholars across disciplines to share ideas and use words with the same meaning to describe information phenomenon found across the academic spectrum.

8 From Perceiving to Knowing

For a person or process to describe or utilize a state of nature, perception must take place, and observation often may be said to take place. These actions are necessary if the information is to leave the system under observation and become known and used by other processes or individuals. Understanding the arrival and use of information in sentient beings requires that the nature of observation, knowledge, and related topics be understood and described if information is be understood, defined, and measured. We make some initial suggestions here as to ways of modeling these phenomena as the result of transforming and hierarchical processes, allowing them to be incorporated into our general conceptual framework for information.

A low level of information reception in humans is *perception*, the physical receipt of information about a state or situation occurring elsewhere. Perception, often understood to be a biological phenomena, may involve a substantial amount of processing by neural networks [Nor93]. Visual perceptions, for example, involve the sensing of light and dark images by rods in the back of the eye, with some animals detecting colors through reactions taking place within cones in their retinas. These sensors are linked through complex networks to portions of the brain.

Some higher level perceptual processes involve cognitive processing and require memory. For example, the detection of motion obviously requires that one remember the previous position of the object if one is to detect a change in position. Similarly, detecting the presence of a particular shape (e.g., a square) requires the ability to recognize a set of individual retinal points representing a square as being a square despite the variations in sizes for squares and the different orientations (e.g., rotation) in the visual field that a square may take.

An entity may be said to perceive another entity when a process or chain of processes receives "input," that is, something at the "causal" end of the process changes, and a representative output is physically produced by the perceiver. Information is available in the perceiver about the object of perception only when the perception is physically necessitated by the presence of a characteristic's values at the input and the information carrying or producing process.

The processes that take physical events to human sense organs is often unaffected by higher level cognitive processes as well as by other human biases. On the other hand, empirical evidence strongly supports the notion that human background and experience strongly determine the nature of what is *observed* by humans. The "biased" nature of information processing by humans is such that all processing, whether of observations of the outside world or "thought process" running from rumination to formal deductions are likely to be biased. The ability of human biases to manipulate what is perceived is one of the factors separating perception from observation.

An observation takes place when a sentient being directly receives the output of a process whose input is said to be "observed." Observation and perception require some degree of accuracy, with scholars disagreeing on the amount of error allowed, if any, for a valid perceptual or observational act to have taken place. Some suggest that the observation must take place "without interference" [Sha82], while others suggest that a "fitness condition" for observation is that identical inputs to the function should produce identical outputs and similar inputs should produce similar outputs or observations [Pal89]. Observations take place within the context of the observer's beliefs and states; an electronically knowledgeable individual will take away far more from viewing the insides of a desktop computer than will most adults or any child.

Observation is understood by some to be limited to "humanly-accessible information" which is eventually used "by a human being" [Sha82]. Perception functions as a lower level function, with observation requiring a higher level of cognition and reasoning on the part of the observer. Abstract concepts such as emotions are usually described as observed rather than perceived. For example, people often feel more comfortable saying they "observe hostility" rather than they "perceive" hostility.

Observing is like receiving the output of any other function [Sha82]; its primary difference is in the nature of the function itself, the high degree of reproduction of the input in the output, and its higher location in the hierarchy, usually including what is referred to as "reason." If an observation is viewed as the output of a single process linking the observed and the observer, the information provided to the observer concerns the state presented to the input of the process as well as about the observation transmitting process itself. This observation transmitting process may also be viewed as a series of processes stacked in a hierarchy, such as those described earlier. The output of each of these processes provides information about its input as well as about the process itself. The output from the terminating process provides information about its input as well as about the terminating process and provides some information about the processes that preceded it. Similar arguments can be made about direct and indirect perceptual processes.

8.1 Belief and Knowledge

Perceiving and observing by a sentient being (and in many non-sentient mechanisms) produce output having some relationship to the state of the world outside the observer. The characteristics of the output of the process serve as input to memory structures that store beliefs. A belief is an idea, or statement, that has one or more characteristics' values that match the values for representandums. Belief may thus be understood as a representation that is not necessarily fully justified and is not necessarily completely true, but must be true in part. A belief is an idea that is held based on some support. Thus, Swinburne has suggested that if a person believes proposition p then p must be more probable than $\neg p$ [Hel94]. Unfortunately, this implies that if there is a proposition, one holds a belief either in the proposition or in its negation. The holding of incorrect or weak beliefs becomes problematic, as does the imposition of a logical formulation on this sort of problem.

A statement of belief contains one or more characteristics' values matching in full or in part the values for representandums. The output of a process or set of processes provides a representation of the input to the processes. We can therefore describe a *percept* as the set of values in this output; it is essentially the information in the output of a process about the input. A *perceiving function*, $f()$, provides a percept, $f(x)$, about input x. Belief is thus transmitted through the hierarchy.

Knowledge has been frequently described as "justified true belief," a belief held by an individual that is both true and for which they have some justification. Thus, for a belief to be knowledge, it must be the case that the belief is, in fact, true, and the believer must have justification for the belief. A belief that is true but for which we have no evidence cannot be described as knowledge. If there are homunculi inside computers performing operations, those who have long believed in their presence cannot be said to have had knowledge of this, since their belief, while true, has never been justified (we assume.)

It had become common to describe knowledge as "justified true belief" when Gettier [Get63] wrote a brief article that raised a problem with this definition. As a result of Gettier's work, we can be certain that "knowledge is not, or is not merely, justified true belief" [Dre81]. There have been several responses to Gettier's argument against accepting knowledge as being only justified true belief. One possible approach is to add a condition requiring that the grounds for believing a proposition do not include any false beliefs [Pol86], to the requirements of justification, truth, and belief. However, this addition and several other modifications that have been proposed fail to avoid counterexamples in which "knowledge is lacking despite the believer's not inferring his belief from any false beliefs" [Pol86]. Other approaches to understanding knowledge have been proposed and supported, such as having a disposition to behave or a disposition to feel a certain way [Ack72, Hel94].

We accept here that knowledge is something like "justified true belief." A belief is an internally accepted statement, the result of an observation or an inferential or deductive product combining observed facts about the world with reasoning processes. To understand knowledge in a way consistent with a hierarchical notion of information, it becomes necessary to understand the notions of "truth" and "justification" in a manner consistent with the hierarchical context.

A statement may be understood as "true" if it exactly represents what it is describing. This is referred to as the "correspondence" theory of truth [All93, Joh92]. This applies not only to statements but to the representation and belief. The coherence theory of truth, on the other hand, suggests that truth is essentially derived from a system. A statement is true when it is consistent with a system of accepted statements. Truth may also be viewed as a representation that is learned and that will not be altered, even given additional experiences. William James thus defined truth as the vanishing point toward which we imagine that all our temporary truths will some day converge.

The justification of a belief is based on internal considerations concerning the qualities of the function producing the belief. A belief is "justified" if and only if the input to the function is *accurately* represented in the output. Consider a handheld calculator which accepts the keystrokes "2" "+" "2" "=" and then displays the digit "4." We note that the digit displayed is not of the same form as the input, e.g., a keystroke. Instead, an accurate function takes keystrokes and produces a displayed number. If the calculator is broken and produces the digit "3" given the above set of keystrokes, we clearly don't have knowledge that $2 + 2 = 4$. Consider a different case where the calculator is broken but the above set of keystrokes produces, through erroneous subprocesses, the digit "4" in the display. While the output is correct or "true" and may be interpreted as a belief, it is not justified—the function is not accurate in that it does not operate as the user intends or *understands* the calculator to operate.

Other models of knowledge have been proposed, such as the notion that knowledge is one's "image," what one subjectively believes to be true [Bou56]. This is close to what we have referred to as a belief, and choosing to call it "knowledge" appears to only confuse the issue. Yet, like the more conventional philosophical idea of knowledge, it can be understood as the values in the output of a process, actually, the hierarchical series of processes that range from low level atomic processes up to sophisticated intellectual processes.

Perception and observation can be understood as conveying information about the input to certain processes (for humans, sensory processes such as seeing, hearing, smelling, etc.) The output of such a process may be understood as a belief. Such a belief may constitute knowledge about the input when the process or set of processes producing the belief operate in a manner consistent with the understanding of the process. These definitions of knowledge and belief are broader than the common language notions of the terms and less human-centered, in the case of belief, making the concepts more objective and more easily studied. We note that knowledge is information that is both true and justified. These perceptual, observational, and processing functions take as input sensory data from the real world, as well as personal beliefs and cultural biases, when producing information bearing output. This conceptual framework for understanding information provides a mechanism for understanding both the cultural influence on information, as well as the most minute phenomena studied by physicists.

9 Errors, Misinformation, and Bad Data

When discussing knowledge, truth, and information, people often begin speaking of "misinformation" or "bad data." Scientists often speak of "bad data," produced by faulty measurement or poor observations. Misinformation [Fox83] often refers to information that is "false," that is, the information does not directly reflect the "true" state of the world. Consider a "lie" told by an individual or an organization. The person making the lie knows the truth and, instead of repeating it, chooses to produce a lie for some purpose. The lie is then information about the process that produced it. It is misinformation or "false information" only in the sense that we may not know the nature of the full hierarchy of processes, that is, the characteristics of the function that did eventually produce the lie. The exact nature of misinformation, etc., is subject to a wide range of interpretations (and misinterpretations). Dretske notes that "no structure can carry the information that s is F unless, in fact, s is F. False information, misinformation, and disinformation are not varieties of information" [Dre83].

Misinformation and related concepts may be defined consistent with the hierarchical model. When what is transmitted is not received as sent, i.e., $x \neq f^{-1}(f(x))$ for some x that is input to an $f()$, an *information loss* has occurred in the process. This can occur when the function f can be said to be a partially random or noisy process. When this happens, what is transmitted was not reproduced by the function; some of what was transmitted has been "lost."

When information has been lost in producing a particular output characteristic, the value taken on by the characteristic is determined, in part, by a random or error component. When there exists a non-null error component in determining a characteristic or variable's value, the "information" contained in the variable may be referred to as "misinformation." The value of a variable is information about the input; when the information is only partial and is tainted by error, it is better understood as misinformation. Essentially, this is information that is partly or wholly false.

Economists and scholars interested in decision theory [HR92, WY73, YK93] often refer to error free information as *perfect information* [AW81, BG54, Mar84, HM87, Ner81, Phl88, Sen93]. Perfect information about a source domain X exists when there is a one-to-one mapping in a noiseless environment for the source X onto the destination set Y [Los90]. Information may be said to be *incomplete* when the mapping from X is into Y, not onto Y.

We may define another form of misinformation as information that isn't justified. If one believes something for the wrong reasons, one may be said to be "misinformed." In these cases, there is a perception that something is wrong with the recipient of the information, and it is this faulty nature of the receiver that makes something "misinformation."

10 Conclusions

Information may be understood as the value attached or instantiated to a characteristic or variable returned by a function or produced by a process. We note that the value returned by a function is informative about the input to the process and about the process itself. Using the proposed hierarchical model of stacked processes one may model existing ideas about information, including the communication model proposed by Shannon, information as “thing” or information as “knowledge.” The information hierarchy provides a satisfactory link between physical processes and consistent ideas about information and higher level mental functions discussed by psychologists and philosophers. This allows information scientists and others to examine information in a uniform way across the breadth of information phenomena, providing a level of precision to some interdisciplinary discussions of information, and serving as a base to which additional limiting assumptions may be added within specific disciplines, such as the concept of “value” for economic studies of information.

Consider the application of our conceptual framework for studying information to a telephone conversation. An electrical engineer might choose to examine the electrical pulses and waveforms on the phone line by treating the telephone line as a process that produces output whose content can be studied. A linguist might wish to study the phonetic or syntactic processes and the information they produce. The student of cultural studies might note how specific cultural biases affect the quantity and quality of information produced in beliefs by specific observational processes. Each scholar can examine the same situation and the same set of data from a discipline specific perspective, yet using common terminology and measures to describe information.

Although information science is seen as a social science by some [HA95] and as a branch of physics by others [Bri56], a single broad discipline-independent definition of information may serve as the basis for a rigorous and inclusive view of information. Accepting such a model simplifies the discussion of several field dependent concepts, such as knowledge and belief. Use of this hierarchical model allows us to focus our studies on the information phenomenon in all its possible incarnations.

References

[Ack72] Robert J. Ackermann. *Belief and knowledge.* Anchor Books, Garden City, NY, 1972.

[AD75] J. Aczél and Z. Daróczy. *On Measures of Information and Their Characterizations.* Academic Press, New York, 1975.

[AFN74] J. Aczél, B. Forte, and C. T. Ng. Why the Shannon and Hartley entropies are ‘natural’. *Advances in Applied Probability*, 6(1):131–146, 1974.

[All93] Barry Allen. *Truth in Philosophy.* Harvard, Cambridge, Mass., 1993.

[Art73] Susan Artandi. Information concepts and their utility. *Journal of the American Society for Information Science*, 24(4):242–245, 1973.

[AW81] Niv Ahituv and Yair Wand. Information evaluation and decision makers’ objectives. *Interfaces*, 11(3):24–33, June 1981.

[BC92] Cristina Bicchieri and Maria L. D. Chiara, editors. *Knowledge, Belief, and Strategic Interaction.* Cambridge U. Press, Cambridge, 1992.

[BEHW87] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Occam's razor. *Information Processing Letters*, 24:377–380, 1987.

[Bel73] Jack Belzer. Information theory as a measure of content. *Journal of the American Society for Information Science*, 24(4):300–304, 1973.

[BG54] David Blackwell and M. A. Girshick. *Theory of Games and Statistical Decisions.* Wiley, New York, 1954.

[BH55] Yehoshua Bar-Hillel. An examination of information theory. *Philosophy of Science*, 22:86–105, 1955.

[Bou56] Kenneth E. Boulding. *The Image: Knowledge in Life and Society.* U. of Michigan Press, Ann Arbor, 1956.

[BR76] Nicholas J. Belkin and Stephen E. Robertson. Information science and the phenomenon of information. *Journal of the American Society for Information Science*, 27:197–204, July-August 1976.

[Bri56] Leon Brillouin. *Science and Information Theory.* Academic Press, New York, 1956.

[Buc91] Michael K. Buckland. Information as thing. *Journal of the American Society for Information Science*, 42(5):351–360, 1991.

[Cha86] Elfreda A. Chatman. Diffusion theory: A review and test of a conceptual model in information diffusion. *Journal of the American Society for Information Science*, 37:377–386, November 1986.

[Cof91] J. Alberto Coffa. *The Semantic Tradition from Kant to Carnap to the Vienna Station.* Cambridge U. Press, 1991.

[Cum91] Robert Cummins. *Meaning and Mental Representation.* MIT Press, Cambridge, Mass., 1991.

[Cum96] Robert Cummins. *Representations, Targets, and Attitudes.* MIT Press, Cambridge, Mass., 1996.

[Dre81] Fred I. Dretske. *Knowledge & the Flow of Information.* MIT Press, Cambridge, Mass., 1981.

[Dre83] Fred I. Dretske. Précis of 'Knowledge & the flow of information'. *Behavioral and Brain Sciences*, 6:55–90, 1983.

[Far80] Jason Farradane. Knowledge, information, and information science. *Journal of Information Science*, 2:75–80, 1980.

[Fox83] Christopher J. Fox. *Information and Misinformation: An Investigation of the Notions of Information, Misinformation, Informing, and Misinforming.* Greenwood, Westport, Conn, 1983.

[Get63] Edmund L. Gettier. Is justified true belief knowledge? *Analysis*, 23:121–123, 1963.

[Gil92] Grant Gillett. *Representation, Meaning, and Thought.* Clarendon Press, Oxford, 1992.

[HA95] Birger Hjorland and Hanne Albrechtsen. Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science*, 46(6):400–425, 1995.

[Har28] R. V. L. Hartley. Transmission of information. *Bell System Technical Journal*, 7:535–563, 1928.

[Har84] Glynn Harmon. The measurement of information. *Information Processing and Management*, 20(1–2):193–198, 1984.

[Hel94] Paul Helm. *Belief Policies.* Cambridge University Press, Cambridge, 1994.

[HM87] Karl Heinz Hoffman and Natalia Meshkov. Problem from empirical economics: II. Determining uncertainties arising from incomplete data using information theory. *Resources and Energy*, 9(4):379–395, 1987.

[HR92] Jack Hirshleifer and John G. Riley. *The Analytics of Uncertainty and Information.* Cambridge, 1992.

[Joh92] Lawrence E. Johnson. *Focusing on Truth.* Routledge, London, 1992.

[KMA82] A. J. Kfoury, Robert N. Moll, and Michael A. Arbib. *A Programming Approach to Computability.* Springer-Verlag, New York, 1982.

[Kri80] Saul A. Kripke. *Naming and Necessity.* Harvard U. Press, Cambridge, Mass., 1980.

[LB75] Nan Lin and Ronald S. Burt. Differential effects of information channels in the process of innovation diffusion. *Social Forces*, 54(1):256–274, September 1975.

[Lev77] Marilyn M. Levine. The informative act and its aftermath: Toward a predictive science of information. *Journal of the American Society for Information Science*, 28:101–106, March 1977.

[Los90] Robert M. Losee. *The Science of Information: Measurement and Applications.* Academic Press, San Diego, 1990.

[Los95] Robert M. Losee. The development and migration of concepts from donor to borrower disciplines: Sublanguage term use in hard & soft sciences. In Michael E. D. Koenig and Abraham Bookstein, editors, *Proceedings of the Fifth International Conference on Scientometrics and Informetrics*, pages 265–274, June 1995.

[LR90] Harvey S. Leff and Andrew F. Rex. *Maxwell's Demon: Entropy, Information, Computing.* Princeton Univ. Press, Princeton, NJ, 1990.

[LW86] Bruno Latour and Steve Woolgar. *Laboratory Life: The Construction of Scientific Facts.* Princeton Univ. Press, Princeton, NJ, 1986.

[Mar84] Bertil Marksjo. Simple aggregation and disaggregation subject to minimal information loss and other criteria. *Regional Science and Urban Economics*, 14(3):465–478, 1984.

[Mil87] Gordon L. Miller. *Resonance, Information, and the Primacy of Process: Ancient Light on Modern Information and Communication Theory and Technology.* PhD thesis, Library and Information Studies, Rutgers, New Brunswick, N.J., May 1987.

[MM83] Fritz Machlup and Una Mansfield, editors. *The Study of Information: Interdisciplinary Messages.* Wiley Interscience, New York, 1983.

[Moo93] A. W. Moore. *Meaning and Reference.* Oxford U. Press, Oxford, 1993.

[Mul74] Michael Mulkay. Conceptual displacements and migration in science. *Science Studies*, 4:205–234, 1974.

[NCB72] Francis Narin, M. Carpeter, and N. C. Berlt. Interrelationships of scientific journals. *Journal of the American Society for Information Science*, 23(5):323–331, 1972.

[Ner81] Manfred Nermuth. *Information Structures in Economics.* Springer-Verlag, Berlin, 1981.

[Nor93] Kenneth H. Norwich. *Information, Sensation, and Perception.* Academic Press, San Diego, 1993.

[Nyq24] H. Nyquist. Certain factors affecting telegraph speed. *Bell System Technical Journal*, 3:324–346, 1924.

[Pal89] Bonnie T. Paller. Visual perception, observation systems, and empiricism. *Philosophical Studies*, 55:65–80, 1989.

[Pet88] John Durham Peters. Information: Notes toward a critical history. *Journal of Communication Inquiry*, 12(2):9–23, Summer 1988.

[Phl88] Louis Phlips. *The Economics of Imperfect Information.* Cambridge, Cambridge, U.K., 1988.

[Pol86] John L. Pollock. *Contemporary Theories of Knowledge.* Rowman & Littlefield, Totowa, N.J., 1986.

[Pra82] Alan D. Pratt. *The Information of the Image.* Ablex, Norwood, NJ, 1982.

[Res89] Howard L. Resnikoff. *The Illusion of Reality.* Springer-Verlag, New York, 1989.

[Rit86] David Ritchie. Shannon and Weaver: Unraveling the paradox of information. *Communication Research*, 13(2):278–298, April 1986.

[Rit91] David L. Ritchie. *Information.* Sage, Newbury Park, CA, 1991.

[Ros94] Jay F. Rosenberg. *Beyond Formalism: Naming and Necessity for Human Beings.* Temple U. Press, Philadelphia, 1994.

[RR89] Zh. I. Reznikova and B. Ya. Ryabko. Experimental proof of the use of numerals in the language of ants. *Problems of Information Transmission*, 24(4):334–338, April 1989.

[Rus37] Bertrand Russell. *The Principles of Mathematics.* Allen & Unwin, London, second edition, 1937.

[Sen93] Jati K. Sengupta. *Econometrics of Information and Efficiency.* Kluwer, Dordrecht, 1993.

[Sha82] Dudley Shapere. The concept of observation in science and philosophy. *Philosophy of Science*, 49:485–525, 1982.

[Sto90] Tom Stonier. *Information and the Internal Structure of the Universe.* Springer Verlag, Berlin, 1990.

[SW49] Claude E. Shannon and W. Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, Ill., 1949.

[TS95] Jean Tague-Sutcliffe. *Measuring Information: An Information Services Perspective.* Academic Press, San Diego, 1995.

[Wil92] Barbara Marie Wildemuth. An empirically grounded model of the adoption of intellectual technologies. *Journal of the American Society for Information Science*, 43:210–224, April 1992.

[WY73] B. J. Whittemore and M. C. Yovits. A generalized conceptual development for the analysis and flow of information. *Journal of the American Society for Information Science*, 24(3):221–231, May-June 1973.

[YK93] M. C. Yovits and R. M. Kleyle. The average decision maker and its properties utilizing the generalized information system model. *Journal of the American Society for Information Science*, 44(6):352–363, July 1993.