

## What is information?

As the Chinese say, 1001 words is worth more than a picture.  
--John McCarthy

### 1. Welcome to the Information Age

If we were told we were living in the Age of Analysis, we might wonder just which kind of analysis was being celebrated? Psychoanalysis? Chemical analysis? Statistical analysis? Conceptual analysis? That term has many uses. We live in the Information Age, say the pundits, and people seem to agree, but few notice that this term also has several equally distinct meanings. *Which* Information Age, the age of megabytes and bandwidth, OR the age of intellectual and scientific discovery, propaganda and *disinformation*, universal education and challenges to privacy? The two concepts of information are closely related, and are often inextricably intertwined in discussion, but they can be made distinct. Consider a few examples:

- (1) "The information is in the light." (J. J. Gibson, ecological psychologist)(1966)
- (2) We are *informavores* (George Miller, psychologist)
- (3) The brain is an information-processing organ.
- (4) The task of a nervous system is to extract information from the environment to use in modulating or guiding successful behavior.
- (5) We are drowning in information.
- (6) We no longer can control our personal information.
- (7) The task of the Central Intelligence Agency is to gather information about our enemies.
- (8) *Humint*, intelligence or information gathered by human agents in clandestine interaction with other human beings is much more important than the information obtainable by satellite surveillance, and other high-tech methods.

None of these uses of the term "information" have any *direct* connection to Claude Shannon's brilliant mathematical theory of information (Shannon, 1948, Shannon and Weaver, 1949), which underlies today's *digitization* of information by computers. This may be the *Digital Age* now that CDs, DVDs and cell phones have replaced LP phonograph records, analog radio, telephone and television transmission, but the *Information Age* was born when people began writing things down, drawing maps, and otherwise recording and transmitting valuable

information they couldn't keep in their heads with high fidelity. Or we could place the beginning of the Information Age earlier, when people began speaking, and passing on accumulated lore, history and mythology.<sup>1</sup> We may call these prior phenomena *semantic* information, since we identify the information of interest to us on a particular occasion by specifying what it is *about* (events, conditions, objects, people, spies, products, . . .). Other terms have been used, but "semantic information" is the dominant term. The information that Tom is tall is about Tom and his height, and the information that snow is white is about snow and its color. These are different items of semantic information (don't say "bits of information," because "bits" belong to the other, Shannon, sense of the term). Even before writing was widespread, people invented ways of improving their control of semantic information, using rhyme, rhythm, and often musical tone to anchor valued formulations in their memories, mnemonic crutches we still encounter today, in Every Good Boy Deserves Fudge (the notes on the lines of the treble clef in musical notation), HOMES (Great Lakes Huron, Ontario, Michigan, Erie and Superior) and My Very Educated Mother Just Served Us Nachos (it used to be: Nine Pies), for the order of the planets, starting at the Sun.

Shannon idealized and simplified the task of moving semantic information from point A to point B by breaking the task down into a Sender and a Receiver (two rational agents, note) with a channel between them and a pre-established or agreed-upon code, the alphabet or ensemble of permissible signals. The channel was susceptible to *noise* (which was anything that interfered with transmission, degrading the signal), and the task was to achieve reliable transmission that could overcome the noise. Some of the designs that accomplish this were already well understood when Shannon devised his theory, such as the Alpha Bravo Charlie system of alphabetical letters (the NATO Phonetic Alphabet) used in voice radio transmission that could minimize the confusion between the rhyming letters (in English) Bee, Cee, Dee, Eee, Gee, Pee, Tee, Vee, Zee, and so forth. Shannon showed how, by converting all codes, including words in ordinary language, to a binary code (with an alphabet containing only two symbols, 0 and 1), noise reduction could be improved indefinitely, and the costs (in terms of coding and decoding and slowing down the transmission speed) could be measured precisely, in *bits*, which is short for *binary digit*. Like the parlor game Twenty Questions, where only yes/no questions are permitted, all information transmission could be broken down into binary decisions, yes or no, 1 or 0, and the number of such decisions required to recover the message could be given a measure, in bits, of the amount of (Shannon) information in the message. "I'm thinking of a number between 0 and 8. What is it?" In the game of Twenty Questions, how many questions do you have ask to be sure of

---

<sup>1</sup> Or with some justice we could say the Information Age began over 530 million years ago, when eyesight evolved during the Cambrian Explosion, triggering an arms race of innovation in behavior and organs that could respond swiftly to the information gathered. Or we could insist that the Information Age began when life began; even the simplest reproducing cells survived thanks to parts that functioned by discriminating differences in themselves and their immediate surroundings.

answering this question? Not eight questions (“Is it 0, is it 1, is it 2. . . ?) but only three: (“Is it greater than 4? Is it greater than 6 [or 2, depending on the first answer]? Is it greater than 7?). So to represent the number 8 you need exactly 3 bits of information: yes, yes, yes = 111. A byte is eight bits, and a megabyte is eight million bits, so you can send a 2.5 megabyte jpeg picture file by playing Twenty Million Questions.

Shannon’s information theory is valuable because semantic information is so important to us that we want to be able to use it effectively, store it, move it, transform it, share it, hide it. Informational artifacts abound—telephones, books, maps, recipes—and information theory itself began as an artifact for studying important features of those artifacts. What began as an engineering discipline has subsequently proven useful to physicists, biologists and others not concerned with the properties of informational artifacts. We will touch lightly on some of these further applications of Shannon information, but our main quarry is semantic information.<sup>2</sup>

Shannon’s mnemonic tricks (and their descendants) are not just good for sending information from one Agent to another Agent, but for “sending” information from an Agent *now* to the same Agent in the future. Memory can be conceived as an information channel, just as subject to noise as any telephone line. In principle, *digitization* could be done with an alphabet of three or four or seventeen or a million distinct signals, or in other ways (as we shall see in a later chapter) but binary coding proves for many reasons to be superior for most purposes and is always available. *Anything* can be coded in a scheme of 0s and 1s, or 0s, 1s, and 2s, or . . . . but the binary code is physically simpler to implement (On/Off; High-voltage/Low-voltage; Left/Right) so it has pretty much gone to fixation in human technology, although there is still competition among secondary codes composed of binary code. (ASCII code for printable characters has been surpassed by UTF 8, which includes ASCII as a subset, and HTML, the code used for websites, has two different color codes in use: HEX code, and RGB triplets, for example.)

The task of translating or transducing semantic information into the binary bit-string format (the format directly measurable as Shannon information) is a mature technology now, and there are many varieties of *A-to-D converters* (ADCs— analog to digital converters) that take continuous or analog variation in some physical event (an acoustic wave hitting a microphone, light striking a pixel in a digital camera, temperature change, acceleration, humidity, pH, blood pressure, etc.) and transduce it into one string of bits or another. These devices are parallels to the sensitive cells that accomplish transduction on the outer input edges of the nervous systems: rods and cones in eyes, hair cells in ears, heat sensors, nociceptors for damage (pain), stretch sensors on muscles, and all manner of internal monitoring cells that feed the nervous system, including the autonomic nervous

---

<sup>2</sup> Colgate and Ziock (2010) provide a brief, useful summary of some of the history of definitions of information growing out of the work of Shannon and Weaver.

system. The conversion is not into bit strings but neuronal spike trains, voltage differences passing rather slowly—millions of times slower than bit strings moving in computers—from neuron to neuron. In 1943 (before there were any working digital computers!) the neuroscientist Warren McCulloch and the logician Walter Pitts proposed a way these neuronal signals might be operating. It seemed that when a spike train from one neuron arrived at another neuron, the effect on it was either excitatory (Yes!) or inhibitory (No!). If the receiving neuron had a *threshold* mechanism that could sum up the Yes votes and subtract the No votes and then trigger its own signal depending on the net result, it could compute a simple logical function (it could be an “AND-gate” or an “OR-gate” or a “NOT-gate”, to take the simplest cases). If a cell’s threshold could be raised or lowered by its history, then a neuron could “learn” something that changed its local behavior. McCulloch and Pitts proved that a network of these units could be wired up or “trained” to represent any proposition whatever, based on logical operations on its input.

This was an inspiring idealization, one of the great oversimplifications of all time, since while it has turned out that real neurons in interaction are *much* more complicated than the “logical neurons” defined by McCulloch and Pitts, they had demonstrated the logical possibility of a general purpose representing-and-learning-and-controlling network made out of units that performed simple, non-miraculous, clueless tasks—a comprehender of sorts made of merely competent parts. Ever since then, the goal of computational neuroscience has been to determine just which, of the infinite variety of more complicated networks, are operating in nervous systems. The wiring diagram of the nematode worm, *C. elegans*, with its 302 neurons of 118 varieties, is now just about complete, and its operation is becoming understood at the level of individual neuron-to-neuron actions. The Human Connectome Project aspires to make an equally detailed map of the tens of billions of neurons in our brains, and the Human Brain Project in Europe aspires to “simulate the complete human brain on supercomputers” but these mega-projects are in their early days. The brain is certainly not a digital computer running binary code, but it is still a kind of computer, **and I will have more on this in later chapters.**

Fortunately, a lot of progress has been made on understanding the computational architecture of our brains at a less microscopic level, but it depends on postponing answers to almost all the questions about the incredibly convoluted details of individual neuron connectivity and activity (which probably differs dramatically from person to person, in any case, unlike the strict uniformity found in *C. elegans*). Learning that a specific small region (containing millions of neurons) is particularly active when looking at faces (Kanwisher, 1997, 2013) is a valuable breakthrough, for instance. We know that *information about faces* is somehow being processed by the activity of these neurons even if we as yet have scant idea about just what they are doing and how they are doing it. This use of the term “information,” which is ubiquitous in cognitive science (and elsewhere), does *not* refer to Shannon information, even if some day far in the future we will be able to measure the bandwidth or storage capacity, in bits, of whatever physically encodes

whatever it is that is encoded.<sup>3</sup> This is semantic information *about* something specific: faces.

## 2. How can we characterize semantic information?

Information is information, not matter or energy. No materialism that does not admit this can survive at the present day.

--Norbert Wiener

The ubiquity of ADCs in our lives, now taken for granted in almost all informational transmissions, probably plays a big role in entangling Shannon's mathematical concept of information with our everyday concept of semantic information, in spite of numerous alerts issued. A high-resolution color photograph of confetti on a sidewalk, broken down into 8 million pixels, might fill a file ten times larger than a text file of, say, Adam Smith's *The Wealth of Nations*, which can be squeezed into two megabytes. Depending on the coding systems you use (GIF or JPEG or . . . . Word or PDF or . . . . .) a picture can be "worth a thousand words", measured in bits, but there is a better sense in which a picture can be worth a thousand words or more. Can that sense be formalized? Can semantic information be quantified, defined, theorized about? Robert Anton Wilson, an author of science fiction and writer on science proposed the Jesus unit, defined as the amount of (scientific) information known during the lifetime of Jesus. (Scientific information is a subset of semantic information, leaving out all the semantic information then available about who lived where, what color somebody's robe was, what Pilate had for breakfast, etc.) There was exactly one Jesus of scientific information in AD 30, by definition, an amount that didn't double (according to Wilson) until the Renaissance fifteen hundred years later. By 1750 it doubled again to 4 Jesus, and doubled to 8 Jesus in 1900. By 1964 there were 64 Jesus, and Lord knows how many Jesus (Jesuses?) have accumulated in the meantime. The unit hasn't caught on, thank goodness, and while Wilson is undoubtedly right to dramatize the information-explosion theme, it is far from clear that any scale or measure of the amount of (scientific) information will be an improvement on such precise-but-tangential measures as number of peer-reviewed journal pages or megabytes of text and data in on-line journals.

Luciano Floridi, in his useful primer (2010), distinguishes *economic* information as whatever is *worth some work*. A farmer is wise to take the time and trouble to count his cows, check the level of the water in the well, keep track of how

---

<sup>3</sup> Giulio Tononi (2008) has proposed a mathematical theory of consciousness as "integrated information" that utilizes Shannon information theory in a novel way, and has a very limited role for *aboutness*: it measures the amount of Shannon information a system or mechanism has about its own previous state—that is the states of all its parts. As I understand it, Tononi's theory presupposes a digital, but not necessarily binary, system, since it has a countable repertoire of output states.

efficiently his farm hands labor. If he doesn't want to do this supervisory work himself, he ought to hire somebody to do it. It "pays you" to find out something about your products, your raw materials, your competition, your capital, your location, . . . . to take the obvious categories only. You can canvass public information about markets and trends, and buy examples of your competition's products to reverse engineer and otherwise test against your own products, or you can try to engage in industrial espionage. *Trade secrets* are a well established legal category of information which can be stolen (or carelessly leaked or given away) and the laws of patent and copyright provide restraints on the uses others can make of information developed by the R&D of individuals and larger systems. Economic information is valuable, sometimes very valuable, and the methods of preserving information and protecting it from the prying eyes of the competition mirror the methods that have evolved in nature for the same purpose.

The mathematical theory of games (Von Neumann and Morgenstern, 1944) was another brilliant wartime innovation. It highlighted, for the first time, the cardinal value of keeping one's intentions and plans secret from the opposition. Having a poker face isn't just for playing poker, and too much transparency is quite literally death to any individual or organization that has to compete in the cruel world. (Dennett, 2013, Dennett and Roy, 2015). Survival, in short, depends on information, and moreover depends on differential or asymmetric information: I know some things that you don't know, and you know some things that I don't know, and our wellbeing depends on keeping it that way. Even bacteria—even non-living viruses—engage in devious ruses to conceal or camouflage themselves, in an arms race with their intrusive, inquisitive competitors.<sup>4</sup>

So let's consider, as a first pass, defining semantic information as *design worth getting*, and let's leave the term "design" deliberately noncommittal for the time being, allowing only for the point I stressed in chapter III, that design without a designer (in the sense of an intelligent designer) is a real and important category. Design always involves R&D work of one sort or another, and now we can say what kind of work it is: using available *semantic information* to improve the prospects of something by adjusting its parts in some appropriate way. *Learning useful facts* (about where the fishing is good, who your friends are) is one of the most common ways of improving one's design as an agent in the world, but it is only one. Learning how to make a fishhook, or avoid your enemies, or whistle is another kind of design improvement. All learning—both learning what and learning how—can be a valuable supplement or revision to the design you were born with. (What about

---

<sup>4</sup> For instance, the Ebola virus mimics an apoptotic cell fragment in order to get itself "eaten" by a phagocytic (trash-collecting) cell, thereby hitching a safe ride in a cell that will carry it around in the body (Misasi and Sullivan, 2014). There are many well-studied examples of viral and bacterial camouflage and mimicry, and biotechnologists are now copying the strategy, masking nano-artifacts that would otherwise be attacked by the immune system.

misinformation, which can also accrue, and disinformation, which is deliberately implanted in you? These phenomena seem to be simple counterexamples to our proposed definition, but as we shall soon see, they really aren't.) The definitions of "semantic information" and "design" are thus linked in circular definitions, but it's a virtuous, not vicious, circle: some processes can be seen to be R&D processes in which some aspect of the environment is singled out somehow and then exploited to improve the design of some salient collection or system of things in the sense of better equipping it for thriving/persisting/reproducing in the future.

Semantic information, then, is "a distinction that makes a difference." Floridi (2010) reports that it was D. M. MacKay, who first coined this phrase, later enunciated by Gregory Bateson (1973, 1980) and others as *a difference that makes a difference*. MacKay was yet another of the brilliant theoreticians who emerged from the intense research efforts of World War II alongside Turing and Shannon (and von Neumann and John McCarthy, among others). He was a pioneer information theorist, physicist, neuroscientist and even a philosopher (and one of my personal heroes, in spite of his deep religious convictions).<sup>5</sup> My proposed definition of semantic information is close to his 1950 definition of "information in general as that which *justifies representational activity*." (p158 of MacKay 1968). MacKay's focus at the time was on what we now call Shannon information, but he ventured wise observations on information in the broader, semantic sense, also defining it as *that which determines form* (1950, diagram, p159 of MacKay, 1968), which points away from representation (in any narrow sense) while keeping the theme of justification or value in place.

If information is a distinction that makes a difference, we are invited to ask: a difference to whom? *Cui bono?* Who benefits?—the question that should always be on the lips of the adaptationist, since the answer is often surprising (Dennett, 1995). It is this that ties together economic information in our everyday human lives with biological information, and unites them under the umbrella of semantic information. And it is this that permits us to characterizing misinformation and disinformation as not just kinds of information, but *dependent* or even *parasitic* kinds of information. Something emerges as *misinformation* only in the context of a system that is designed to deliver—and rely on—*useful* information. An organism that simply ignores distinctions that might mislead (damage the design of) another organism has not been misinformed, even if the distinction registers somehow (ineffectually) on the organism's nervous system. In Stevie Smith's poem, "Not Waving but Drowning" (1972), the onlookers on the beach who waved back were misinformed, but not the seagulls wheeling overhead. We can't be misinformed by distinctions we ignore. Understanding what disinformation is benefits doubly from our asking *cui bono?* Disinformation is the designed exploitation (for the benefit of one agent) of

---

<sup>5</sup> The Ratio Club, founded in 1949 by the neurologist John Bates at Cambridge University, included Donald MacKay, Alan Turing, Grey Walter, I.J. Good, William Ross Ashby and Horace Barlow, among others. Imagine what their meetings must have been like!

another agent's systems of discrimination, which themselves are designed to pick up useful information and *use it*. This is what makes the Ebola virus's design an instance of camouflage.

Kim Sterelny (personal correspondence) has raised an important objection:

Humans are representational magpies — think of how incredibly rich forager fauna and floras are — much of their natural history information has no practical value. Once information storage (in the head and out of the head) became cheap, the magpie habit is adaptive, because it is so hard to tell in advance which of the informational odds and ends will turn out to be valuable. But that does not alter the fact that most of it will not.

He goes on to claim most of what anybody knows is “adaptively inert. But that does not matter, since it is cheap to store, and the bits that do matter, *really* matter.” Setting aside the misleading use of the term “bits” here, I largely agree, but want to issue one proviso: even the dross that sticks in our heads from the flood that bombards us every day has its utility profile. Much of it sticks because it is *designed to stick*, by advertisers and propagandists and other agents whose interests are served by building up outposts of recognition in other agents' minds, and, as Sterelny notes, much of the rest of it sticks because it has some non-zero probability (according to our unconscious evaluations) of being adaptive someday. People—real or mythical—with truly “photographic memories” are suffering from a debilitating pathology, burdening their heads with worse than useless information.<sup>6</sup>

Evolution by natural selection churns away, automatically extracting tiny amounts (not “bits”) of information from the interactions between phenotypes (whole, equipped organisms) and their surrounding environments. It does this by automatically letting the better phenotypes reproduce their genes more than the less favored.<sup>7</sup> Over time, designs are “discovered” and refined, thanks to these encounters with information. R&D happens, designs are improved because they all have to “pay for themselves” in differential reproduction, and Darwinian lineages “learn” new tricks by adjusting their *form*. They are, then, *in-formed*, a valuable step up in local Design Space. In the same way, Skinnerian, Popperian and Gregorian creatures *inform* themselves during their own lifetimes by their encounters with their environments, becoming ever more effective agents thanks to the information

---

<sup>6</sup> Science has hugely expanded our capacity to discover differences that make a difference. A child can learn the age of a tree by counting its rings, and an evolutionary biologist can learn roughly how many millions years ago two birds shared a common ancestor by counting the differences in their DNA, but these pieces of information are not playing any role in the design of the tree or the bird; the information is not *for* them, but has now become information for us.

<sup>7</sup> Colgate and Ziock (2010) defend a definition of information as “that which is selected,” which is certainly congenial to my definition, but in order to make it fit the cases they consider, the term “selected” has to be allowed to wander somewhat.



they can now *use* to do all manner of new things, including developing new ways of further informing themselves. The rich get richer. And richer and richer, using their information to refine the information they use to refine the information they obtain by the systems they design to improve the information available to them when they set out to design something.

This concept of useful information is a descendant of J. J. Gibson's concept of affordances, introduced in chapter V. I want to expand his account to include not only the affordances of plants and other non-animal evolvers, but also the artifacts of human culture. Gibson says "the information is in the light" and it is by "picking up" the information that animals perceive the world.<sup>8</sup> Consider the reflected sunlight falling on the trunk of a tree, and on the squirrel clinging to the trunk. The same *potential* information is in the light for both tree and squirrel, but the tree is not equipped (by earlier R&D in its lineage) to make much of anything from *the information in the light*. (The tree uses the *energy* in the light to make sugars, by photosynthesis.)<sup>9</sup> and it is the utility to the *receiver* that matters.

We can think about potential utility if we like: suppose a man with a chain saw is approaching, visible to any organism with eyes, but not to the tree. Eyes are of no use to a tree unless it also has some way of using the information (if not to run and hide, perhaps to drop a heavy limb on the lumberjack, or secrete some sticky sap that will gum up the saw). The presence of the information in the light might someday "motivate" a trend in the direction of eyesight to trees, if a behavioral payoff were nearby! Unlikely, but such unlikely convergences are the heart of evolution. There has to be a difference that could make a difference if only something were present to hinge on it somehow. As just noted, evolution by natural selection is astonishingly good at finding needles in haystacks, almost invisible patterns that, when adventitiously responded to, yield a benefit to the responder. Just as the origin of life depends on getting the right 'feedstock' molecules in the right place at the right time, there has to be raw material in the variation in the population that includes, by sheer coincidence, some heretofore functionless (or under-utilized or redundant or vestigial) feature that happens to be heritable and that covaries with the potentially useful information in the world.<sup>10</sup>

---

<sup>8</sup> Notoriously, Gibson doesn't just ignore the question of what internal machinery manages to do this pick-up; he often seems to deny that there are any difficult questions to answer here. The slogan of radical Gibsonians says it all: "It's not what's in your head; it's what your head is in." I am not at all endorsing this view.

<sup>9</sup> Thanks to Kim Sterelny for drawing this nice contrast to my attention. As Norbert Wiener once said,

<sup>10</sup> Useful information emerges even at the molecular level. David Haig, in his fascinating essay, "The Social Gene" (1997), exploits the agential perspective all the way down to what he calls the *strategic gene* concept. As he notes, "The origin of molecules that were able to discriminate between themselves and closely similar molecules greatly expanded the strategies available to genes and made possible the evolution of large multicellular bodies." (p294). When there is no information

Not everything “possible in principle” is automatically available, but given lots of time, and lots of cycles, there are *likely* to be paths of happenstance that lead to the Good Tricks in the neighborhood (Dennett, 1995). But not always. That is why plausible “just so stories” (Gould & Lewontin, 1979, Dennett, 1995) are only *candidate* explanations, in need of confirmation. Every well-confirmed evolutionary hypothesis (of which there are thousands) began as a just-so story in need of supporting evidence. And just as most organisms born die childless, the majority of just-so stories that get *conceived* never earn the right to reproduce. The sin of adaptationism is not conceiving of just-so stories—you can’t do evolutionary biology without this---but uncritically reproducing just-so stories that haven’t been properly tested.

Consider a less fantastical possibility: brilliant autumn foliage is understood to be not an adaptation, but merely a functionless byproduct of the chemical changes that occur in deciduous leaves when they die. The leaves stop making chlorophyll when the sunlight diminishes, and as the chlorophyll decomposes, other chemicals present in the leaves—carotenoids, flavonoids, anthocyanins—emerge to reflect the remaining light. Today, human beings, especially in New England, value the brilliant colors of autumn foliage, and—usually unconsciously—foster the health and reproduction of the most impressive trees by cutting down other trees first, saving the nice colors for another season. Having bright foliage is already an adaptation in the trees of northern New England, if not yet measurable directly. This is how adaptations start, imperceptible to all except the unblinking process of natural selection. There is considerable difference in how long deciduous trees hold their leaves in the autumn; the dull brown leaves of oaks are the last to fall, long after the brilliant maples have become bare. Any chemical changes that permitted some strains of maple to hold their foliage longer would become an adaptation in any environment where people exert a non-negligible selective effect (knowingly or not). And now let’s stretch our imaginations a few steps further: suppose that this foliage-prolongation capacity is itself quite energetically costly, and only pays when there are people around to appreciate the colors. The evolution of a people-present-detector (more likely a pheromone sniffer than a rudimentary people-detecting eye) could roll onto the scene. This would be a step towards *self-domestication* by the tree lineage. Once offspring are favored, once reproduction is fostered or prevented by us, we’re on the way to a domesticated species, alongside date palms and avocado trees. The opening moves need not be the result of conscious, deliberate, *intelligent* choice by us (and certainly not by the trees). We can, in fact, be entirely oblivious to the lineages that become *synanthropic*, evolving to thrive in human company without belonging to us or being favored by us. Bedbugs and mice are synanthropic, and so is crabgrass, not to mention the trillions of tiny things that inhabit our bodies, flying beneath the radar if possible, adapted to living in or on or next to the human body niche.

---

available to exploit, genes are unable to defect, or form coalitions, with better than chance success.

In all these cases, *semantic information about* how best to fit in has been mindlessly gleaned from the cycle of generations, and notice that it is not *encoded* directly in the organisms' nervous systems (if they have any) or even in their DNA, except by something akin to *pragmatic implication*. Linguists and philosophers of language use the term *pragmatics* to refer to those aspects of meaning that are not carried by the syntax and "lexical" meanings of the words but conveyed by circumstances of particular utterances, by the *Umwelt*, in effect, of an utterance. Think of all the different "things" you might say by uttering "Dream on!" in one context or another. If I burst into a house and yell to all assembled "Put on the kettle!" I have uttered an imperative English sentence, but some will probably infer that I would like to have a cup of tea or other hot beverage, while another may further surmise that I feel myself at home here, and may in fact be the occupant of this house. Yet another person present, a monoglot Hungarian, may infer only that I speak English, and so does whomever I am addressing (well, it sounds like English, doesn't it?), while somebody *really* in the know will be instantly informed that I have decided after all to steam open that sealed envelope and surreptitiously read the letter inside in spite of the fact that it isn't addressed to me; a crime is about to be committed. What semantic information can be gleaned from the event depends on what information the gleaner already has accumulated. Learning that somebody speaks English can be a valuable update to your world knowledge, a design improvement that may someday pay big dividends. Learning that somebody is about to commit a crime is also a valuable enhancement to anyone in a position to put it to good use. The local design improvements each can make on the basis of this interaction are wildly different, and it would be an utter mistake to think the semantic information could all be extracted by looking closely at the structure of the signal as an acoustic wave *or* as an English sentence ("p-u-t-space-o-n-space. . ."). *There is no code* for all these different lessons-learned.

Similarly, the DNA of a bird in a lineage that has "learned" to make a hanging nest when the time comes for breeding will not have codon sequences that describe the nest or how to build it step by step, but will consist of sequences of imperatives along the lines of "then attach a lysine, then a threonine, then tryptophan, . . ." (the recipe for making one protein or another out of a string of amino acids) or "fiddle-dee-count-to-three" (a bit of "junk" DNA used as a timer), or just "blahblahblah" (a genomic parasite or other bit of junk). In any event, don't hope to "translate" any sequence of codons as "nest" or "twig" or "find" or "insert". Still, thanks to the particular string of codons inherited by the bird from its parents, and thanks to the developmental systems that had already "learned" in earlier evolutionary steps to interpret those codons, the knowhow to build the hanging nest will have been transmitted from parent to offspring. The offspring inherit a manifest image with an ontology of affordances from their parents, and are born ready to distinguish the things that are most important to them. Reflecting on how much one would have to know in the kettle case to understand the message "fully" is a good way of appreciating how impenetrable to analysis the transmission of knowhow via DNA from one generation to the next is apt to be.

There will be (so far as I can see) no privileged metric for saying *how much* semantic information is “carried” in any particular signal. We will have to make do, I think, with disciplined hand-waving, relying on our everyday familiarity with how *we* can communicate so much to our fellow human beings by saying things and doing things. (Yes, *disciplined* hand-waving. There may not be any algorithms for measuring semantic information, but we can erect any number of temporary structures for *approximating* the information content of the topics of interest to us, relying on posited ontologies—the ontologies that furnish the *Umwelten* of organisms, “possible worlds” or models.) We already do this every day, of course, using our human categories and ontologies in a rough and ready way to say quite rigorously controlled things about what categories animals distinguish, what tasks they perform, what they fear and like and avoid and seek. We just need to restrain ourselves from *assuming* what many theorists would like to assume: that if an organism is competent with the categories of *predator, edible, dangerous, home, mother, mate, . . .* they must have a “language of thought” with terms in it for each of these categories. If DNA can convey information about how to build a nest without any terms for “build” and “nest,” why couldn’t a nervous system do something equally inscrutable?<sup>11</sup>

All evolution by natural selection is design revision, and for the most part it is design improvement (or at least design-maintenance). Even the loss of organs and their functions counts as improvement when the cost of maintaining them is factored in. The famous cave fish that have abandoned vision are engaged in cost-cutting, which any company executive will tell you is design improvement. Don’t acquire and maintain what doesn’t pay for itself. It is no accident that biologists often speak of a lineage as “learning” its instinctual behaviors over the generations, because all learning can be similarly seen to be processes of self-redesign, and in the default case, improvement of design. We see the acquisition of knowhow, or factual information, as learning, and it is always a matter of using the base of competence/knowledge you already have to exercise quality control over what you acquire. Forgetting is not usually considered learning any more than discarding is usually seen as design improvement, but sometimes it is. More is not always better. The legal distinction between flotsam and jetsam is apropos here: flotsam is cargo that has inadvertently or accidentally been swept off the deck or out of the hold of a ship, while jetsam is cargo that has been deliberately thrown overboard—jettisoned. Intentional mind-clearing, jettisoning information or habits

---

<sup>11</sup> The definition of Colgate and Ziock (2010) has a further condition that I am explicitly denying: “To usefully select information, information must be stored (written); otherwise there is no way to decide what has been selected.” (p58). It all depends on what “stored (written)” means. I would say that information about nest building is stored and transmitted in the bird lineage but not written down (as information about nest building).

that endanger one's welfare, is not an unusual phenomenon, sometimes called unlearning.<sup>12</sup>

Semantic information is not always valuable to one who carries it. Not only can a person be burdened with useless facts, but often particular items of information are an emotional burden as well. This doesn't cancel the link to utility in the definition of semantic information; it complicates it. (The value of gold coins is not put in doubt by the undeniable fact that pockets full of gold coins may drown a strong swimmer.) Still, defining semantic information as *design worth getting* seems to fly in the face of the fact that so much of the semantic information that streams into our heads every day is *not* worth getting, and is in fact a detestable nuisance, clogging up our control systems and distracting us from the tasks we ought to be engaged in. But we can turn this "bug" in our definition into a "feature" by noting that the very existence of information-handling systems depends on the design value of the information that justifies the expense of building them in the first place. Once in place, an information-handling system (a pair of eyes or ears, a radio, the internet) can be exploited—parasitized—by noise, of several species: sheer meaningless "random" *white noise*, the raspy "static" that interferes with your transistor radio when the signal is weak, and semantic information that is useless or harmful to the receiver. **Spam and phishing emails on the internet are obvious examples (and more on such phenomena in chapter N)**, both dust clouds and (deliberately released) squid ink are others. The malicious items depend for their effect on the trust of the sender in the medium. Since Aesop we've known that the Boy Who Cried Wolf stops commanding attention and credence after a while. Batesian mimicry (such as a nonpoisonous snake with markings that mimic a poisonous variety) is a similar kind of parasitism, getting a benefit without going to the cost of manufacturing poison, and when the mimics outnumber the genuinely poisonous snakes Aesop's moral takes hold and the deceitful signal loses its potency.

Any information-transmitting medium or channel can set off an arms race of deception and detection, but *within an organism*, the channels *tend* to be highly reliable, since all "parties" have a common fate, sinking or swimming together, so that aside from some fascinating internal rebellions (such as genomic parasites), trust reigns (Sterelny, 2003 [Hostile World], McKay and Dennett, 2009). Error is always possible, the result of simple breakdown—wear and tear—of the system, or misapplication of the system to environments it is ill-equipped to handle. This is why delusions and illusions are such a rich source of evidence in cognitive neuroscience, providing hints about what is being *relied upon* by the organism in the normal case. It is often noted that the brain's job in perception is to filter out, discard, ignore all but the noteworthy features of the flux of energy striking one's sensory organs. Keep and refine the ore of (useful) information, and leave all the

---

<sup>12</sup> Robert Mathai notes (personal correspondence) that evolutionary unlearning is never really jetsam; it is flotsam that, over time, comes to resemble jetsam. It wasn't jettisoned with foresight, but its washing overboard proves to have been a good thing, saving the ship.

noise out. Any non-randomness in the flux is a real pattern (Dennett, 1991 [real patterns]) that is *potentially useful* information for some possible creature or agent to exploit in anticipating the future. A tiny subset of the real patterns in the world of any agent comprise the agent's *Umwelt*, the set of its affordances; these patterns are the *things* that agent should have in its ontology, the things that should be attended to, tracked, distinguished, studied. The rest of the real patterns in the flux are just noise *as far as that agent is concerned*. From our Olympian standpoint (we are not gods, but we are cognitively head-and-shoulders above the rest of the creatures), we can often see that there is semantic information in the world that is intensely relevant to the welfare of creatures who are just unequipped to detect it. The information is indeed in the light, but not for them.

### 3. Trade secrets, patents, copyright, and Bird's influence on bebop

My claims, then, are that:

- (1) Semantic information is valuable—misinformation and disinformation are either pathologies or parasitic perversions of the default cases.
- (2) The value of semantic information is receiver-relative and not *measurable* in any non-arbitrary way, but can be *confirmed* by empirical test.
- (3) The *amount* of semantic information carried or contained in any delimited episode or item is also not usefully *measurable in units* but roughly comparable in local circumstances.
- (4) Semantic information need not be *encoded* to be transmitted or saved.

All these claims get clarified and supported when we turn to human “economic” information, and consider the way human societies have enshrined these claims in their laws and practices. Consider a case of stealing a trade secret. Your competitor United Gadgets, has developed a new widget, an internal component of its very powerful new strimpulizer, but you can't get a good look at it, since it is “potted” in an x-ray-opaque casing that can only be cracked open or dissolved by something that destroys the widget in the process. A well kept secret indeed. You go to great lengths to embed a spy in United Gadgets, who eventually encounters a bare widget. Almost home. Now how to get the information out? A plaster piece-mold cast, a negative, would be wonderful but too bulky to sneak out. Drawings, photographs, or blueprints would be good, but also hard to carry out, since security is fierce and radio signals closely monitored. A very precise recipe, in English words, for making the device, might be encrypted and then concealed in an otherwise innocuous message, a rambling memo about health insurance options and their many conditions, for instance.

Another recipe system would be a CAD-CAM file; put the widget in a CAT-scanner (a Computer Aided Tomography machine) and obtain a suitably high-resolution slice-by-slice tomograph to use as a recipe for your 3D printer. Depending on the resolution, this could be ideal, providing, in the limit, a recipe for making an atom-for-atom duplicate (a fantasy much fancied by philosophers and other fans of teleportation). One virtue of this extreme variation is that it arguably produces the maximally detailed specification of the widget in a file whose size *can* be measured in bits! You can send “perfect” information about the widget, specifying its every atom, in a file of only umpteen million zetabytes. Today a widget, tomorrow the world. The idea, generalized, of the whole universe being exhaustively (?) describable in one cosmic bitmap, lies at the heart of various largely speculative but fascinating proposals in physics. And of course it doesn’t stop at big old atoms, a relatively “low-res” recipe for reality these days. Such an application of Shannon information theory does permit, “in principle” but not remotely in practice, saying exactly how much (Shannon) information there is in the cubic meter of ocean and ocean floor surrounding a particular clam, for instance, but it says nothing about how much of this information--a Vanishingly (Dennett, 1995) small portion—is semantic information for the clam.

Back to the theft of the widget design. The CAD-CAM file, of suitably high resolution, could be stored on a tiny digital memory device that might be safely swallowed. But if you can’t arrange to do the tomography, your spy may have to simply study the widget intensely, turning it every which way, hefting and twisting it, sniffing and tasting it, trying to dent it with his fingernails, and so forth, and then do what he can to memorize the shape somehow—a tall order if the shape is highly irregular and made to very fine tolerances—and carry the information out in his bare brain. Probably the best way to steal the information, if you can get away with it, is to borrow a widget, take it home, examine and record it in any way you like so you can make a copy that meets your requirements, and then return it to United Gadgets. *All you took was the information.*

Suppose your spy succeeds. The information you acquired enables you to improve the design of your own strimpulizer, and thereby improve your market share, and make a living. United Gadgets finds out (you didn’t bother potting yours), and sues you—or worse, has you arrested for industrial espionage. It can be beyond reasonable doubt that you stole the design of the widget, even if the prosecution cannot definitely establish how you did it. If the widget is that unusual, the case will be just like plagiarism, which can be proven on the basis of design-replication (with improvements and variations) alone. In fact, if United Gadgets anticipates theft from the outset, they would be wise to incorporate a distinctive but non-functional knob or pit or slot in its design, which would be a dead giveaway of theft if it appears in your version. (This is the tactic long used by encyclopedias to catch competitors copying their entries illicitly; a fictitious animal or poet or mountain that appeared with roughly the same description in both books would be hard for the culprit to explain. Look up Virginia Mountweazel in Google for details.) Notice that this telltale feature carries information of great utility to the original “sender”

only so long as the original receiver does not recognize it as an intended signal (like the signal of the injury feigning bird), while the copier unwittingly “sends” a self-damaging message back, conveying the information that it has illicitly copied the original.

We can learn another lesson, I think, from the law of patent and copyright. First, these laws were enacted to protect designs, in this case designs created by intelligent designers, people. People are often designers, and designing takes time and energy (and a modicum of intelligence unless you are an utter trial-and-error plodder, an R&D method which almost never bears interesting fruit except over evolutionary time). The designs that result are typically valuable (to somebody), so laws protecting the owners/creators of these designs are reasonable.

A few features of these laws stand out. First, you have to demonstrate the *utility* of your brainchild to get a patent. And you have to demonstrate that nobody else has already invented it. How useful and original does it have to be? Here is where there is ineliminable hand-waving in the law. For instance, the Canadian law of patent excludes as not sufficiently novel (to warrant a patent) any invention for which there has been *anticipation*, which can be demonstrated if any of eight conditions obtains (Wikipedia, article on *novelty (patent)*). Two of the conditions are officially described thus: anticipation may

convey information so that a person grappling with the same problem must be able to say "that gives me what I wish"

give information to a person of ordinary knowledge so that he must at once perceive the invention.

It should not surprise us that patent law has this awkward issue of definition. Novelty, like semantic information in general, depends very directly and heavily on the competence of the parties involved. In the Land of the Dunces you could patent a brick as a doorstop; in Engineers' Heaven, a flying house powered by sunlight might be regarded as a trivial extension of existing knowledge and practices. What can be seen “at once” by a “person of ordinary knowledge” will vary with what counts as ordinary knowledge at any time and place.

You can patent an idea (for a process, a gadget, tool a method) without actually building a working prototype, though it certainly is advisable to build one, to support your sketches and descriptions. You can't copyright an idea, but only a particular expression of an idea. A song can be copyrighted, but probably not a four-note sequence. If Beethoven were alive today, could he copyright the first four notes of his Fifth Symphony (Ta-ta-ta-DAH)? Has NBC copyrighted its three-note chime, or is that just a trademark—a different legally protected informational item? A book title by itself cannot be copyrighted. A short poem can. How about this one:

This verse



is terse.

It is no doubt covered by the copyright on this entire book. As a stand-alone “literary work,” I wonder. Copyright law has been amended many times, and vexed issues continue to be problematic. Books and articles, music and painting, drawing, sculpture, choreography, and architecture, can be copyrighted *as long as there has been some fixed expression*. An improvised jazz solo line in an unrecorded performance cannot be copyrighted, nor can an unchoreographed and unfiled dance routine. This is in line with Colgate and Ziock’s definition of information, but seems to be driven by legal requirements of evidence, not by any natural requirement for information to be “stored (written)”. Charlie (“Bird”) Parker couldn’t copyright his solos, but many saxophone players and other jazz musicians who heard them went on to be heavily influenced by him, which is to say: valuable semantic information flowed from his performances to theirs (and not to the tin ears in his audiences who couldn’t pick up the affordances made available). You can’t copyright an idea or discovery, but “where do we draw the line?” As Judge Learned Hand once said,

Obviously, no principle can be stated as to when an imitator has gone beyond copying the ‘idea,’ and has borrowed its ‘expression.’ Decisions must therefore inevitably be ad hoc.” [Peter Pan Fabrics, Inc. v. Martin Weiner Corp., 274 F.2d 487 (2d Cir. 1960). *cited in Wikipedia on copyright.*]

Interestingly, when considering whether or not an item is copyrightable, utility or function works *against* a creation, since copyright is intended to protect “artistic” creation, which must be “conceptually separable” from functional considerations, where the law of patent applies (under more stringent conditions). This is an interestingly truncated conception of “function,” since *aesthetic* effects are obviously functional in most if not all contexts. This reminds me of the myopic rejection by some early evolutionists of Darwin’s ideas about sexual selection, since it implied—mistakenly in their eyes—that the perception of beauty could have a functional role to play. But of course it does; females have evolved heightened competence in discernment of good properties in prospective mates while the males have evolved ever more impressive (aesthetically impressive, good for nothing but show) displays. Mating successfully is not an optional adventure in life; it’s the finish line, the goal, and whatever it takes to hit the tape is functional whatever cost or burden it places on its bearer.<sup>13</sup> The law of copyright tries to separate “utilitarian” function

---

<sup>13</sup> I find that some people resist the quite obvious free-floating rationale for why it is females that do the evaluating and males that do the strutting and costly advertising: asymmetrical parental investment. In species where females normally invest more time and energy per offspring than males do (making eggs vs making sperm, lactating, incubating, rearing. . . .) the females should be more choosy when accepting a mate. A female can make only so many eggs or bear only so many litters of young, and if she chooses to mate with a second-rate male she may waste most or all her bearing capacity, while a male who chooses to mate with a second-rate

from aesthetic function, and while there are good legal reasons for trying to make this a bright line distinction, there are good theoretical reasons for seeing this as an ad hoc undertaking, as Learned Hand observes about the “idea”/“expression” distinction.

When we turn to cultural evolution we will encounter this sort of codeless transmission on a grand scale. As I have often noted, a wagon with spoked wheels doesn't just carry grain or freight from place to place; it carries the brilliant idea of a wagon with spoked wheels. (Dennett, 1991, p204, 1995, p348) The wagon no more carries this idea—this information—to the dog in the road than the sunlight carries the information about the approaching lumberjack to the tree. You have to be informed to begin with, you have to have many competences installed, before you can avail yourself of this information, but it is there, embodied in the vehicle that rolls by. One of the tasks remaining for us is to understand how come we human beings are so much better at extracting information from the environment than any other species.

Are we really better? Beyond a doubt, if we measure our prowess in numbers of affordances. In addition to those we share with our mammalian cousins (water to drink, food to eat, holes to hide in, paths to follow, . . . .) there are all our recognizable, familiar artifacts. A hardware store is a museum of affordances, with hundreds of different fasteners, openers, closers, spreaders, diggers, smoothers, reachers, grippers, cutters, writers, storers, . . . all recognizable and usable in appropriate circumstances, including novel circumstances in which we invent and construct new affordances ad lib, using the pre-designed, pre-manufactured parts. Richard Gregory, whose reflections on intelligence and equipment inspired me to name Gregorian creatures after him, emphasized that it not only took intelligence to use a pair of scissors, the pair of scissors *enhanced* the “intelligence” of users by giving them a large increase in available competence. These tools, like the hermit crab's shell, the bird's nest, the beaver's dam, are a part of our extended phenotypes, acquired design improvements that are neither innate nor organic; what is innate is the talent to recognize such things and put them to good use.

I sometimes suggest to my students that evolution by natural selection is nothing but “universal plagiarism”: if it's of use to you, copy it, and use it. All the R&D that went into configuring whatever it is that you copy is now part of your legacy; you've added to your wealth without having to “reinvent the wheel.” That's what Nature has been doing for billions of years, refining and articulating and spreading billions of good design features into every corner of the planet. This tremendous creativity would not be possible without almost unimaginable amounts of copying. Nature's Good Trick is against the law, for us, and there is a good reason

---

female has lost only a little of his precious time and can always make more sperm. In species where parental investment is roughly equal, the males and females look and act pretty much the same.

for this: semantic information is costly to create and valuable, so unauthorized copying is theft. It is worth noting that this is not a free-floating rationale. The first laws of patent and copyright (and trade secret and trade mark) were devised and enacted following extended explicit rational debate and discussion about the need for them. They themselves are products of intelligent design, designed to protect other intelligent designs.

*Chapter VI Summary*

TK

*Chapter VII Preview*

TK

*Hanging Questions*

If brains aren't digital computers what kind of computers are they? [ms p4]

Cui bono? Who benefits?—the question that should always be on the lips of the adaptationist, since the answer is often surprising (Dennett, 1995). [ms p7]

Spam and phishing emails on the internet are obvious examples (and more on such phenomena in chapter N), [ms p11]