The Value of Meaning

by Robert Adámy Duisberg

I sit with my bees. Their incessant activity is hypnotic, like staring into a fire.

Foragers fly and return. Packed with pollens, they pause for a moment before entering the redolent darkness, taking an instant to check with sentries stationed at the threshold, antennae touching, fluttering, as information, some pheromonal status report, is shared.

A hive behaves with a kind of intelligence. It makes collective decisions about where to collect food, how to manage its internal populations, whether to swarm, and if so to where. It bases these decisions on information collected and shared by individual bees about its place, the seasons, what's in bloom and where. Knowledge is conveyed symbolically through their famous "waggle dances," through pheromones, and individual exchanges between bees. How can I engage with an intelligence so alien and distributed?

Once as I sat with them, one hive's energy at their doorstep began to surge. In a moment bees were streaming out like a poured liquid, covering their entire landing board and taking to the air by the thousands. The air around me was electric, roaring with the energy of cascading bees in their vibrance. This was the beginning of a days long journey to find and choose a new home. Having decided to divide, the hive had made new queens. I'd seen evidence of that, and now the old queen was leaving with half of her many thousands of daughters. I stood in the middle of that upsurging swarm, and I was riveted by the conviction that I was amidst a mind, whose very cognitive elements seethed about me, knowing without a doubt that the cognition of that mind was happening exactly where I stood, in that space between the bees.

How could I feel so certain? Whether a hive as a whole achieves consciousness is likely not knowable, inasmuch as we are scarcely able to grasp what exactly it is in ourselves. But just observing its behavior tells me the hive clearly considers its circumstances and acts decisively. Another time, for example, I harvested three frames full of honey from a particularly strong hive. I couldn't process the honey right away, so I left the frames in an empty hive box on the front porch until I could get to them. But the message got back to the robbed hive about where the frames had been moved to. The meaning of this message, however it was represented within the hive, was so compelling that immediately all forces, tens of thousands of bees, urgently mobilized. Suddenly we could no longer use the front door because the storm of activity out there was too furiously intimidating. By the end of the day the colony had emptied out those harvested frames, and cleanly ferried all fifteen pounds of honey back home! Decisive action, indeed, in response to a recognized crisis and a foolish beekeeper.

But more of my hives now have been collapsing in recent years. Collapsing is so different from merely dying. For in dying, the dead are all around; in collapse they simply vanish, as if raptured.

Collapse is always sudden and surprising. A strong and burgeoning hive will, over just three or four days, effectively evaporate. I've watched them during such brief declining hours. After the initial chunk of the population has abandoned home, those who remain, and the newly hatched, mill about and circle the air aimlessly. No purposeful work is done, no foraging. There are no sentries or reports, and wasps even probe the space to feed unchallenged. It is as if they had lost their collective mind.

###

Of course collapsing societies of pollinators is just one of the many collapses we can anticipate as we continue in this Anthropocene epoch. We have relied so heavily on models of prosperity that require growth as the absolute necessity, that we have forgotten the implications of ever exponential increase. Always scaling geometrically, we have been led directly, necessarily and mathematically toward resource and environmental extinction. Exponentials just don't sustain steady states. But exponential growth is the *only* model we have. Clearly, we must think anew.

Our own society is less harmonious than a well-ordered hive, and we are decidedly less able to respond both decisively and cooperatively to our shared and recognized crises. How then might we struggle to live, hoping to thrive, despite collapsing systems around us? What kind of story can we narrative-loving *sapiens* tell ourselves, about flourishing toward some livable, fulfilling future — one in which our lives are rich and meaningful, even prosperous, while not destroying our own world, this world, in all its exquisite particular creation?

Thinking toward a rich and meaningful future entails understanding what we mean in wishing it to be so. How do we construct compelling meanings in and amongst ourselves? Unavoidably, we must take on the following question, namely:

What do we mean when we talk about meaning?

A foraging bee returns from a bush in bloom, and tells of it through dances in the dark. She encodes information about the direction and distance to the nectar and its quality, in the angle, length and enthusiasm of her dance. Thus she recruits her sisters to join in the harvest of her discovery. Her sisters attune to her dancing, and the meanings they derive inform the foraging behavior of the whole colony.

We contemporary humans are awash in information, but how does it become meaningful? The noise of news feeds, data mined for monetization, endless streams of texts and images enable people to experience lives mediated by digital information. Whether that's a good thing is beyond this discussion. But we certainly know what we mean when we talk about Information.[1] When we define it formally, we measure it in bits and bytes. It has a mathematical description due to Claude Shannon, which yields handy ways to compress data into things like jpg, mp3 and pdf files. Without results from the mathematical theory of Information, computers couldn't do what they do.

Shannon's equation defining Information in communication channels happens to be identical to one written down by Ludwig Boltzmann 75 years earlier to define Entropy in thermodynamics, except for a minus sign. Information is precisely negative Entropy; this is not a metaphor. The equation involves probabilities, the likelihoods of each character to appear in a message stream. It's a familiar enough idea if you've ever played Scrabble[TM]. In that game, more unlikely letters win a higher score. 'Q' gets you more points than 'E' because it has a bigger surprise value. It is more unlikely, so it reduces the Entropy of your set of letters, which is just to say it packs more Information. Child's play. Though many adults enjoy Scrabble too.

But what does it all mean? Shannon was careful to only describe the statistics of Information coming down wires. He pointedly said nothing at all about interpretation, semantics, what the Information refers to, what or how it means. In my mind's ear I hear young Brigitta's voice in *The Sound of Music* as Maria teaches the do-re-mi of a scale, "But it doesn't *mean* anything!" However, when the doe becomes a deer in a ray of golden sun before me, as I like to call myself, Maria has created a web of meaningful relationships among otherwise arbitrary syllables, besides offering us a piece of genius song writing.

It's easy to feel overwhelmed by our information flood, because in and of itself, it can be meaningless. Someone or something needs to make sense out of it, and that takes effort. Meaning has to be found in and constructed out of raw Information.

There's a sweet spot between too little and too much Information. A dripping faucet is completely predictable, so each new drip conveys no new information, nor is it meaningful; it's just annoying. At the other extreme, utter noise could carry maximal Information according to the statistical definition, because of the high unpredictability of each new event, but that message too would be obviously meaningless.

---

1 I adopt the convention of capitalizing terms which refer to a mathematical definition, to distinguish them from the same words used in common parlance. It is admittedly rather germanic, but is helpful in distinguishing technical from demotic usage. So, for example, the heading of the section heading, "The meaning of Meaning," can be understood as indicating it is about implications and consequences of having a mathematical definition of Meaning. And while Entropy is not so common an every day term, I capitalize it also for consistency.

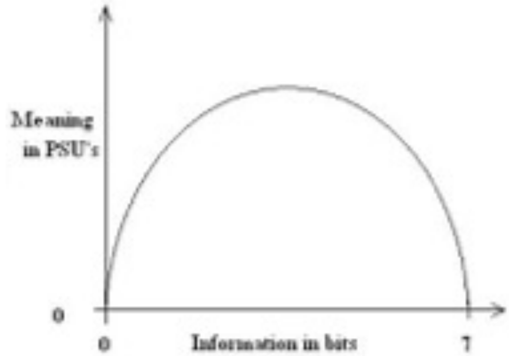

Figure 1: Meaningfulness of a message stream, as a function of its Information content. For now, while we have yet to define it, meaningfulness is suggested in Potter Stewart Units (PSU), in other words "I know it when I see it." Note that while meaningful messages have a certain redundancy (reduced Information density), that is a necessary but not sufficient condition. Something with the "right" amount of redundancy still might not be meaningful. So perhaps in the diagram, the area under the curve should be shaded, to indicate the range of meaningfulness possible in an Information stream with certain statistics.

The sweet spot for meaning lies somewhere in between such extremes. We find messages meaningful when we partly anticipate what's coming, according to similar things we've known, but can encounter something unexpected within this framework of the familiar.

Partial predictability amounts to a certain degree of redundancy, which is only to say that the quantity of Information delivered in such a message is reduced somewhat from the maximum possible. Think along with me, of where Beethoven takes us with his relentlessly redundant, yet always developing, four note motif in the Fifth. Now, there's epic meaning making!

Astonishingly, we can begin to see particular ways by which meaning making actually happens in artificial cognitive systems, as well as in neurological systems, like our own minds. In recent decades, with the theory and practice of artificial neural networks and their explosive new applications in so-called "deep learning" systems, we have a laboratory for studying how cognitive processes can emerge from networks of interconnected responsive nodes, like neurons, simulated or biological.

When an artificial neural network is trained on a data set to learn a task like facial recognition, it discovers the redundancies — the common patterns — in the data. All faces have typical elements, such as eyes, nose and mouth, facts which enable basic recognition that an image is of a face at all. But meaningful distinctions are found among individual variations in such expected features.

These computer systems are able to make meaningful inferences based on patterns they have found in their training data. So taking a dive toward understanding how these things work could offer insights into ways that minds create meaning out of information.

Though emergent intelligence may seem mysterious, even mystical, in fact amazing behaviors can arise from rather simple rules among connected responsive parts. Strikingly vivid examples appear in schools of fish or murmurations of birds. Their breathtaking display of coordination as a whole, as if some guiding direction were at play, instead demonstrably emerges just through slight navigational adjustments amongst nearest neighbors.

But neural networks are not schools of fish. Neurons are more richly connected beyond nearest neighbors, and minds are capable of learning.

By building model networks and then observing the changes that occur among a the network's connections as learning happens, we can see precisely how what is learned becomes embodied. Stated simply, learning becomes manifest as changes in connections among nodes in a network.

In an artificial system, such a change can be simply adjusting a number, as we shall see in our simple model below. In the synapses between living nerve cells, these changes are embodied in shifting concentrations of neurotransmitters, in the case of short term memory. In long term memory and learning, connections are strengthened by actually growing new synapses between neurons.[2]

Now in real living intelligences, changes in connections can be as richly complex as the living nodes themselves, be they neurons, slime mold cells, trees of a forest, birds in murmuration, or bees in a hive. Still it is truly surprising the degree to which intelligent-seeming behavior emerges even among simple nodes making simple adjustments. For the sake of discussion, let's keep it simple for now.

A simple model

So imagine with me a model which has nodes characterized by just one property, a single number representing something. Anything really, but let's just call it an "activation level." Our model will want to have many of these nodes, all tallied and indexed, with many connections among them. These connections can also be characterized by just a single number representing some other property. This one is usually called its "weight." That's it. Pretty simple really.

How is this thing to work, then? Well, each connection is, of course, between two nodes, as in the figure. There is a directionality here, and we speak of the connection as being "from *i* to *j*," (or the other way around). We can think of a connection as propagating the activation level from its incoming node to its next, according to its weight.

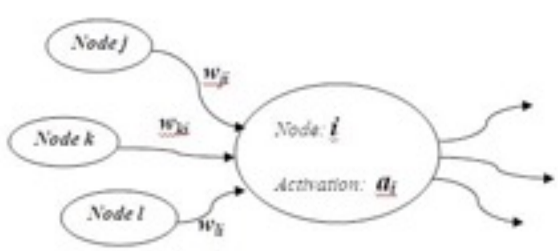

Figure 3: Elements of a simple model artificial neural network. Nodes are ovals. Connections are arrows. Nodes are indexed, and have an activation property given by the number "a_i". for node number "i". Connections have weights, labeled with the indices of the two nodes it connects. Thus, "w_{ji}" is the weight of the connection from node "j" to node "i". This connection propagates the activation of node "j" on downstream to node "i".

We want to present some input to a front end, and observe some output at the back, so let's designate some nodes to be input and some output. Imagine input activations propagating through

[2] Kandel, Eric R., "In Search of Memory, the emergence of a new science of mind," W.W.Norton, 2006.

the network along connections until the spreading wave reaches some outputs. To propagate an activation from one node to the next we can simply multiply by the weight of the connection. The receiving node in turn can adjust its own activation level by scaling weighted activations from all the connections coming into it. That node's newly adjusted activation level can then propagate into all the outgoing connections from it. And so on.

That such a relatively uncomplicated model proves to be capable of quite a lot is remarkable. The key lies in what we propose for its dynamics. Put another way, how shall it learn? That in turn is to ask, how shall its connection weights change, and in response to what? If we say our machine should learn from experience, then let us engineer things so that whenever it "gets it right," that is when a pattern at its outputs corresponds correctly to a pattern presented to its inputs, then the system will be "rewarded" by incrementing the weights on all the connections which contributed to the right answer. Conversely when wrong, it is "reproached," and the contributing weights are decremented. After each teaching cycle, increments or decrements are processed in a wave that moves back upstream through the web, changing weights of connections as it goes, in an operation called "back propagation." That's how our model can learn. The algorithms described as "deep learning" are but refinements and elaborations of models like this little one, with multitudinous nodes and applied to massive bodies of information.

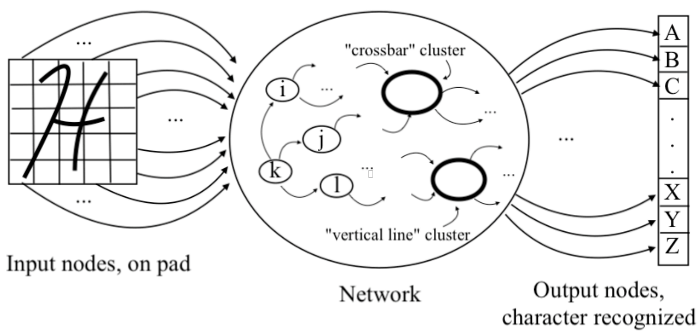

Figure 4: Input pixels from a pad come into the network, and outputs are wired to letters of the alphabet it is to recognize. Ellipses indicate many more connected nodes. After training, strongly connected sub-clusters of nodes, suggested by dark ovals, tend to form, like "ganglia" representing patterns in the data.

To see how an abstract model like this can really work, let's think about a concrete example: imagine we want our machine to recognize hand written letters drawn on a pad. Input nodes could be squares on the pad's grid, and outputs could be one for each character in the alphabet. We add any number of internal nodes and hook them up with connections from the inputs, among each other, and ultimately to the outputs, initialized with some default set of weights, mostly the same value, as a sort of *tabula rasa*.

We start the teaching cycles by setting input nodes to activations corresponding to what was drawn on the pad, and propagate activations forward through the network until some changes appear at the outputs. Then we have to decide if the answer is "right." An easy way is to take the most highly activated output, and call it the system's choice of letter. Back propagate the appropriate reward or reproach upstream into the network, rinse and repeat. Many times. Little by little, it turns out that a system this simple makes coherent changes in its connections, and can eventually do a creditable job in performing hand writing recognition.

We can truly say that our toy machine has in some sense learned what shapes mean. Its operation results, after all, in an interpretation, in this case a simple categorization, of unspecified patterns at its inputs. It ascribes a meaning, right or wrong, to what it has seen. Our question now becomes, how do changes in connection weights embody how a machine has learned what data means?

I described the collective changes as coherent, and it is the nature of that coherence that we want to characterize. When these networks become large, as they naturally do, coherence appears in statistical measures we can define over the set of connections, analogous to changes in Entropy. Think of order condensing, as it were, out of chaos; the relations between thermodynamics and Information theory are so intimate.

What we observe is that connections evolve through learning so that nodes tend to become clustered into interconnected subgroups, mutually self-excitatory subsets, like neurons grouped into ganglia. These condensations properly represent elements of patterns a machine has learned to find in the training data, so that when a cluster becomes generally activated it means that the element it represents has been recognized. For example, a cluster might tend to become activated when there are vertical lines present in the input figure, or when there is a central cross-bar, such as in letters like ‘A’, ‘E’ or ‘H’. Such condensations into interconnectedness, the formerly featureless *tabula rasa* now having differentiated into tight subsets, constitute reductions in Entropy measured over the weights of connections, which, given a definition like the one I propose in the next section, we can calculate.

Just as the golden sun’s energy streams through our world driving the exquisite reductions of entropy into the rampant beauty of life itself, and as the shapes of this blossoming world can be danced into what a hive knows, so a stream of information flowing through a cognitive network, through our toy model even as through the mind of a child, drives a condensation into the relative order of understanding, out of the chaos of confusion. And that order embodies the meaning which we ascribe to experience.

The meaning of Meaning

What I am proposing therefore is a statistical definition of what is to be meant by Meaning, analogous to Shannon’s definition of Information and Boltzmann’s of Entropy. But instead of being defined with respect to a one dimensional thing, like Shannon’s message on a wire, Meaning is defined over the two dimensional matrix (column *i* by row *j*) of the connection weights in a cognitive network which has learned something through its experience with a body of Information.[3]

Equipped with a definition like this, we can imagine directly measuring the making of Meaning in a mind, broadly construed as any cognitive network. We can ascribe a measurable, physical value to Meanings made in the manner sketched out in our model. It measures the amount of order that the cognitive widget has managed to find in patterns within the Information it has been trained on.

Moreover, this description of where the meaningful nuggets reside in a trained network could enable us to isolate the clusters and give them names. This has the potential of giving a deep learning system the explanatory power that comes with self-reflection. For example, when you identify an image of your mother in a batch of pictures, you do so in a flash, as the inputs propagate through your neural networks sparking recognition. When I ask you how you did that, you will construct a tale perhaps citing such elements as her hair style, the arch of her eyebrow, the dimple in her smile, or other salient elements in the picture. But that's of course not how you did it, really. Now, when a deep learning algorithm makes a decision about hiring or insurance or financing, or denying parole, and you ask why, there can be no answer. Even if the algorithms weren't proprietary, they have no ability to reflect on what happened. But if learned pattern elements are given names (such as "crossbar" or "vertical line" in our toy example), then a system could detect which such clusters became highly activated in the course of the decision, then it could possibly offer a constructed explanation of its reasoning (such as, "I called it the letter 'H' because I saw a 'crossbar' between two 'vertical lines'.")

The subliminal secret here is that there is power in being able to associate a precise numerical value with Meaning. For while what we have defined here is a mere physical value, let us perform the trick of eliding our meaning of "value" into its sense as used in economics. Because there is no reason why our statistical number, our measure of order in an understanding of information, shouldn't be closely associated with our human values which, as *homo economicus*, we express in the marketplace. As humans at our most humane, we do undeniably value our own understanding; why else do we work so hard to achieve it?

There are compelling reasons why these values emphatically *should* be closely related. Indeed, what would it mean if Meaning had not only a value, but a price? It is common now to read about the "Information economy." What would a Meaning economy look like?

Consider the economic and social costs represented by a surge in what are called "deaths of despair," of people, for whatever reasons, simply unable to find meaning in being alive. Think of

---

[3] The mathematical expression and details of this definition are presented elsewhere, and omitted here. See http://robduisberg.org/pdfs/statMechOfMeaning.pdf, In applying such models to complex living systems and ecologies, the trick is that nodes and connections can be characterized by numerous parameters, instead of just one. This adds another layer of summation to the calculations, but they remain the same in principle.

the costs of declining communities in which people experience a lack of meaningful connectedness. And contemplate our tragic, rapacious plunder of natural resources, driven by an economics of solely pecuniary values, where there is no competing value in the marketplace representing such a thing as the value of an ecosystem simply being intact, which is the value of the Meaning measured in the richly diverse connections inherent in that system as a whole.

Perhaps it was ever so with our species. But I would prefer attempting to envision a more hopeful path, since we do excel at being so adaptive. I try to think of ways in which social incentives toward stewardship and husbandry could work, replacing our more perverse incentives to merely exploit. I imagine a system of financial incentives wherein the measurable value which derives from a system being Meaningfully whole competes directly in the marketplace with the short term value to be gained from destructive extraction. Economic activity follows incentives, so this could be promising.

From the bubble of our American culture we tend to forget that some other cultures do manage to value things besides money. An example presents itself to me in the tidy little socialist society of Austria, where contemporary artists, musicians of all sorts, singers of Lieder, dancers, public intellectuals are able to actually earn reasonable livings practicing their arts and scholarship. A lower middle class resident of Vienna can regularly attend the Vienna State Opera for just a few Euros, a price orders of magnitude less than it costs here, where such enrichment is reserved thereby for only upper classes. Viennese society at large values its wonderfully rich cultural heritage and vibrant contemporary creativity, so the state expresses this valuation through generous support of cultural organizations at all levels. State administration enables the assessing of costs and securing financing to support a decided cultural value. Similarly, we are now learning how a state actor can administer the measurement of carbon production, and associate appropriate costs thereto, since a free market fails to assess the real costs onto a producer whom we thus enable to ignore his degradation of the commons with impunity. There is nothing politically wrong with the state simply administering such choices. It is only a matter of our choosing what has value for us, and managing that value through pricing in the marketplace.

As we face impending changes resulting from a history of exclusively materialistic choices, we may find ourselves beginning to value more what is truly meaningful to us in our lives, rather than squandering it all for that hollow feeling of just having some more stuff. Value choices can change the culture.

Were cultures to evolve, as they adapt to Anthropocene pressures, toward embracing a solid set of Meaningful values, it would surely be transformative. We would be guided by more expansive aspirations. Extinctions would be experienced as the profound and irrecoverable loss of real value that they most surely are! We might move toward living more in balance with our world. We might even come to know what we really mean, when we talk about how we might meaningfully live our one precious life.