Recherches sémiotiques
Semiotic Inquiry

# Siri Fails the Turing Test: Computation, Biosemiosis, and Artificial Life

Victoria N. Alexander

Volume 39, numéro 1-2, 2019

Biosémiotique. Tome 2
Biosemiotics. Tome 2

URI : https://id.erudit.org/iderudit/1076234ar
DOI : https://doi.org/10.7202/1076234ar

Aller au sommaire du numéro

Éditeur(s)

Association canadienne de sémiotique / Canadian Semiotic Association

ISSN

0229-8651 (imprimé)
1923-9920 (numérique)



Citer cet article

Alexander, V. N. (2019). Siri Fails the Turing Test: Computation, Biosemiosis, and Artificial Life. *Recherches sémiotiques / Semiotic Inquiry*, *39*(1-2), 231–249. https://doi.org/10.7202/1076234ar

Résumé de l'article

Les concepteurs d'intelligence artificielle (IA) tentent d'imiter les aptitudes des cerveaux humains au moyen de réseaux neuronaux qui apprennent par eux-mêmes grâce à des processus de sélection. Mais même après des décennies d'efforts, l'IA n'en continue pas moins d'échouer le test de Turing. Alors que des ordinateurs utilisent des codes et développent des *algorithmes* hors contexte, les cellules vivantes utilisent des *signes* et auto-organisent des *habitudes sémiotiques* de manière contextualisée. Je soutiens que cette différence s'explique, en partie, par les activités collectives des neurones biologiques qui produisent des ondes, lesquelles contraignent l'activité neuronale. Il appert que les motifs ondulatoires fonctionnent comme des contextes, et qu'ils informent le contenu des connexions locales. Au moment de sa mort, Alan Turing l'inventeur original de l'IA, s'intéressait au rôle organisateur des motifs ondulatoires sur le développement biologique. S'il avait vécu et poursuivi ses travaux, il aurait peut-être réorienté la recherche sur l'IA, laquelle est devenue un outil servant simplement la régularisation et la création de stéréotypes, et non un outil de pensée.

# Siri Fails the Turing Test : Computation, Biosemiosis, and Artificial Life

Victoria N. Alexander
Dactyl Foundation

## Introduction

The ubiquity of technologies using artificial intelligence (AI – Google learning algorithms, Apple smart phones and weaponized robots – should give us pause. What *is* intelligence? What might be the difference, if any, between intelligence in machines and organisms? Both can obtain goals, set either by evolution or design. Machines can be programmed to perform computations, seek objects, read signs, and even preserve themselves. But do organisms and machines use different methods for learning, remembering and interpreting in order to perform these intelligent actions?

Alan Turing is celebrated for designing the first universal computers and decrypting the German Enigma code during World War II. He is also known for the hubris of believing that AI could eventually develop chat bots that could pass for humans in limited exchanges, otherwise known as the Turing Test. Few are aware that Turing was a pioneer in the field of Artificial Life, working out equations to describe the process underlying the formation of dappled animal fur patterns, root growth, and embryonic differentiation. Turing intended to apply this research to the design of a self-learning network computer. He died of cyanide poisoning after writing his second preliminary paper on the subject. These long-neglected papers reveal how organisms *construct themselves* as universal Turing machines. His next step would have been to understand how organisms *learned* by the same methods.

I approach these questions through the lenses of biosemiotics, emergence theory and second-order cybernetics, newly developing fields indebted to Turing. Biosemioticians claim sign-use emerged with the first life forms, which are able to respond to gradients and transform them in ways that may not lead directly to self-preservation, but which lead to other gradients, and perhaps other gradients, which ultimately contribute to self-preservation. Thus even for the most primitive life forms, gradients can function as *signs* of what is needed. Biological computation, the rule-bound transformation of matter and energy, uses *signs* and runs *semiotic habits* that transform gradients, while machine computation uses codes and runs *algorithms* that transform 1s and 0s. Both signs and codes are gradients/differences that *stand for* certain objectives or functions (Alexander 2013). But signs are both more flexible and more robust than codes. Additionally, biological computation is structured by the self-organizing laws of physics and by external selection, whereas all artificial computation (with rare experimental exceptions) depends entirely upon external selection for its organizational structure. And finally, whereas computer algorithms are normally activated by external command, each step in a semiotic habit helps reset the cycle so the habit can reactivate automatically.

That is my thesis stuffed into a carry-on. Now my task will be to unpack it.

Turing's death at forty-one was ruled a suicide, but he may have accidentally produced cyanide gas in his small, untidy en suite lab, which he called the "nightmare room". Turing had lost security clearance as the UK entered the cold war, and paranoid bureaucracies began to enforce mechanized procedures for managing human affairs. The authorities (ironically so-called) jumped to the conclusion that because Turing was gay and he had got caught, the supposed shame had depressed him. His death is doubly tragic when we realize that Turing was on the brink of a great discovery about the fluid nature of human intelligence.

Had Turing lived perhaps our understanding of human learning would not have become so distorted by inadequate machine computing metaphors, with the field of neuroscience liberally borrowing language from AI. Perhaps public education would not be approached as if children can be programmed to know. Although Turing had initially helped to promote such misconceptions, his thinking had evolved. This is about a man who was not just a computer.

## What is Computation?

Before machine computers, there were "human computers" who performed logical transformations using memorized set procedures : multiply A times B, write the result in line C. A multiplication table is essentially a computer program : connect vertical column 4 to horizontal row 6 and get 24. Turing called computer programs "instruction tables", and the first computers were the physical embodiment of a look-up table, repurposed telephone switchboard equipment with plugs in rows and columns that could be connected by cables in different ways to do different procedures.

Turing's work in decryption follows similar computational logic. Every letter represents a different letter according to a rule, such as "move two places to the left in the alphabet"; accordingly, A = C, B = D.... Z = B. The Enigma code used much more complicated multi-step transformation rules and the outcome of one step was fed-back into the operation to get the next step. To solve this, Turing devised a system of plugs and interconnecting gears that turned each other according to set rules.

Thus we can say that mechanical computer memory is physically embodied, or "programmed", into its gears or look-up table structure. In digital computers the data is encoded in any kind of difference, e.g., 1s and 0s. Switches or gates, replaced cables and directed how encoded differences are transformed.

*Task-specific* mechanical computers existed long before Turing's *universal* machine. Made between 205-100 BC, the Antikythera Mechanism had multiple interacting gears representing the movement of the stars and solar system. The ancient Aztec calendar is a computational program in abstract form (look-up table) rather than physical form (gears). All programs describe the rule-bound transformations of matter/energy. We can say, therefore, that the equations of physicists are programs and nature does computation. Turing was interested in the fact that, while a machine computer requires a person to record the instruction tables, a human computer can *learn* to do procedures through experience.

## Learning Networks

When Turing started working on a self-learning network computer, he assumed humans make guesses when they do not know a procedure for solving a problem. In 1948 in "Intelligent Machinery", he claims, "training a human child depends largely on a system of rewards and punishments" for good and bad guesses respectively (Copeland 2004 : 425). Turing designed a chess-playing program with optional moves that could be tried at random. If a move ultimately led to failure, it would not be reinforced. Turing's neural network was designed to start out unorganized and become organized with appropriate "interference", mimicking eduction.

Similar kinds of *connectionist* approaches are used today in most self-learning algorithms. Depending on the kind of data that flows through them, connections are strengthened or weakened, affecting how nodes are switched on or off. Connection gates become biased with use. In this way, the "instruction table" is embodied in the connections and nodes as they are altered by reward/punishment feedback.

Feedback can be administered by a programmer, who adjusts the biases so that the network develops the desired output (taking the opponent's king), or the biases can be adjusted by sub-algorithms that filter at the nodes, or the feedback can come from crowd sourcing. Internet users train algorithms all the time. Reward and punishment eliminates the wrong algorithm and propagates the right one, according to a desired output. Although some AI developers call this approach "*self*-organizing", this is an incorrect use of the term. As Turing noted, it still requires "interference" from the outside.

The newest phase of AI, which arrived with the 21st century, boasts of "unsupervised" learning. A visual recognition network is exposed to millions of random images. Even though no images are labeled – the desired output is not defined – the network eventually detects recurring patterns. It creates a generalized representation of commons patterns, some of which may belong, for example, to cat faces. (If the fruit fly is the object of contemplation for geneticists, cats are it for programmers, whose datasource is the Internet). Programmers do not tell the network what to find; it just detects common patterns and outputs a generalized picture of the pattern. If the programmers recognize the output as a cat face, that network can be used to find cat faces, even though the programmer does not know what criteria were used to develop the generalization. The pathways and connections will have acquired biases in unknown ways at various levels that may be hundreds of levels deep (Le *et al.* 2013).

These new unsupervised "neural" networks are not so dissimilar to Turing's 1948 notion of a self-learning network. The main difference is the point at which the programmer interferes, during the training process to target a pre-specified pattern or after the network as detected a pattern that is of interest to the programmer. The unit of selection here is the entire network, not individual connections within the network. The trained network has become an "instruction table" for identifying the patterns it has detected.

Unsupervised learning is how most non-human animals learn. They are not taught (not intentionally rewarded and punished). Instead, they imitate what they detect. It's the monkey see, monkey do approach. If they see the correct behavior frequently, they will end up doing the correct behavior.

## Human Memory Storage

Humans can learn by this "connectionist" process. Rote learning is like being programed. Repeating the right procedure over and over until it becomes a habit actually changes neuronal connections. The learned behavior can be recalled automatically, without thinking, as if it were machine intelligence. Neurons that fire together wire together, as Donald Hebb so famously noted. Learning by rote, strengthening connections over time, is statistical in nature. What happens the *most* gets selected. Repetition is one way humans learn, but not the only.

Before machine computers, before written language, people had another kind of program for memory storage, namely poetic narrative. An ancient astronomer without a practical way to record the movement of the stars and planets in a look-up table or a physical model might put the information into an oral metaphorical narrative with superhumans symbolizing the calendar.

A paragraph of a poem is a “stanza”, “room” in Italian. Ancient poets associated each line in a stanza with the random objects in a room of a great palace. To recall a long poem, they imagined themselves walking from room to room, looking at the objects. Similarly, tone, rhyme and rhythm are used as mnemonic devices. Synesthetes can remember long arbitrary lists by associating numbers or words with colors, textures or shapes. One can recall something better, in the right order, if it is associated with something arbitrarily similar or arbitrarily nearby. This way of learning doesn’t follow the logic that Turing initially thought would be important for imitating human learning.

According to American semiotician C. S. Peirce, natural symbolic language emerged from icons (A is a sign of B by virtue of similarity to B) and indices (C is a sign of D by virtue of contiguity with D) – or metaphor and metonymy. In biolgoical systems, cell receptors have a *similar* shape to molecules with which they are able to interact. Chemicals that are by-products of any process are *contiguous* with that process, that is, they are usually found in the vicinity of that process, as smoke is an index of fire. Things, like molecules, that are associated with other things, can form a chain of reactions, and if this chain of reactions is autocatalytic, recreating the conditions that allow it to continue, it creates a habit cycle. If this habit contributes to the survival or functioning of the organism in which it exists, then we can say that it serves a purpose for the organism. Each icon or index, the individual links in this reaction chain, become an arbitrary symbol of the habit’s function.

## Biosemiotics

Whereas, codes/symbols are decrypted in a strict one-to-one manner following a predetermined arbitrary convention, an icon and index might have a different rule attached to each use, depending on context. In other words an icon might be similar to something else and an index might be associated with more than one outcome. Let’s try to imagine what is going on inside a neuron. This Rube Goldberg-inspired machine (Figure 1 below) for loading a toothbrush can be compared to a semiotic habit involving transformations of signs that control neuron activity. If the pin is pulled A, the ball pops B, and rolls down a channel C, etc. Figure 2 shows a Rube-Goldberg-like semiotic habit of a neuron involving a dopamine pathway. (We also note that each neuron has multiple such habits and is also a node in semiotic habits among several neurons and groups of neurons which further constrain firing patterns. It's complex.) I use a Rube Goldberg machine because it illustrates how the process of evolution has repurposed old tools and cobbled steps together. Each step in these processes is determined by the laws of physics, but different materials/devices might have been used at each step; some might have been skipped over (see Faria 2008). Each step is only arbitrarily related to the outcome, that is, each step has been selectively retained in this machine only as a means, any means, to an end, which is the continuation of the cycle.

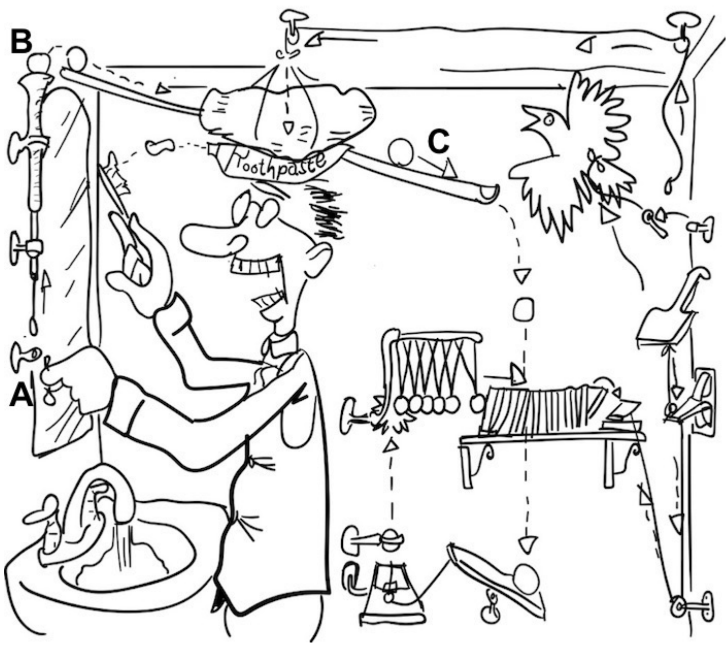


Fig. 1 Rube Goldberg-Inspired Machine as a Semiotic Pathway.

Semiotic habits are more flexible than algorithms that use codes. For instance, the bird in Figure 1 could be replaced by something similar, like a frog. Anything that might jump or fly when startled could work. Organisms don't try out new signs at random like Turing's self-learning computer program for chess. Biosemiotic guesses are more like hypotheses based on experience. A new sign in this habit must be similar/contiguous to the old one. If the substitution fits, the semiotic habit might continue or adapt. In the neuron, similar enzymes might be used at different times to make the signal pathway work in variety of circumstances or to make it follow a different route.

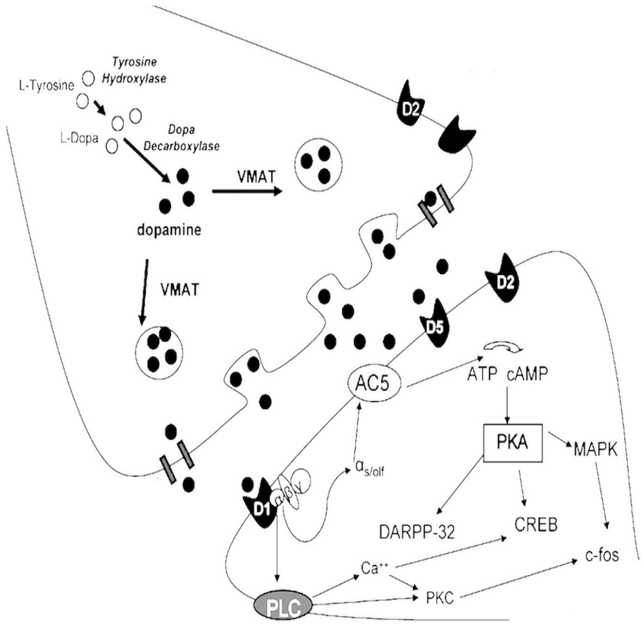


Fig. 2 - Neural dopamine pathway as a Rube Goldberg-like semiotic pathway.

Also, semiotic habits (within and among cells/organs) reset themselves, so we must imagine a Rube Goldberg machine designed by M. C. Escher. Such machines are *autocatalytic* reactions whose by-products restart the machine.

## Artificial Life

To design a self-learning network computer, Turing felt needed to know how the brain's connections and switches developed. In the 1950s, he started to work on what we now call Artificial Life, a field tangent to AI. In "The Chemical Basis of Morphogenesis" (1952) and the "A Diffusion Reaction Theory of Morphogenesis in Plants" (Turing & Wardlaw 1952), Turing explores how an organism develops, how it does computation, how it transforms materials into other materials by set procedures.

Having studied C. H. Waddington's *Organisers and Genes* (1940), Turing had learned that genes are not equivalent to instruction tables for development. Waddington noted that the physical constraints of liquid crystals in random motion allowed them to *fall* into a pattern;

thus no instructions would be needed to *direct* similar kinds of biological organization. Turing had also studied D'arcy Wentworth Thompson's *On Growth and Form* (1942). Thompson held that similarities between different species did not necessarily indicate common descent or common genes, but might merely indicate there are similar physical laws constraining development.

In the first paper, Turing offers simple equations describing chemical reaction-diffusion processes that can artificially create (without genes) animal skin-like patterns (Figure 3) and cause cell differentiation. Turning discovered that the physical structure of the gene does not specify the elaborate, complex structure of the organism. The genes mainly provide the templates for making the materials, in the right order and in the right amounts. But the genes do not contain the instructions for how to put the materials together (see Keller 2002 for a history of the understanding of gene action). The laws of physics act as the transformation rules that help self-organize the gene-produced materials.

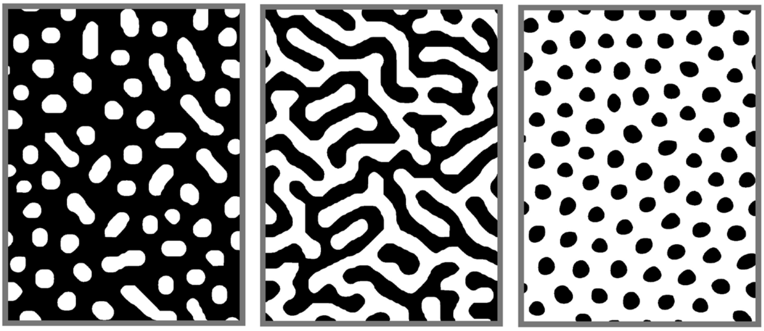

Fig. 3 - Computer-generated Turing Patterns.

As a computer programmer, Turing would have had great admiration for Nature's ingenuity and economy. She did not have to physically record the procedure for development in the DNA. Instead, Nature availed herself of programs that already exist in the world. We can say that zebra stripes and giraffe patterns already existed as patterns before life on Earth, before genes. Many forms in nature are the result of the constraints of physical forces, not the direct outcome of genetic control or gene selection.

## Self-Organization = No External Selection

As noted above, before beginning his work on morphogenesis, Turing had assumed organisms are programed by external interference, be it from computer scientists, school teachers, or the process of natural selection. He discovered nature sometimes works without interference.

Waddington and Thompson continued the work of pre-Darwin theorists who sought to understand the inherent, universal laws biological form, not the externally-driven, particular effects of selection (whether Lamarckian or Darwinian). By the late 1940s and early 1950s, Darwin's theory had almost completely eclipsed pre-Darwinian theories, which were then derided or unknown. To be interested preDarwinian theories in England in 1952 was tantamount to heresy. As an outsider to biology, Turing was probably less bothered by what the reigning conventions dictated. He was poised to bring about the Kuhnian paradigm shift that, after his death, did not begin for another thirty years.

The natural selection for reproductive fitness explains why organisms are so well suited to their environments, but Darwinism is not a theory morphogenesis. Where does form come from?

Turing's reaction-diffusion equations describe autocatalytic processes that are initiated by local fluctuations that become amplified : initially one type of chemical reaction produces some by-product, but when this reaction runs out of materials, another chemical reaction takes over which uses the by-product; repeat. The result is a standing wave pattern. The second paper dealt with how these patterns, in turn, interact non-linearly to cause further differentiation. The salient point is that the unorganized chemical soup spontaneously forms patterns by amplifying small fluctuations in chemical concentrations. Nothing external is driving the changes. As chemicals diffuse across the cells, they leave behind chemical residues, which makes some cells different from others. These chemical differences, in turn, trigger genes within the cell to make different kinds of proteins. Turing proposed that reaction-diffusion is the mechanism whereby an embryo starts out as a collection of perfectly identical cells and then automagically differentiates into specialized cells, heart cells, lungs cells, etc. In the 1950s, scientists offering such theories would either be ignored or come under serious fire from orthodox neoDarwinists who wanted to think that the natural selection of genes and the natural selection of genes alone was responsible for shaping organisms — not physics.

## Self-Organization as Amplified Chance

How could Turing be open to self-organization, which is so unlike the programming he had so far conceived? In 1932, when Turing was twenty, he was thinking about his first love, Christopher Morcom, who had died two years earlier. Turing wrote "Nature of Spirit", influenced by J. M. E. McTaggart, who, though an atheist, believed the human "spirit" preceded and survived the body. McTaggart was committed to the idea that everything, including matter, is fundamentally "spirit" (by which he meant, insofar as I can tell, something comparable perhaps to C. S.

Peirce's notion of Firstness). In the essay, Turing reflects on the way LaPlace's deterministic universe had been overthrown by the discovery of quantum indeterminacy. He goes on to speculate that if nothing in the universe is predetermined, perhaps

> We have a will which is able to determine the actions of the atoms probably in a small portion of the brain, or possibly all over it. The rest of the body acts so as to *amplify* this. There is now the question which must be answered as to how the action of other atoms of the universe are regulated. Probably by the same law and simply by the remote effects of spirit [by which Turing seems to mean *quantum indeterminacy*, as the "new" monism to succeed McTaggart's monism], but since they have no *amplifying apparatus* they seem to be regulated by pure chance (2-3, emphasis added).

It may have been significant for his discovery of morphogenesis by self-organization that Turing had once imagined organisms have an "apparatus" that "amplifies" small fluctuations. Although he did not locate the supposed apparatus, he did find that fluctuations could be amplified to significant effect by physical laws. The idea fits well with McTaggart's idealism in that the "apparatus" turns out to be the immaterial laws of physics. Although this may not be quite what the grieving young Turing had imagined for his friend's immortal spirit, the self-organizing laws of reaction-diffusion processes are "programs" that exist abstractly prior to and survive the body. Scientific advances don't always proceed logically. Sometimes an emotional need opens up the mind to new ideas.

## Emergent Brain Patterns

To recap, Turing had found the complete instructions for organizing cells are not in the genes, but are also encoded in the laws of physics. Is there a program for organizing the actions of neurons?

This question has long been known as the "binding problem" in neuroscience (see Raffone & van Leeuwen 2001). Because we know that neurosurgeons probing around in grey matter may make the patient hallucinate the smell of bacon or the sound of an elevator bell, it is probably not too incorrect to say that raw data seems to be habituated as connectivity among neurons. Connections also appear to serve multiple purposes, habituating different data among the same neurons with different connections. But it is less well understood how sense data is bound or organized into coherent thoughts, later to be recalled or used in different contexts.

In thorough reviews of the literature, Kelso *et al.* (1991), Uhlhaas *et al.* (2009) and De Assis (2015) report that many neuroscientists understand the mechanisms underlying working memory and attention in terms of emergent brain waves that synchronize distant neurons. Synchronization creates virtual neuronal assemblies, without creating permanent circuits (De Assis 2015; see also Freeman 2000; Postle 2006). Generally, neuroscience literature (see Basar *et al.* 2001) has come to associate beta waves with perception, attention, motor control, sensory gating, top-down control; alpha waves with attentional process and consciousness; theta waves in the hippocampus with memory and spatial navigation (see Zhang & Jacobs 2015), and gamma waves with perception, attention, memory, consciousness, synaptic plasticity, and motor control. It appears that waves may provide "the 'contexts' for the 'content' carried by networks of principal cells" providing "the precise temporal structure necessary for ensembles of neurons to perform specific functions, including sensory binding and memory formation" (Buzsáki & Chrobak 1995). The systemic reorganization of neural activity, key to insight, is associated with gamma waves and the anterior superior temporal gyrus, which, significantly perhaps, is linked to language use, understanding of literary themes and metaphors, and getting jokes, which require making distant connections (Jung-Beeman *et al.* 2004).

In addition to acting as a binding mechanism, emergent wave patterns may also define what data gets attention, that is, consciousness (see Thompson & Varela 2001), which in turn affects sensory processing. For years neuroscientists had referred to attention as a "spotlight" shining on the right file folder as needed, but now it is clear that attention is not neutral. It anticipates, fills in details, sharpens and augments data (Uhlhaas *et al.* 2009; Gilbert & Sigman 2007).

All this contradicts 20th century notion that brainwaves are like the sound an engine makes and do not contribute to the operation of the engine.

In this paper, then, I work under the assumption that individual neurons both produce the higher level waves and are controlled by them. Such reciprocity does not exist in current AI computation, which does not produce contextualizing emergent constraints. Even with the latest celebrated update (Levis-Kraus 2016), Google Translate is still bad with puns, jokes and poetry which depend on context. The application tries to deal with context by measuring the statistical probability that words will appear near each other based on past samples. The sentences : "Time flies like an arrow. Fruit flies like a banana" translated into traditional Chinese, produces, "時間就像一個箭頭。 果蠅像香蕉", which omits the all-important first "flies" (飛) because in Chinese, the combination of "time" with "flies" is uncommon. The translation is equivalent to "Time is like an arrow. Drosophila like bananas". Machines just don't have a sense of humor. To design computers that can get jokes, one might need a more fluid medium for traveling gamma waves to emerge. (In the film Ex Machina, the robot's brain is a gel.) Atomic switch networks seem promising; they have been used to create emergent patterns that imitate simple natural systems (Stieg *et al.* 2014). Experimental chemical reaction-diffusion computers have been around for more than a decade (Adamatsky *et al.* 2005), but although they create emergent patterns, they do away more permanent connections. Our brains seem to use both.

It's no wonder AI falls short. The brain is the most complex object known to man. We are closer to understanding the most distant event in the universe, the Big Bang, than we are to understanding our own brains. Current methods for measuring brain activity, such as with EEG net caps, resembling Dr Frankenstein technology, give only a limited picture of electrical oscillations. Although models can be cobbled together from EEG, fMRI, PET, SQUIDs or MEG measurements, we still cannot see three-dimensional topological waves in real time. Even if we could look at three-dimensional picture, the emergent patterns would be difficult to see because neurons are not all connected side by side; the dendrite of one neuron may connect to an axon at a relatively long distance. Nevertheless, we may assume emergent behavior in the brain does occur (see Chialvo 2010). We see emergence in all living and many non-living systems : slime mold organization, video feedback, stock markets, shimmering bees and butterfly wing patterns, to name a few examples.

## Algorithms, Instruction Tables, Habits, Epsilon-Machines

All self-organizing systems follow universal rules that can apply to any system of a certain complexity, regardless of its actual materials/members. Thus, we can try to visualize emergent brain patterns by looking at a flock of starlings. Just as neuroscientists investigate the mechanisms whereby coordination of various neurons occurs, we can investigate the mechanisms whereby coordination of individual starlings occurs.

Starling murmurations can fall into recognizable shapes : funnels, waves, folds, bowls – loosely speaking. Since emergent brain patterns also follow limited variations on themes, they might be reliably reactivated when approximately the same groups of neurons are stimulated as when the memory was habituated. The different shapes and frequencies could trigger different memory structures (Raffone, & van Leeuwen 2001; Tsuda 2001). As the emergent brain waves roll and interact, they affect the local connections associated with information from the senses. We can see that the wave in a flock, as it contracts and diffuses, constrains the behavior of individual starlings, further reinforcing the continuation of the wave pattern.

So what might be going on at the level of the individual bird that leads to emergent patterns at the collective level? Let this instruction table (Figure 4) represent the imagined possible states of a single starling; the arrows stand for transformation rules.

Initially a group of birds flying together will behave fairly chaotically. When a flocking pattern emerges, we can guess that the individuals are beginning to cycle through their various possible routines in a more efficient manner than previously, without getting hung up on "if A then B, if B then A", or crashing into other birds. Complexity scientist J. P.

Crutchfield (1994) calls the simplest possible representation of the local level process of a self-organizing system an Epsilon-Machine. Although the bird does not take the same path every time, the Epsilon-Machine is relatively less noisy.

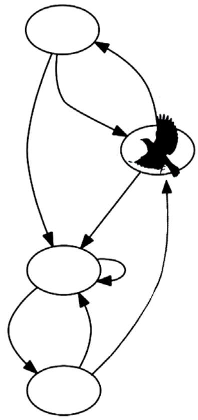

Fig. 4 - Epsilon Machine for Possible State Changes of a Singles Starling in a Flock.

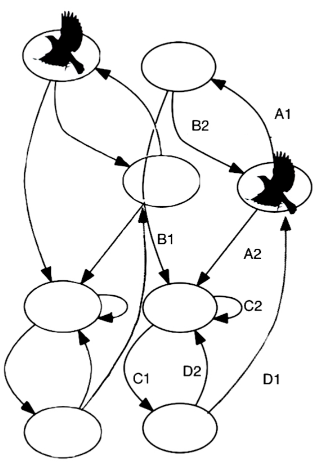


Fig. 5 - Epsilon Machine for Possible Overlapping Routines of Two Starlings in a Flock.

Let this instruction table (Figure 5) represent the possible states for two birds overlaid one upon the other. The birds have similar constraints so they tend to fall into similar routines. Once an efficient routine is discovered by one individual, it tends to pick up neighbors, when, by chance, their cycles coincide. When they all begin to flow together, we have a wave of birds more or less following each other. The wave tends not to lose participants once its gained them. The birds in a flock do not follow each other like clockwork. Our flock behaves organically, not mechanically, not at all like a team of synchronized swimmers nor like a marching band. The individual birds here have to be flexible, if the next bird takes an unexpected option, ready to switch to a different routine almost instantly. Not only would each routine have several options within the routine, but a flock probably has number of routines from which to choose. The different routines might result in different emergent patterns.

Let's consider a much simpler diffusion wave at football stadium. Spectators stand up and raise their arms when their neighbors do, then sit back down. A messy wave travels over the crowd. Heartcells require flexibility and cannot act like a marching band. If one cell's timing is off, the pattern has to be able to flow around the misfiring cell. If we watch just one man in the stadium, we would see that he jumps up and down at fairly regular intervals. Flocking is more complicated. If we could just focus on one bird in a flock and record its motions in space, it would be very difficult to recognize any kind of pattern. Because the bird's instruction tables are flexible, its actions, taken out of context, end up creating non-repeating patterns. A neuroscientist looking at just one neuron won't be able tell that it's following any kind of algorithm.

If we can't figure it out, how do birds learn to flock? Before he started researching reaction-diffusion, Turing had imagined that a self-learning network would be organized by a selection process, by system of rewards and punishments. But I suspect such training would produce a marching band, not an organically flowing flock. If we assume that any of the bird's choices in the routine allow it to avoid crashing into other birds, what would be driving a selection process if all available choices are equally valid? When random interactions among individuals begin to flow, what is happening to define the pathways? This is not the kind of selection that determines connectionist algorithms. It's self-organizing information flow.

Selection for function is not necessary to create a starling's instruction table because it is highly probable that randomly interacting birds will find the most efficient paths. Likewise Turing's chemical reactions flow to the lowest possible energy state as they create higher level patterns. If a bird has taken a path, its neighboring state configurations are timed to allow the flow to occur : one bird gets out of the way, another can get in, the two neighbors become coupled, and will continue to be coupled if nothing knocks them out. If another individual happens to be in a fortuitous state when the efficient flow passes, it will become coupled with the flow.

Biosemiotics offers an explanation for organic flow : the holistic wave patterns emerge when local fluctuations allow stochastic resonance, the similarity and contiguity of possible states, to drive the routine toward efficiency. Artist have long known that they, like birds, must allow their actions to flow.

Waddington had provided Turing with a model of flow in organic development. An epigenetic landscape (Figure 6) is a visual metaphor for the physical forces that guide development beyond the control of genes. The model illustrates how local state changes tend to flow down hill.

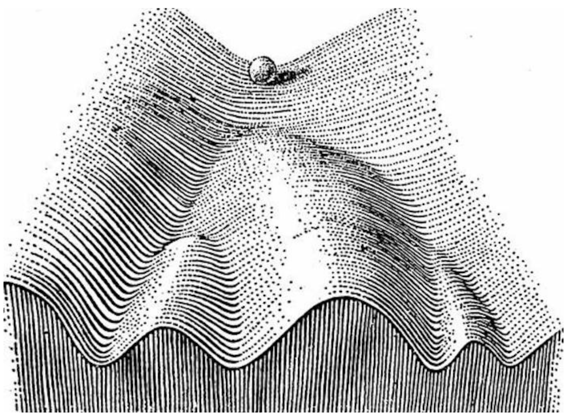

Fig. 6 - C. H. Waddington's epigenetic landscape, from *The Strategy of Genes* (1957). An earlier version of this model appeared on the frontispiece of *Organisers and Genes* (1940).

When a bird is ready to move in one of several directions, or when a neuron is primed to fire, it is in a state of instability, like a ball sitting atop a mountain saddle with various valley features down below. Any slight fluctuation in initial conditions might push it toward one pathway or another from this point of instability. It's very hard to predict exactly which way it will go. These ideas became known as the "catastrophe theory" of early biosemiotician René Thom (see Favareau 2009 : 337-376) and later "chaos theory" (Crutchfield *et al.* 1986). Very small fluctuations in initial conditions can lead to disproportionately large differences in outcomes : the changes are non-linear; trajectories fork. But while unpredictable state changes are going on at the local level, emergent waves travel over the collection of individual nodes constraining them. The Heraclitean never-quite-same wave pattern might amplify fluctuations across an entire neuronal assembly.

Natural selection can then favor one of these emergent patterns over another. Just as natural selection doesn't "see" the genes per se only their outcomes, natural selection probably doesn't see the local state configurations either; it doesn't need to since these are just flowing spontaneously. What it can select are the patterns that emerge from the local interactions (*cf.* Rocha 1998). While local self-organization flows downhill, natural selection is said to drive gene frequencies uphill, that is, to *unlikely* genetic configurations, to high fitness in a fitness landscape (think uncommonly long giraffe necks). Thus self-organization and natural selection work together to develop and keep flexible robust emergent forms that fit well in an environment.

Likewise in the brain, the spontaneous actions of neurons may create emergent waves that constrain attention and thoughts which then may be selectively retained or not.

## AI as Judge

This signal propagation theory of learning, using icons and indices not just symbols, helps explain how people are able to form and use fluid adaptable categories and deal with complex changing environments. Current AI does not imitate the fluid interplay between self-organization and natural selection. Designers are more committed to strictly selectionist, aka connectionist, approaches. Although learning can be accomplished this way, it produces automatons, as does standardized curriculums and relentless testing, reward and punishment.

AI self-learning algorithms regularize, make generalities and stereotype. If they seem to work well pretty often, that's because stereotypes are often true, and they tend to be self-reinforcing. Learning algorithms can make predictions about *groups* as well as any actuarial table, but predicting with certainty *individual* human behavior is still impossible. Due to the non-linearities of the lower level activity, complex emergent phenomena cannot be captured by statistical descriptions, and this is precisely why AI has been disappointing. This is not to say that I think intelligence cannot be created in a lab or factory. Intelligence has emerged from inanimate matter before under certain conditions and there is no reason why it cannot emerge again under certain other conditions. But we cannot trust an algorithm to be more objective than a human semiotic habit. In 2015 Google's algorithm for classifying images labeled black people gorillas (Dougherty 2015).

AI is the bureaucratic approach to decision-making on steroids. We recall here that Alan Turing was punished by the British criminal justice system because he was an "upper class" Brit; he was gay and he had a relationship with a younger "lower class" man. The courts found him guilty of gross indecency. He was forced to take hormones, a kind of chemical castration.

Unfortunately, AI is today being used in U.S. court systems to pass sentences, exacerbating already existing stereotypes. Software is used to predict the likelihood that a criminal will re-offend by categorizing him or her as a type. The result is, as you would expect, blacks get tougher sentences than whites with comparable data points (Angwin *et al.* 2016).

AI is going into other areas it should not go. According to Deputy Director for Digital Innovation at the CIA, Andrew Hallman, the agency now uses Deep Learning AI to better “anticipate the development of social unrest and societal instability…three to five days out” (Konkel 2016). Sounds like the pre-crime unit is up and running.

I’m sorry to see human judgment being taken out of the equation in the criminal justice and national security systems. Computers are good at retrieving specific facts, but in some ways, they can be more prejudiced than the average person. Maybe we will eventually use reaction-diffusion computers to create more humanoid processors, but why would we? We already have lots humans coupled to computers to provide judgement. As many science fiction writers show, consciousness is not something the power elite want in a tool.

We should not pretend that the specific behavior of individuals is predictable. We can’t even predict the weather five days out and never will. At best, like a starling in a flock, we can learn to flow with and/or create an emergent pattern, but we can’t get ahead of it.

## Bibliography

ADAMATZKY, A.; COSTELLO, B.; & ASAI, T. (2005) *Reaction-Diffusion Computers*. Amsterdam : Elsevier.

ALEXANDER, V. N. (2011) *The Biologist’s Mistress : Rethinking Self-Organization in Art, Literature and Nature*. Litchfield Park, AZ : Emergent Publications.

__________. (2013) “Creativity : Self-Referential Mistaking, Not Negating”. *In Biosemiotics* (6)2 : 253-272.

ANGWIN, J.; LARSON, J.; MATTU, S.; & KIRCHNER, L. (2016) “Machine Bias”. *In ProPublica.* http://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (Accessed, February, 6th, 2017)

BASAR, E.; BASAR-EROGLU; C., KARAKAS, S.; & SCHURMANN, M. (2001) “Gamma, Alpha, Delta, and Theta Oscillations Govern Cognitive Processes”. *In International Journal of Psychophysiology* (39)2/3 : 241-248.

BUZASKI, G.; & CHROBAK, J. J. (1995) “Temporal Structure in Spatially Organized Neuronal Ensembles : A Role for Interneuronal Networks”. *In Neurobiology* (5) 4 : 504–510.

COPELAND, J. (2004) *The Essential Turing*. Oxford : Oxford University Press.

CHIALVO, D. (2010) “Emergent Complex Neural Systems”. *In Nature Physics* (6)10 : 744–750.

CRUTCHFIELD, J.; FARMER, J. D.; PACKARD, N.; & SHAW R. (1986) “Chaos”. *In Scientific American* (254)12 : 46-57.

__________. (1994) “The Calculi of Emergence : Computation, Dynamics, and Induction”. *In Physica D* (75)1/3 : 11-54.

DE ASSIS, L. (2015) “Neural Binding, Consciousness, and Mental Disorder : Complexity as a Common Element”. *In Journal for Neurocognitive Research* (57) 3/4 : 110-121.

DOUGHERTY, C. (2015) “Google Photos Mistakenly Labels Black People ‘Gorillas’”.

*New York Times*. July 1. https://bits.blogs.nytimes.com/2015/07/01/google-photos-mistakenly-labels-black-people-gorillas/?_r=0 (Accessed February, 21st, 2017).

FARIA, M. (2008) “Signal Transduction Codes and Cell Fate”. *The Codes of Life*. M. Barberi (Ed.). Amsterdam : Springer : 265-283.

FAVAREAU, D. (2009) *Essential Readings in Biosemiotics : Anthology and Commentary*. Amsterdam : Springer.

FREEMAN, W. J. (2000) How Brains Make up their Minds. New York : Columbia UP.

GILBERT, C.; & SIGMAN, M. (2007) “Brain States : Top-Down Influences in Sensory Processing”. *In Neuron* (54)5 : 677-96.

JUNG-BEEMAN; M., BOWDEN; E.M., HABERMAN; J., FRYMIARE; J. L., ARADMBEL-LIU, S.; GREENBLATT, R.; REBER, P. J.; & KOUNIOS, J. (2004) “Neural Activity When People SolveVerbal Problems with Insight”. *In PLOS Biology* (2)4 : 500-510.

KELLER, E. F. (2002) *The Century of the Gene*. Cambridge : Harvard UP.

KELSO, J.; BRESSLER, S.L.; BUCHANAN, S.; DeGUZMAN, G.C.; DING, M.; FUCHS, A.; & HOLROYDET, T. (1991) “A Phase Transition in Human Brain and Behavior”. *In Physics Letters A* (169) 3 : 134-144.

KONKEL, F. (2016) “The CIA Says It Can Predict Social Unrest as Early as 3 to 5 Days Out”. *In Defense One*. http://www.defenseone.com/technology/2016/10/cia-says-it-can-predict-social-unrest-early-3-5-days-out/132121 (Accessed, January, 29th, 2016)

LE, Q. V.; RANZATO, M.; MONGA, R.; & NG, A. Y. (2013) “Building High-level Features Using Large Scale Unsupervised Learning”. *Proceedings of the 29th International Conference on Machine Learning*. Edinburgh.

LEVI-KRAUS, G. (2016) “The Great A.I. Awakening”. *In New York Times Magazine*. Dec. 14, 2016.

POSTLE, B. (2006) “Memory as an Emergent Property of the Mind and Brain”. *In Neuroscience* (139)1 : 23–38.

RAFFONE, A.; & VAN LEEUWEN, C. (2001) “Chaos and Neural Coding : Is the Binding Problem a Pseudo-problem?”. *In Behavioral and Brain Sciences* (24)05 : 826-827.

ROCHA, L. M. (1998) “Selected Self-Organization and the Semiotics of Evolutionary Systems”. *In Evolutionary Systems : Biological and Epistemological Perspectives on Selection and Self-Organization*. S. Salthe; G. Van de Vijver; & M. Delpos (Eds.). Kluwer Academic Publishers : 341-358.

STIEG, A.; AVIZIENIS, A.; SILLIN, H.; AGUILERA, R.; SHIEH, H.-H.; MARTIN-OLMOS, C.; SANDOUK, E.; AONO, M.; & GIMZEWSKI, J. (2014) “Self-Organization and Emergence of Dynamical Structures in Neuromorphic Atomic Switch Networks”. In *Memristor Networks*. A. Adamatsky & L. Chua (Eds.) New York and London : Springer : 173-209

THOMPSON, D. W. (1942) *On Growth and Form*. Cambridge : Cambridge University Press.

THOMPSON, E.; & VARELA, F. (2001) “Radical Embodiment : Neural Dynamics and Consciousness”. *In Trends in Cognitive Sciences* (5)10 : 418–425.

TSUDA, I. (2001). “Toward an Interpretation of Dynamic Neural Activity in Terms of Chaotic Dynamical Systems”. *In Behavioral and Brain Sciences* (24)5 : 793-810.

TURING, A. (1932) “Nature of Spirit”. http://www.turingarchive.org/browse.php/C/29 (Accessed, February 3rd, 2017).

__________. (1952) “The Chemical Basis of Morphogenesis”. *In Philosophical Transactions of the Royal Society of London*. Series B (237) 641 : 37-72.

TURING, A.; & WARDLAW, C. (1952) “A Diffusion Reaction Theory of Morphogenesis in Plants”. http://www.turingarchive.org/browse.php/C/7 (Accessed, January 20th, 2016).

UHLHASS, P.; PIPA, G.; LIMA, B.; MELLONI, L.; NEUENSCHWANDER, S.; NIKOLIC, D.; & SINGER, W. (2009) “Neural Synchrony in Cortical Networks : History, Concept and Current Status". *In Frontiers in Integrative Neuroscience* (3)17 : 1-19.

WADDINGTON, C. (1940) *Organisers and Genes.* Cambridge : Cambridge University Press.

WINFREE, A. (1987) *When Time Breaks Down : The Three-Dimensional Dynamics of Electrochemical Waves and Cardiac Arhythmias.* Princeton : Princeton University Press.

ZHANG, H., & JACOBS, J. (2105) "Traveling Theta Waves in the Human Hippocampus". *In The Journal of Neuroscience* (36)35 : 12477-12487.

## Abstract

Artificial Intelligence (AI) designers try to mimic human brain capabilities with "self-learning" neural networks trained by selection processes. Yet decades on, AI still fails the Turing Test. While computers use codes and develop *algorithms* apart from contexts, living cells use *signs* and develop *semiotic habits* within contexts. This difference, I argue, is partly due to the collective activities of biological neurons that produce traveling waves, which, in turn, further constrain neural activity. It appears wave patterns function as contexts shaping the content of the local connections. At the time of his death, Alan Turing was investigating the organizing role of emergent wave patterns on biological development. Had he lived to continue this work, he might have reoriented AI research, which instead has become merely a tool for stereotyping and regularizing, not thinking.

**Keywords** : Semiotics Habits; Emergence of Semiosis; Alan Turing; Biological Computation; Poiesis.

## Résumé

Les concepteurs d'intelligence artificielle (IA) tentent d'imiter les aptitudes des cerveaux humains au moyen de réseaux neuronaux qui apprennent par eux-mêmes grâce à des processus de sélection. Mais même après des décennies d'efforts, l'IA n'en continue pas moins d'échouer le test de Turing. Alors que des ordinateurs utilisent des codes et développent des *algorithmes* hors contexte, les cellules vivantes utilisent des *signes* et auto-organisent des *habitudes sémiotiques* de manière contextualisée. Je soutiens que cette différence s'explique, en partie, par les activités collectives des neurones biologiques qui produisent des ondes, lesquelles contraignent l'activité neuronale. Il appert que les motifs ondulatoires fonctionnent comme des contextes, et qu'ils informent le contenu des connexions locales. Au moment de sa mort, Alan Turing l'inventeur original de l'IA, s'intéressait au rôle organisateur des motifs ondulatoires sur le développement biologique. S'il avait vécu et poursuivi ses travaux, il aurait peut-être réorienté la recherche sur l'IA, laquelle est devenue un outil servant simplement la régularisation et la création de stéréotypes, et non un outil de pensée.

**Mots-clés** : Habitudes sémiotiques; émergence de la sémiose; Alan Turing; calcul biologique; poïèsis.

VICTORIA N. ALEXANDER is a philosopher of science, author of *The Biologist's Mistress : Rethinking Self-Organization in Art, Literature and Nature.* She is a member of the Third Way of Evolution group, and a director at the Dactyl Foundation (New York). She is on the editorial boards of *Biosemiotics* journal (Springer Publishing) and *Meaning Systems* book series (Fordham University Press). Her work in saltational evolutionary theory appears in Fine Lines : *Vladimir Nabokov's Scientific Art*, published by Yale University Press. She is also a novelist and is currently serving a three-year term on the Fulbright Specialist roster as an expert in art-science interactions.