
Prologue

There may be nothing new under the sun, but
permutation of the old within complex systems
can do wonders.

STEPHEN JAY GOULD, 1977

THIS IS A BOOK about thought, memory, creativity, consciousness, narrative, talking to oneself, and even dreaming. In a book that parallels this one, *How Brains Think*, I explored those subjects in a general way but here I treat them as some of the predicted outcomes of a detailed darwinian theory for how our cerebral cortex represents mental images — and occasionally recombines them, to create something new and different.

This book proposes how darwinian processes could operate in the brain to shape up mental images. Starting with shuffled memories no better than the jumble of our nighttime dreams, a mental image can evolve into something of quality, such as a sentence to speak aloud. Jung said that dreaming goes on continuously but you can't see it when you are awake, just as you can't see the stars in the daylight because the sky is too bright. Mine is a theory for what goes on, hidden from view by the glare of waking mental operations, that produces our peculiarly human type of consciousness with its versatile intelligence. As Piaget emphasized, intelligence is what we use when we don't know what to do, when we have to grope rather than using a standard response. In this book, I tackle a mechanism for doing this exploration and improvement offline, how we think before we act and how we practice the art of good guessing.

Surprisingly, the subtitle's *mosaics of the mind* is not just a literary metaphor. It is a description of mechanism at what appears to be an appropriate level of explanation for many mental phenomena — that of hexagonal mosaics of electrical activity, competing for territory in the association cortex of the brain. This two-dimensional mosaic is predicted to grow and dissolve, much as the sugar crystals do in the bottom of a supersaturated glass of iced tea. Looking down on the cortical surface, with the right kind of imaging, ought to reveal a constantly changing patchwork quilt.

A closer look at each patch ought to reveal a hexagonal pattern that repeats every 0.5 mm. The pattern within each hexagon of this mosaic may be the representation of an item of our vocabulary: objects and actions such as the cat that sat on the mat, tunes such as Beethoven's dit-dit-dit-dah, images such as the profile of your grandmother, a high-order concept such as a Turing Machine — even something for which you have no word, such as the face of someone whose name you haven't learned. If I am right, the spatiotemporal firing pattern within that hexagon is your cerebral code for a word or mental image.

THE OTHER PHRASE IN THE BOOK'S TITLE that is sure to be mistaken for literary license is, of course, the *cerebral code*. The word "code" is often only a short way of saying "unlocking the secrets of" and newspaper headline writers love such short words. Neurobiologists also speak loosely about codes, as when we talk of "frequency codes" and "place codes," when we really mean only a simple mapping.

Real codes are phrase-based translation tables, such as those of bank wires and diplomatic telegrams. A code is a translation table whereby short abstract phrases are elaborated into the "real thing." It's similar to looking up *ambivalence* in a dictionary and getting an explanatory sentence back. In the genetic code, the RNA nucleotide sequence CUU is translated into leucine, the triplet GGA into glycine, and so on. The cerebral code, strictly speaking, would be what we use to convert thought into action, a translation table between the short-form cerebral pattern and its muscular implementation.

Informally, *code* is also used for the short-form pattern itself, for instance, a nucleotide chain such as *GCACUUCUUGCACUU*. In this book, *cerebral code* refers to the spatiotemporal firing pattern of neocortical neurons that is essential to represent a concept, word, or image, even a metaphor. One of my theoretical results is that a unique code could be contained within a unit hexagon about 0.5 mm across (though it is often redundantly repeated in many neighboring hexagons).

It was once thought that the genetic code was universal, that all organisms from bacteria to people used the same translation table. Now it turns out that mitochondria use a somewhat different translation table. Although the cerebral code is a true code, it surely isn't going to be universal; I doubt that the spatiotemporal firing pattern I use for *dog* (transposed to a musical scale, it would be a short melody, perhaps with some chords) is the same one that you use. Each person's cerebral codes are probably an accident of development and childhood experience. If we find some commonality, for example, that most people's brains innately use a particular subset of codes for animate objects (say, C minor chords) and another subset (like the D major chords) for inanimate objects, I will be pleasantly surprised.

An important consequence of my cerebral code candidate, falling out of the way in which cortical pattern-copying mechanisms seem capable of generating new categories, is that ascending levels of abstraction become possible — even analogies can compete, to help you answer those multiple-choice questions such as “A is to B as C is to D,E,F.” With a darwinian process operating in cerebral cortex, you can imagine using stratified stability to generate those strata of concepts that are inexpressible except by roundabout, inadequate means — as when we know things of which we cannot speak. That's the topic of the book's penultimate chapter, “The Making of Metaphor.”

AS A NEUROPHYSIOLOGIST with long experience doing single neuron recordings in locales ranging from sea slug ganglia *in vitro* to human cerebral cortex *in situ*, I undertook this theoretical venture about a decade ago. I didn't set out to



explain representations, or even the nature of working memory. Like most people in neurobiology, I considered such questions too big to be approached directly. One had to work on their foundations instead.

Back then, I had a much more modest goal: to seek brain analogies to the darwinian mechanisms that create higher-order complex systems in nature, something that could handle Kenneth Craik's 1943 notion of simulating a possible course of action before actually acting. We know, after all, that the darwinian ratchet can create advanced capabilities in stages, that it's an algorithmic process that gradually creates quality — and gets around the usual presumption that fancy things require an even fancier designer. We even know a lot of the ins-and-outs of the process, such as how evolution speeds up in island settings and why it slows down in continental ones.

However attractive a top-down cognitive design process might be, we know that a bottom-up darwinian process can achieve sophisticated results, given enough time. Perhaps the brain has invented something even fancier than darwinism, but we first ought (so I reasoned) to try the darwinian algorithm out for size, as a foundation — and then look for shortcuts. In 1987, I wrote a commentary in *Nature*, "The brain as a Darwin Machine," proposing a term for any full-fledged darwinian process, in analogy to the Turing Machine.

Indeed, since William James first discussed the matter in the 1870s during Charles Darwin's lifetime, darwinian processes have been thought to be a possible basis for mental processes, a way to shape up a grammatically correct sentence or a more efficient plan for visiting the essential aisles of the grocery store. They're a way to explore the Piagetian maze, where you don't initially know what to do; standard neural decision trees for overlearned items may suffice for answering questions, but something creative is often needed when deciding what to do next — as when you pose a question.

When first discovered by Darwin and Wallace and used to explain the shaping up of new species over many millennia, the darwinian ratchet was naturally thought to operate slowly. Then it was discovered that a darwinian shaping up of antibodies also

occurs, during the days-to-weeks time course of the immune response to a novel antigen. You end up with a new type of antibody that is a hundred times more effective than the ones available at the time of infection — and is, of course, far more numerous as well. What would it take, one asks, for the brain to mimic this creative mechanism using still faster neural mechanisms to run essentially the same process? Might some milliseconds-to-minutes darwinian ratchet form the foundation, atop which our sophisticated mental life is built?

As Wittgenstein once observed, you gain insights mostly through new arrangements of things you already know, not by acquiring new data. This is certainly true at the level of biological variation: despite the constant talk of “mutations,” it’s really the random shuffle of grandparent chromosomes during meiosis as sperm and ova are made, and the subsequent sexual recombination during fertilization, that generates the substantial new variations, such as all the differences between siblings. Novel mental images have also been thought to arise from recombinations during brain activity. In our waking hours, most of these surely remain at subconscious levels — but many are probably the same sorts of juxtapositions that we experience in dreams every night. As the neurophysiologist J. Allan Hobson has noted:

Persons, places, and time change suddenly, without notice. There may be abrupt jumps, cuts, and interpolations. There may be fusions: impossible combinations of people, places, times, and activity abound.

Most such juxtapositions and chimeras are nonsense. But during our waking hours, they might be better shaped up in a darwinian manner. Only the more realistic ones might normally reach “consciousness.”

THE MECHANISTIC REQUIREMENTS for this kind of darwinian process are now better known than they were in the 1870s; they go well beyond the selective-survival summary of darwinism that so often trivializes the issue. Charles Darwin, alas, named his theory *natural selection*, thus leading many of his followers to focus on

only one of what are really a half-dozen essential aspects of the darwinian process. Thus far, most "darwinian" discussions of the brain's ontogeny, when examined, turn out to involve only several of the darwinian essentials — and not the whole creative loop that I discuss in later chapters.

I attempted to apply these six darwinian attributes to our mental processes in *The Cerebral Symphony* and in "Islands in the mind," published in *Seminars in the Neurosciences* in 1991, but at that time I hadn't yet found a specific neural mechanism that could turn the crank. Later in 1991, I realized that two recent developments in neuroscience — emergent synchrony and standard-length intracortical axons — provided the essential elements needed for a darwinian process to operate in the superficial layers of our cerebral cortex. This neocortical Darwin Machine opens up a broad neurophysiological-level consideration of cortical operation. With it, you can address a range of cognitive issues, from recognition memory to higher intellectual function including language and plan-ahead mechanisms — even figuring out what goes with the leftovers in the refrigerator.

DESPITE THE HERITAGE from William James and Kenneth Craik, despite the recent interdisciplinary enthusiasm for fresh darwinian and complex adaptive systems approaches to long-standing problems, any such darwinian crank is going to seem new to those scientists who have little detailed knowledge of darwinian principles beyond the crude "survival of the fittest" caricature.

For one thing, you have to think about the statistical nature of the forest, as well as the characteristic properties of each type of tree. Population thinking is not an easily acquired habit but I hope that the first chapter will briefly illustrate how to use it to make a list of six essential features of the darwinian process — plus a few more features that serve as catalysts, to turn the ratchet faster. Next comes a dose of the local neural circuits of cerebral cortex, as that is where the triangular arrays of synchronized neurons are predicted, that will be needed for both the coding and creative complexity aspects. This is also where I introduce the hexagon as the smallest unit of the Hebbian cell-assembly and

estimate its size as about 100 minicolumns involving 10,000 neurons (it's essentially the 0.5 mm macrocolumn of association cortex, about the same size as the ocular dominance columns of primary visual cortex but perhaps not anchored as permanently). This is where compressing the code is discussed and that puts us in a position to appreciate how long-term memory might work, both for encoding and retrieval.

About halfway through the book, we'll be finished with the circuitry of a neocortical Darwin Machine and ready to consider, in *Act II*, some of its surprising products: categories, cross-modality matching, sequences, analogies, and metaphors. It's just like the familiar distinction we make between the principles of evolution and the products of evolution. The products, in this case, are some of the most interesting ways that humans differ from our ape cousins: going beyond mere category formation to shape up further levels of complexity such as metaphor, narrative, and even agendas. I think that planning ahead, language, and musical abilities also fall out of this same set of neocortical mechanisms, as I've discussed (along with their "free lunch" aspects, thanks to common neural mechanisms) in my earlier books.

SOME READERS MAY HAVE NOTICED BY NOW that this book is not like my previous ones. They were primarily for general readers and only secondarily for fellow scientists, but that order is reversed here. To help compensate, I've provided a glossary starting at page 203 (even the neuroscientists will need it for the brief tutorials in chaos theory and evolutionary biology). Consult it early and often.

And I had the general reader firmly in mind as I did the book design (it's all my fault, even the page layout). The illustrations range from the serious to the sketchy. In *Three Places in New England*, the composer Charles Ives had a characteristic way of playing a popular tune such as "Yankee Doodle" and then dissolving it into his own melody; even a quote of only four notes can be sufficient to release a flood of associations in the listener (something that I tackle mechanistically in *Act II*, when warming up for metaphor mechanisms). As a matter of writer's technique,

I have tried to use captionless thumbnail illustrations as the briefest of scene-setting digressions, to mimic Ives. I have again enlisted the underground architect, Malcolm Wells, to help me out — you won't have any trouble telling which illustrations are Mac's! Furthermore, a painting by the neurobiologist Mark Meyer adorns the cover. For some of my own illustrations, alas, I have had to cope with conveying spatiotemporal patterning in a spatial-only medium (further constrained by being grayscale-only and tree-based!). Although I've relied heavily on musical analogies, the material fairly begs for animations.

I have resisted the temptation to utilize computer simulations, mostly for reasons of clarity (in my own head — and perhaps also the reader's). Simulations, if they are to be more than mere animations of an idea, have hard-to-appreciate critical assumptions. At this stage, simulations are simply not needed — one can comprehend the more obvious consequences of a neocortical Darwin Machine without them, both the modular circuits and the territorial competitions. Plane geometry fortunately suffices, essentially that discovered by the ancient Greeks as they contemplated the hexagonal tile mosaics on the bathhouse floor.