Neosentience: Sensitivity, Persistence, and Insight

ISIS 266S Spring 2011

Elliott Hauser

Preface: What am I ‘up to’?

Introductions and prefaces are often where authors share with us their motivations for creating their work, and offer us motivations for engaging in it. I find these segments of scholarly works essential to, and perhaps more important than, the content of the scholarly product in all its technical detail in assessing what has been done in a work and whether it relates to my needs at the time. I’m indebted to Professor Alan Nelson for introducing me to a mode of scholarly inquiry that begins with the question, “What’s this person up to?” In other words, ‘what is their aim and why have they chosen it?’ This allows a perspicuous integration of different scholars’ words and ideas, even when they disagree or are disparate. By way of introducing my work in the present paper I would like to both tell you what I’m “up to” and relate the program I’ve set myself to to various other scholars, clarifying the relationship of my ideas to others’.

I’m seeking a powerful and unified explanatory language for the phenomena I’ve observed. This places me within the broad tradition of seeking explanation, but puts me at odds with others also well within it. By way of example, I’d like to quote the opening paragraph from Hava Siegelmann’s Neural Nets and Analog Computation:

> Humanity’s most basic intellectual quest to decipher nature and master it has led to numerous efforts to build machines that simulate the world or communicate with it. The computational power and dynamic behavior of such machines is a central question for mathematicians, computer scientists, and, occasionally, physicists.
>
> Siegelmann 1999, p. ix

Thus grounded, she begins her inquiry into the properties of artificial neural networks. While I will support and expand upon much of what she writes about discrete, digital computation versus continuous, analog computation, an exegesis of her first sentence should clearly demonstrate how my ideas come from a different quarter and thus have a very different tack than hers. Siegelmann speaks of '[h]umanity's most basic intellectual quest' with an even tone, but, as is apparent in the rest of the work, she is engaged in this quest. Siegelmann's equations and technical mastery, evident throughout this and other works, are employed to master the portions of nature involved in computation, as defined by the quantitative methods of the field of computer science. She envisions her work supporting "efforts to build machines that simulate the world or communicate with it". Siegelmann seeks to enable the scientific quest of deciphering and mastery through mathematical and computational improvement of engineered systems.

There is a crucial difference between the deciphering and mastery of nature that Siegelmann writes about and the powerful, unified explanatory language that I am after. Siegelmann seeks a deciphering; I explicitly seek a code. Siegelmann seeks mastery, the total predictive power of perfect science; I seek *explanatory* power and a unification of language. These differences, subtle in compact form but extreme in implication, will motivate my divergence from her, when it happens, below. I'll elaborate.

In explicitly seeking a language I'm surrendering to Siegelmann's cipher. I believe that intellectual inquiry occurs within the bounds of language, and that it is a process of linguistic creation and refinement. When I grant Lakoff and Johnson that our knowledge is deeply founded in the ways in which our experience is embodied,[1] I am not seeking to make an argument. I am saying that this is the way that we *should* view knowledge, that this way of speaking and thinking has very precise things to say to us, and that, in conjunction with other, similarly powerful ideas, it is able to offer us deep insight into ourselves and what we know.

---

[1]See Lakoff and Johnson 1980 and 1999; I'm unable to give their ideas due consideration in this work, but am generally receptive to their contention that human knowledge is built us from the particular ways in which we sense and move using our bodies. Seaman and Rössler's functional requirement that the Neosentient have a deep, embodied knowledge of context seems to adequately reflect Lakoff and Johnson's basic insights, or so it seems from my current level of understanding.

This is not a deciphering, but a re-ciphering. Kuhn's insights in the Structure of Scientific revolutions can guide here. In describing the trajectory of inquiry that led to The Structure of Scientific Revolutions, he writes, in that work's Preface:

> A fortunate involvement with an experimental college course treating physical science for the non-scientist provided my first exposure to the history of science. To my complete surprise, that exposure to out-of-date scientific theory and practice radically undermined some of my basic conceptions about the nature of science and the reasons for its special success.
>
> Kuhn 1970 p. v

He goes on to write that these "failures of verisimilitude" he had uncovered were "thoroughly worth pursuing". Though Kuhn does not say it explicitly, I would like to term his reason for shifting from the practice of science to the study of its history and language as a search for 'insight'. Nowhere does he mention anything akin to Siegelmann's decoding of or mastery over nature. He, and I, grant science a "special success"; but we are both, I think, interested in examining science with a view towards crafting and refining the language within which insight and understanding are represented, rather than viewing science as an enterprise of [de]cryptography.

So, though the presentation above is schematic and simplified, I hope it has clearly articulated what I seek to get 'up to' in this paper. I'm interested in insight and understanding, and I think that these things, in an intellectual sense, are produced by and within language. Therefore, I'd like to present and advocate a way of speaking and thinking that, I believe, is powerfully insightful. That insight itself plays a central role in the conception I will advance below should not be problematic; I'm attempting to articulate a series of interconnected concepts that are tautologically related. Tautologies form the basis of all systems of thought; their presence below is the indication that I am suggesting a basis. "Justification must come to an end somewhere",[2] and I am suggesting that our justifications in the enterprise of Neosentience research (and elsewhere) end in the concepts of sensitivity, persistence, and insight. These terms have emerged as I've interfaced with insightful works like those of Siegelmann, Kuhn, and the others I discuss below. For me, they are able to unify their powerful works and ideas and produce very beautiful, albeit speculative, insights. I'll now demonstrate what I've found.

[2]Wittgenstein 1969 p.28e: "204. Giving grounds, however, justifying the evidence, comes to an end;--but the end is not certain propositions' striking us immediately as true, i.e. it is not a kind of *seeing* on our part; it is our *acting*, which lies at the bottom of the language-game"

Sensitivity, Persistence, and Insight

The terms sensitivity, persistence, and insight as I seek to use them need clarification and definition and in this section I'll attempt each in turn, while I demonstrate the reasons I have for thinking them useful. The terms have definite relations to existing others, and I'll mention other ideas in the production of both positive and negative definitions. In general, I'm advocating examining systems with respect to their 'sensitivity' and 'persistence' characteristics, and viewing 'insight' as the outcome of sensitive systems persisting. This basic relationality should hold throughout the paper, and hopefully is clarifying as we begin to dig deeper into the terms.

Persistence

I've arrived at the term persistence in a search for a durable and powerful way of contemplating phenomena as diverse as human language, the evolution of species, quantum indeterminacy, and the structure of cultures. Even cognition, the means by which I'm generating these words and by which you're reading them, can be described in terms of how it persists through time. Below, I attempt to reduce persistence, at least in theory, to entropic characteristics to find a common language in which to describe the phenomenon. But regardless of the success of this project we're quickly led to ask the question 'persists how?' The answer to that question we'll eventually have to resort to a discussion of the systems' sensitivity. But before we delve into sensitivity it we'll try and distinguish some basic types of persistence.

Types of persistence: Physical

Persistence in its simplest sense can be thought of a propagation through space-time. Something can remain three dimensionally motionless while propagating through time (such as that chair over there), or move in a complicated way. Relativistic considerations here don't change the basic conception.

But there are types of persistence that are interesting to us. There are elements, for instance, which seem to have a kind of stasis and that, outside of nuclear reactions, remain essentially indistinguishable from others. These elements can be said to be in a condition of persistent stasis through time in the period between formation and fission, etc. Other types of persistence described by the physical and chemical sciences can likely also be described as static (or, perhaps, dynamic equilibrium).

Types of persistence: Biological

Biological and biochemical systems possess many important qualities of persistence. Many of these are subject to natural selection. Gould (2002, p. 608-613) identifies many hierarchies of potential 'Darwinian individuals': alleles (genes), genotypes, cells, organisms, demes, species, and clades. We'll now review these individuals from the perspective of their persistence through time and try to distinguish them based upon persistence characteristics.

The types of persistence observed in these individuals can perhaps be divided into broad categories: stasis (equilibrium), and reproduction. Reproductive persistence involves at least two discrete parts: heredity and ontogeny. Heredity, as we know, is centered around the transmission of DNA, either from mitosis or after the sexual recombination following meiosis and fertilization. Ontogeny, discussed so thoroughly by Gould in Ontogeny and Phylogeny (1977), is the period of growth and development where an organism uses the chain of metabolic processes, guided by DNA, and facilitated by whatever biotic resources available to it from its parent[3], to begin to develop. These chains of metabolic processes ‘shunt’ off entropic forces that would otherwise cause decay.[4] Then, the organism may die. Or reproduce. Theoretically, it may do neither. But senescence (which could be viewed as a defect in metabolism) is ubiquitous. In the case of asexual, unicellular organisms, persistence can be seen as only minimally interrupted by the processes of reproduction. And death, quite literally, may never come.[5] But for multi-cellular, sexual organisms like humans, there is a discrete death for individuals, when persistence comes to an end.

Species, too, have ‘births’ and ‘deaths. The lines between progenitor and descendant are likewise imprecise, and composed of heredity and ontogeny. But species death, like organismal death, is discrete and identifiable (setting aside, of course, the practical problems of confirming an absence; biological death is a hard line). We’ll return to the phenomenon of death below but, for now, let’s examine the heredity and ontogeny of speciation. There is no strand of DNA that is passed from progenitor to descendant species. Rather, the sum of all the alleles in the deme that is participating in speciation form the genome of the daughter species. This is why speciation is, as Gould emphasizes, highly contingent. Sexual reproduction and reproductive isolation of demes both serve to recombine alleles and change relative frequencies. Mutations may alter the alleles themselves, but this process, Gould reminds us, must be isotropic. In fact the sexual recombination and population dynamics tend also to produce isotropic phenomena. It is this ‘equilibrium’ that is of interest to us in terms of

---

[3]Of course these terms don’t apply to unicellular mitosis, but for simplicity we’ll ignore this.
[4]Note that ‘entropic shunting’ may be related to autocatalysis. Regardless, the micro/quantum definition of entropy is in such a state of flux that this term is very vague at the moment. By it I am trying to indicate a chemical property, often possessed by metabolic processes, where entropy is minimized within an organism, maximized in molecules that can be excreted (waste) and yet the total system’s entropy is always increasing (thus preserving the Second Law). The mechanisms aside, it is systems with these properties (and not all of them are living systems) that I wish to indicate.
[5]See Isalan et al 2008 for an indication that recombined ot inserted alleles are often non-deleterious in E. Coli, providing a mechanism for persistence of genetic material through time.

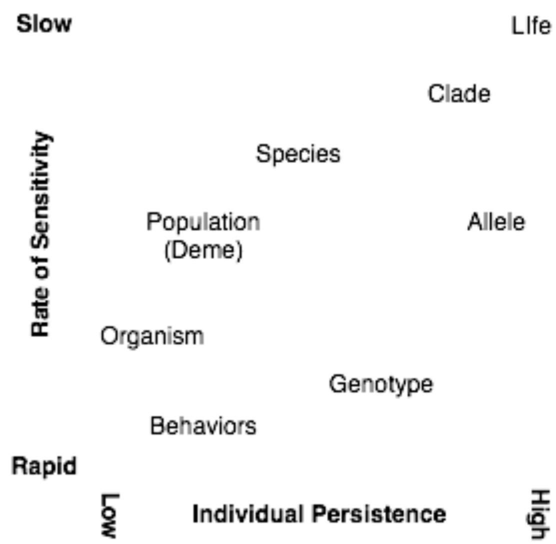
persistence and- interestingly- grounds Gould and Eldridge's theory of speciation, Punctuated Equilibrium (Eldridge and Gould, 1972). It would seem that organisms and species each tend to develop and maintain long-term (10s or 100s of millions of years) morphological stases.

Death, or the phenomenon of biotic systems ceasing to persist, is an interesting place to examine this relationship between persistence and life. There are many theories as to what causes death, whether it is necessary, and its role in evolution. For our purposes, I'd like to concentrate on its entropic qualities.

Death is thermodynamically spontaneous, like all other chemical processes. But so, then, is metabolism and life. What could be the change that breaks the metabolic chain irretrievably? By the logic above, it is not a massive increase in entropy (as some have suggested- see Brooks and Wiley 1988), but rather a loss of the ability to shunt entropy. Once the localization of low-entropy states and the ability to metabolize and maintain those states has dissipated, the chains quickly decay.

Clades, or groups of species, also can be seen as Darwinian individuals. This level of 'identity' is taxing the limits of colloquial understanding, but is still able to bear fruit. Mass extinction events tend to affect species by clade, rather than species independently of one another. This is because clades of species tend to share characteristics, and the certain characteristics that catastrophic extinction affects negatively will lead to clade-based death.

Finally, at the largest scope, we can attempt to view the phenomenon of life on Earth as a Darwinian individual. But this designation breaks down upon further examination. It's not clear what kind of reproduction such a system could be capable of (another method of heredity?). It is, however, possible to view life as an intriguing system that has persisted for billions of years, with the potential to persist for billions more. Other persistent entities, such as stars, have persisted for billions of years, but don't have the potential to propagate the way that life on earth does.[6]



A provisional plotting of biological hierarchies by sensitivity and persistence. This is a draft.

I have separated the axis of Sensitivity into Rapid and Slow to indicate rates of response to change.

I have indicated 'individual persistence' to indicate that the plotting is for any one individual of the named group, on average. For instance 'Organism' would be 'Elliott', while 'Species' would be 'Human'.

Note that in this formulation there is only one known strain of life, so 'life' has membership of one, 'life on earth'

So in our tour of Gould's hierarchy we have been able to identify two main types of persistence: stasis and reproduction. Reproduction further involves two main components, heredity and ontogeny. Can these concepts be expanded into realms beyond biology? Let's briefly test them on human cognition before moving our inquiry into sensitivity.

Types of persistence: Cognition

Human cognition can be seen as an ontogenic process.[7] In psychology, the ontogeny of cognition in particular is usually referred to as 'development'. Human cognition can't very often

---

[6]See Gould's brief treatment of the non-Darwinian individuality of planets and stars, in 2002 p. 608.

[7]As an aside, I have long found it fruitful to think of human education as an ontogenic process. Following Collins argument in The Sociology of Philosophies, the 'great' thinkers undergo a developmental process that is more steeped in a certain intellectual tradition, and are able to generate ideas that are most relevantly creative (likely to find acceptance by generating a new set of philosophical problems). The 'ideas' that philosophers are known for within the community are not their solutions to problems, but their uncovering of problems. Intelligent humans who haven't had the opportunity to be apart of a certain tradition during their formative

be described as in a period of stasis. Like homeostasis during ontogeny, cognitive stasis has to be viewed as a dynamic equilibrium involving flows of balancing and restorative actions, within a larger trajectory. Within homeostasis, a human may sweat or shiver to maintain a certain body temperature. But this limited stasis is a prerequisite for the larger ontologic arc of infancy, adolescence, adulthood, etc. Similarly, cognitive 'stasis' could be maintained by processes of sleep (esp REM sleep), within a larger developmental trajectory.

Perhaps most relevant to persistence and cognition is the phenomenon of memory. Currently a subject of intense research, the mechanisms of memory are relatively poorly understood. It is known that humans can loose portions of memory tied to specific time frames (e.g. short, long, etc) through trauma or degenerative development. These different levels combine to form the substrate of conscious experience. Extreme short term memory is required to aggregate sensory data and form basic awareness. Short term memory, on the order of a few seconds, is the seat of attention and immediate context, while long-term memory is the source of most of human knowledge. The ways in which memories persist are intimately related to the ways in which humans are sensitive to the world, and how these sensations can evoke

---

years will not be acculturated enough to have their creativity be *relevant*. The problems they uncover won't be compelling. And so we find the emergence of intellectual lineages, where teachers and students develop philosophical discourse most fruitfully. Philosophical speciation is perhaps more continuous than biological speciation, but it seems to me that it is similarly precipitated by the development of a flourishing and permissive equilibrium state, adapted to some set of rules ('normal' philosophy, after Kuhn), then punctuated by a state of 'different rules'. These 'different rules' in philosophy are similar to those in science: a crisis where previously unknown or ignored anomalies within the prevailing paradigm are skirted in more and more ad-hoc modifications. The new paradigm is formulated in a state of intense creativity, which is facilitated by the previous, more permissive state of flourishing equilibrium. In philosophy it is also possible for individual thinkers or schools to 'create' crises by attacking other thinkers' ideas. Without the semi-grounding of empirics, philosophy is a much more sociologically fruitful area of study than Kuhn's sciences, though, as I've tried to indicate, similar forces seem to be at work in each, but to different degrees. I'll also note here that Collins has dealt with what he terms 'rapid-discovery science' in Collins 1994.

memories, which alter sensations, etc, in a chain of intra-sensitivity. Let's leave sensitivity alone for just awhile longer, to examine persistence further.

The larger trajectory of cognitive development, within humans, is characterized by a process of learning. Humans are sensitive to changes in their environment (in ways explored below) and, crucially, have a mechanism for preserving those sensitivities. In considering the amazing computational power of the Human brain, we can view each persistence as an 'advice function' for computation. Wiedermann 2010 has described how such advice functions might enable super-Turing computation out of automated systems. I simply wish to point out here that the ability of the human memory to preserve certain features or potentials through time is in effect an advice function, or billions of such advice functions, so that, even if neurons were simple and deterministic (which is under debate), the computational power of the human machine is drastically augmented. This formulation is important for Neosentience research if the paradigm seeks to employ bio-mimetic strategies.

The remaining exposition will consider persistence in combination with sensitivity.

Sensitivity

In "Pattern Flows: Hybrid Accretive Processes Informing Identity Construction", Seaman lays out a powerful re-framing of linguistic inquiry to include not just the word, but all "constellation[s] of sensual perturbations." In re-conceiving meaning productions in humans as an intensely complex, accretive process, Seaman is able to more fully describe an expressive and communicative aesthetic, one built directly around his reconceptualized, recombinant meaning-making mechanism. Seaman employs the term "intra-active" here to indicate the complex interrelationships between hybridizing pattern flows over time. I'd like to import Seaman's discussion of meaning production as an excellent example of the intense complexity of the human. Also, taking "intra-active" as inspiration, I'd like to describe the import of a notion of “intra-sensitivity” within complex sensitive persistent systems such as humans, as this will be an important component in the generation of insight.

As described above, the brain is able to preserve sensory inputs in the form of memory[8], and these inputs can have drastic effects upon humans’ behavior and their future actions. For instance, in extreme cases of verbal abuse, even without threat of violence, humans can exhibit semi-permanently changed behavior or personalities, all based upon sound and vision input. To understand this better, it is necessary to identify types of sensitivities.

Humans are sensitive to the world around them; this we can call sensitivity proper. But humans are especially attuned, even from birth, to each other. This we can call inter-sensitivity. And humans, through the sensations they have and build into memories, are able to think in isolation and even modify sensation drastically based on past experience. This we can call intra-sensitivity.

So Humans are inter- and intra-sensitive. These two concepts are very important to delineating concepts important to Neosentience. Two cellular phones may be inter-sensitive, but without intra-sensitivity no creativity or generation takes place during transmission. The Neosentient, much like the human, will have to be both inter- and intra-sensitive.

But what if the Neosentient were only sensitive to other Neosentients, and not to humans? In what respect could we call it intelligent?

Pragmatically, we need to be sensible of a system’s actions before we could call it intelligent. Whether a system could be intelligent without being inter-sensitive to us is doubtful. Like human infants, a system that was capable of intelligence would have to undergo a human-mediated ontogenic process. It might be possible for the Neosentient to develop sociality and

[8]I should note here that by memory I mean, loosely, an altered potential within one or more neurons due to some stimulus, external or internal. This is distinct from the phenomenal experience of memories or remembering.

intelligence endogenously, but this seems unlikely. At the very least, such a process would require a very large amount of time.[9]

But a certain requirement of such a system is the intra-sensitivity we've observed within humans. That is, a complex relationship between sensation and brain state, with the brain able to preserve 'memories' as advice functions for future cognitive states.[10] Seaman and Rössler's diagram of the Neosentient depicts both short and long term memories as 'pattern matching

---

[9]In the book Sociological Insight, Collins argues that "..if we are ever going to build a computer with the intelligence of a human being, it will have to be programmed by sociologists...If human rationality rests on a non-rational foundation of social rituals, then a computer can handle symbolism the way that humans do only if it too can take part in ritual interactions." (p. vii) Notwithstanding my contention that the Neosentient won't be 'programmed', I mostly agree with Collins here. His 'social rituals' arise out of mutual sensitivity, and the 'rhythmic entrainment' they generate serves to increase the sensitivity and persistence of social interactions.

[10]For an extended abstract on how these functions can be generated through manipulation of persistent symbols by simple machines in a virtual environment, see Wiedermann

devices':

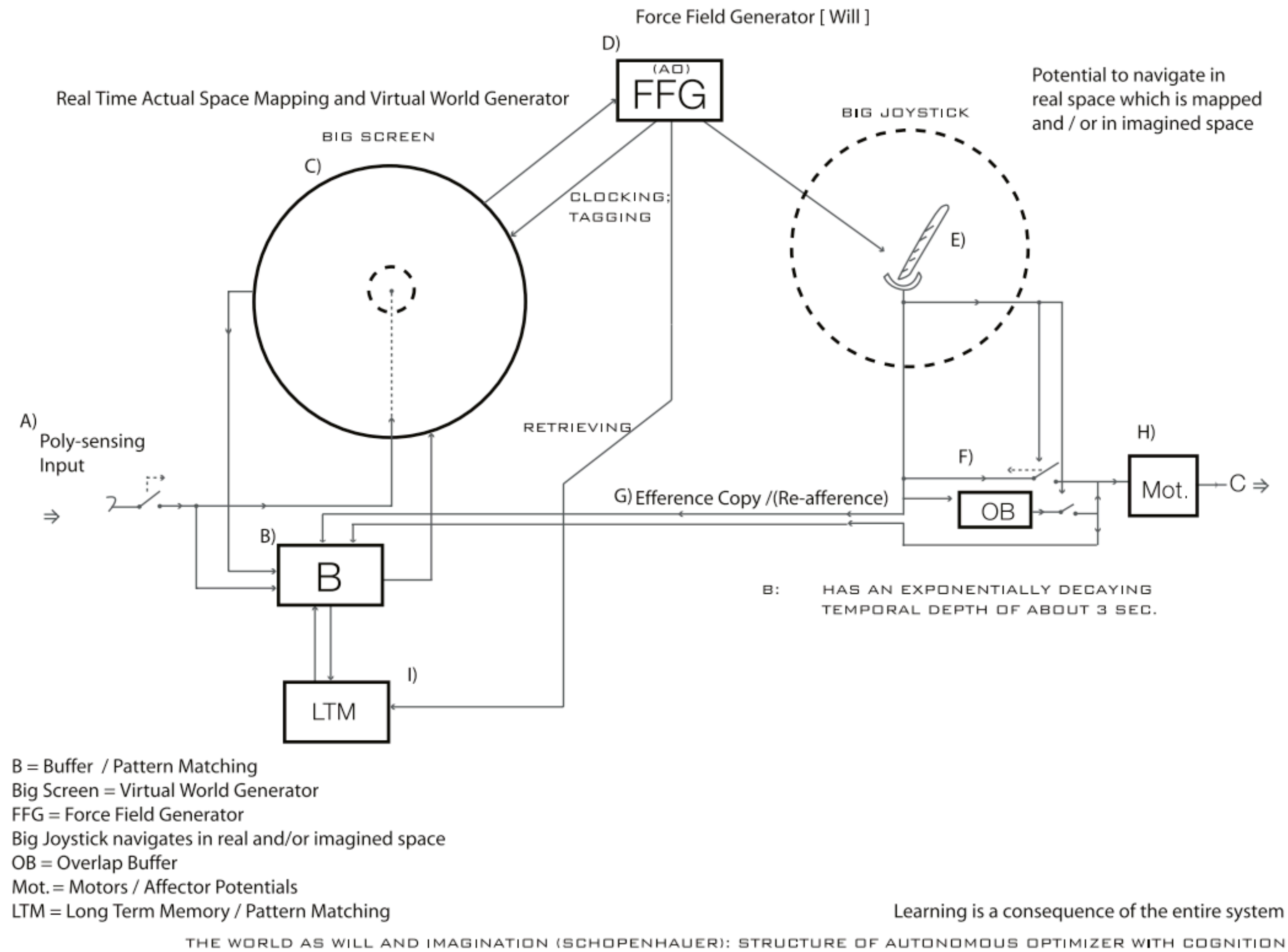


Diagram from Seaman and Rössler 2008

The Neosentient they've designed could potentially exhibit the intra-sensitivities identified here. But further specification is needed to describe how the 'Long Term Memory/Pattern Matching" might influence, and be formed by, the "Buffer/Pattern Matching" module. We won't attempt an in-depth evaluation here, just note the locus of comparison.

Insight

To go deeper with our inquiry, we need to examine what happens when two intra-sensitive systems are sensitive towards each other. Simply, the principle I'm advancing could be stated thus:

The sensitivity of persistent intra-sensitives produces insight

My use of the term insight is idiosyncratically broad, but it is related to many other terms which can provide external support for the breadth I use. For instance, one form of this 'insight' could be Benevolence [/empathy] that Seaman and Rössler have proposed the Neosentient will be capable of. In fact, this could be an important refinement of the above, when intra-sensitive systems interact:

Inter-sensitivity of intra-sensitives can produce empathy

Beyond the pragmatics they utilize in defining the Neosentient, human benevolence, empathy, and wisdom are all highly abstract insights that, I argue, arise from our persistent sensitivity to persistent sensitivity in others.

Maturana's "mutual ontogenic structural coupling" (cited from Seaman 2005) is close to what I mean by the inter-sensitivity of intra-sensitives. Maturana and his student Francisco Varela's theory of cognition is interesting to include here as well, and perhaps more relevant as we transition from describing the ways systems interact to the outcomes of this interaction:

> Living systems are cognitive systems, and living as a process is a process of cognition
>
> (Maturana and Varela, 1980)

A related, but less detailed conception is Randall Collins' term 'Interaction Ritual Chains'.[11] Since Collins' theory is macroscopic in focus, he doesn't need to cover social ontogeny like Maturana, with his more detailed focus, does. But Collins' ritual *chains* contain the mechanism of propagation and his notion of *mutual entrainment* is close to Maturana's structural coupling.

---

[11]Collins develops and employs this term in various works, but the best and most focused introduction to is, not surprisingly, his 2004 book Interaction Ritual Chains.

In seeking a way to link and preserve both thinkers' insights, I propose the phrase 'mutual inter-sensitivities between intra-sensitive persistent entities.' The payoff from this terminological shift is that some component of this phrase has the ability to follow both Maturana and Collins to the depths of their insight, while remaining able to represent our consciousnesses as well. This reflexivity is powerful, and a mechanism of external understanding that is inherently recursive is the most powerful.[12]

Seaman has dealt with Insight at some length in his text *(re)Thinking - the body generative tools and computational articulation*. I should describe the differences between the way I seek to use the term and the way that he does. As mentioned before, Seaman's notion of intra-action within the body to produce consciousness and thought is related to intra-sensitivity. But, in the context of *(re)Thinking*, Seaman relegates insight to the result of a human interacting with an artificial system- the insight engine- with varied properties. Seaman envisions the system as capable of various forms of 'recombinant informatics' that, in concert with a human user steering it, could produce insight within the human that could advance research into the entailments of consciousness and, thereby, Neosentience. This vision if insight is a constrained subset of the large role I wish to give the term. Let's unpack what I mean.

Insight is a property of a subject system in relation to an object system. Without proceeding further, we can already derive the Santiago theory from this definition. We can derive the Santiago Theory and only that by restricting the role of 'subject' system to living systems or, as I seek to do, allow for a broad relation between systems of multiple types.[13] With that first statement granted, there is a definite need to differentiate between types of insight. I suggest that analyzing insights by the modes of sensitivity and persistence of subject and object system is up to this task. So, a human, inter-/intra-sensitive, observing a glass of water, which

[12]Many other systems are reflexive as well and I don't claim anything other than that this reflexivity is a necessary condition satisfied by the current formulation.
[13]Living systems are of course most capable of intelligible relations with object systems, but rudimentary or degenerate relations are possible as well.

has only physical/chemical sensitivities, will be able to develop insight on the water. This insight will be unbounded and heavily depended upon the 'advice' function his/her memory is able to play. The human is sometimes able to be creative, generative, because of this unboundedness (although a glass of water rarely invokes creative advice functions out of the memory). The glass of water, however, is only able to develop a continually degrading insight as to the human. The human's physical or chemical properties may somehow enter the water-glass-system, but they would there degrade over time. In this rendering, insight does not require consciousness. In fact, insight could be a name for the causal processes that have interconnected two systems (with the caveat that the unbounded creativity of the Human would subvert determinism here).

How would this conception render Seaman's Insight Engine? At first, the majority of the insight would be generated within the human, with the engine acting as a catalyst. But as more of the functionalities Seaman has articulated are implemented, the engine could begin to develop a richer relation to the human. By sensing the human in various ways (profiling, for instance, on a very prosaic level), the insight engine could start to develop a mutually reinforcing relationship. Regardless, the recombination the engine is able to produce will alter the ways in which the human views his/her relationship with other systems, generating insight.

The powerful thing about generalizing insight out of Seaman's engine, is that we can seek insight generation out in the places where it already exists and is thriving: face to face interaction. Viewing existing human activities that alter relationships with external systems as insight generation opens up a rich view of the potential range of insights a successful system could generate. Other activities humans perform, such as reading, or viewing or practicing art are also insight engines in this sense. Seaman has envisioned a simple, theoretically realizable artificial system that could deepen our insight into artificial systems while deepening its insight into individual users- a bootstrapping strategy towards Neosentience. But by expanding the view of what insight is an how it is generated, we can glimpse a future when an inter-,intra-sensitive system will be able to replicate, augment, or surpass the experience gained from deep conversation. And, in doing so, the system will move closer towards generating its own robust insights- themselves intra-sensitive and persistent.

So, to review, we've examined the ways in which certain types of sensitivity produce insight. We haven't been able to parse out these different types of insight that might be generated, beyond mentioning the possibility of empathy, with Seaman and Rössler's Benevolence as an example of how this might be operationalized in an artificial system (notwithstanding questions as to whether their design would be sufficiently intra-sensitive for this ability). But what about moving beyond general insight, and the special case of empathy? This is where Maturana and Varela's Santiago Theory of Cognition can start to lead. If the process of living is a process of cognition, and living systems generate insight through their inter- and intra-sensitivities, then the types of sensitivities of a system can be a reliable guide as to what cognitions it can experience.[14] Though the prototypical intra-sensitive system is a human, types of systems may generate other insights through their respective processes (processes, it should be noted, of persistence). So species' insight could be termed Evolution, or perhaps, a genome. Cultures' insights, which are manifold, could be termed art, language, or custom. Each of the intricate emergences that are most interesting to us, where science must stretch out into theory, or in the regions abandoned to the humanities and social sciences, in each of these places we see inter-, intra-sensitive systems generating insight through their processes of persistence. Maybe, with work, this language could unify these phenomena in to one beautiful, anticipated whole.[15]

---

[14]Note here that by sensitives I am not only indicating perceptual abilities, such as sight, but all of the particular sensitives of, for example, the visual cortex that allow pattern matiching, edge detection, and the association of concepts and other sensations with visual sensations. Some sense of the word 'dynamism' might better indicate my intention here.

[15]I must explain my use of 'anticipated' here. I do not mean 'predicted' but rather I'm trying to indicate a certain disposition we have towards familiar phenomena. It's still wonderful for me to see a city dweller actually see the night sky. The pleasure of that and the pleasure of actually looking into that sky myself are anticipated: they have a place within my views of the world. It is

Conclusion

Returning to the style of the Preface, I'll here note that it's in Conclusions where authors generally state what they think they've said, or proven, and why this is important. I'll do the same, and again in relation to other authors I've cited.

Roger Penrose, in the conclusion to Shadows of Mind, writes

> Genuine consciousness involves an awareness of an endless variety of qualitatively different things--such as of the green color of a leaf, the smell of a rose, the song of a blackbird, or the soft touch of a cat's fur; also of the passage of time, of emotional states, of worry, of wonder, and of the appreciation of an idea. It involves hopes, ideals, and intentions, and the actual willing of innumerable different bodily movements in order that such intentions may be realized...There is no question of our being able to understand such matters merely in terms of the physics of critical amounts of coherent mass movement. Yet without such an opening into a new physics, we shall be stuck within the straight-jacket of an entirely computational physics...

In reading these words, it's important to remember what he's trying to do. For Penrose, his mission is to carve out space within science for these very important, subjective experiences. He implies that without his efforts, or those similar to his, there is no way to reconcile phenomenal experience with scientific laws. My goal in this paper is again similar- a unification of cognition with scientific understandings- but my efforts are less urgent. Penrose reminds me of the scientists and philosophers that really believe that they must 'ground' their systems of thoughts, or 'carve out space' for those things that are important and real to us in what Collins terms 'banal reality'. I don't see any urgency in this project (as I once did), but rather an

---

just this sort of anticipation that Penrose was seeking for 'emotional states' 'worry' 'wonder' and 'the appreciation of an idea' (see the Conclusion for more on Penrose). For, in the language of his science, each of these was a phenomenal encounter, a puzzle, without a name or a firm place within the rest of knowledge. When we encounter phenomena with warmth and familiarity, we have anticipated them. We are able to give them a place within a continuous language. This does not require us to surrender wonder, awe, or excitement. Hopefully, it will increase all of these things by providing for a more full experience: the universe, as visible in a dark sky, and worry, wonder, and ideas, all at once and in the same language.

opportunity, an invitation to think in new, broader, reconciling ways. And this is the attitude and perspective I bring to this paper. My ideas are admittedly speculative, and not properly grounded. But they are groundable, and tinged with excitement and, for me, expectant with insight.

So, in conclusion to this paper, I'd like to speak about the benefits of traveling through the very dense and beautiful forest of words that surrounds cognition, artificial intelligence, computation, and Neosentience. I've not sought to construct anything firm or inhabitable here. Away from the heavy machinery of science and mathematics, where things are said and done with purpose and certainty, fitting tightly, and concrete foundations are poured and reinforced, there is a place of great beauty where we can walk, see, and gently intervene. This is the place I hope to have led us too. More Andy Goldsworthy than Frank Gehry, I want to gently intervene in these other scholars' words and, hopefully, point at something that is only sayable in a fleeting way but concrete and lasting in the minds of those who've gone there, who've experienced what I've done.

Postscript: Moving beyond algorithms

This portion of the paper I am least confident of. I include it to outline further work that I believe will be needed to advance Neosentience research. The ideas in this section seek to advance Seaman and Rössler's project of articulating a "new paradigm of computation", but they will do so in the terms discussed above. If I'm successful, I'll have shown how the sensitivity|persistence|insight paradigm 'pays off' in clarifying the requirements of producing the "well-defined functionalities" that Seaman and Rössler have laid out for the Neosentient.

A Crisis in Information and Computation?

Kuhn's account of a scientific crisis in The Structure of Scientific Revolutions (1969) involves an aging scientific paradigm, whose tools have been adequate to solve its problems for some years, encountering anomalies of experience or calculation that require elaborate theoretical explanation. Kuhn's examples of the Copernican revolution, Newton's theory of light and color, and Galileo's discoveries regarding motion all occurred at times of great theoretical creativity, mostly directed at 'patching up' existing theories. Surveying the landscape of Information Theory, it's tempting to see the same patterns emerging. Shannon's theory of information [transfer] (see Shannon 1993 for a full account) has created a solveable set of problems and provided usable tools to solve them. But the question of what information *is* isn't addressed in his theory, as he noted (p. 5, Shannon 1993). Many Theories of Information have proliferated, especially in new domains of computation like DNA or quantum. More recently, Floridi has championed a new philosophy of information (Floridi 2004). The far-from-equilibrium phenomena studied by the humanist and social disciplines are difficult to represent using current Information Theory. Also, the type of 'computations' possible by individual minds and collections of them are difficult to describe at all, much less represent in a language usable within Information or Computer Science.

It's against this backdrop that I want to place my 'post-algorithmic' answer to Seaman and Rössler's call for a new paradigm of computation. This paradigm is needed for Neosentience, true, but is also needed to re-orient computer science and ground Information Science in a definition of its subject more versatile than Shannon's and more specific than Webster's Dictionary.

There have been an interestingly large number of recent attempts to answer the question 'What is information?'[16] Bob Losee at UNC has been working on his text, 'Information from Processes' for some years now, and his contention- that information comes from and is given its content and character from processes- is intriguing (Losee 2011). Since his work is still in development, I'll not comment on it in too much detail, but I mention it here because it has the potential to link Information and Computer Science in a deeper way. Shannon's theory represented information in bits, which is useful for communications and digital computers, but inadequate for describing thoughts, languages, or cultures. Information Science, which has positioned itself at the intersection of people and information systems, is in definite need of a definition of information that is at home both in systems and in people. The systems above each seem at home in one or the other.

How is this all related to my call for a 'post-algorithmic' theory of computation? If, after a basic reading of Losee, information is generate by, and derives it character from, processes, a 'post-algorithmic' theory of information should follow neatly from the theory of computation. And many of the answers to the question 'What is Information?' don't involve algorithms at all. But here, before we get too far afield, we should end our speculation about crisis in Information and Computer Science. Suffice it to say that at the moment the field seems to be filled with a cacophony of theories and many anomalies. Ripe, perhaps, for revolution.

---

[16]See, for instance Lombardi 2004, Israel 1998, Israel and Perry 2000, and, for context, Borko 1968.

Algorithm as Logic + Control

Let's start with a classical definition of algorithms and examine their capabilities. Robert Kowalski's 1979 paper *Algorithm = Logic + Control* is a good statement of a traditional definition of an algorithm. Remember the source of this work within computer science, which in 1979 was, as now, seeking both greater efficiency of execution and greater precision within computational systems. Kowalski's use of 'control' establishes it as the method of accomplishing Logical instructions, where the logic is representable as Boolean components and, thereby, executable within a Turing computer. Different methods of control of the same algorithmic logic will yield algorithmically equivalent results. Beyond noting that this equivalence is a unique property of algorithmic systems, we can for the most part leave control and concentrate on Logic. What Logic was Kowalski speaking of? What are its limitations?

Essentially, Kowalski views algorithms as ways of implementing predicate logic in a machine-readable format. Predicate logic was widely viewed by the logical atomists, and later logical positivists, as being the key to infallible understanding of nature through precise language.[17] Predicate logic is unable to express much of the complexity of human thought, though, since it is designed to describe *states of affairs,* which are verifiable through experience. Much of human communication is about expression, not of states of affairs, but of personal states, which are inherently unverifiable by us. Furthermore, verification is not the outcome, or use, of expressions. Music, painting, and dance, for instance, can be seen as expressive acts not describable in predicate logic. Human cognitive states are involved in the production and

---

[17]See, for instance, Ayer's *Logical Atomism*. Bertrand Russell was associated with this earlier atomistic view, while Carnap and the Vienna Circle are generally credited with developing these views into the widely influential Logical Positivism. Interestingly, Wittgenstein's Tractatus was an attempt to show that much of religion, ethics and art fell outside of the grasp of this logical atomism, while the later works of Wittgenstein were focused on describing the limits of language as such. It is the tack taken by Wittgenstein I wish to follow here.

perception of these arts, and an artificial system would have to be able to move beyond predicate logic to engage with these in the way that humans do.

So, if digital algorithms are unable to interface adequately with human cognition, what might? This, as I see it is a basic question that must be answered in the course of Neosentience research. My contention that this answer will not come in the form of an algorithms has one serious barrier left: analog computation and the possibility of analog algorithms.

Algorithms: From Digital to Analog?

Siegelmann (1999) was interested in describing the limitology of analog computations using a specific idealized analog computer. This work can be viewed as the analog equivalent of the Church-Turing Thesis (and on p.154 she does in fact propose an analogous thesis). As Siegelmann notes, the dominant paradigm within computer science, the Church-Turing thesis, "identifies the notion of an algorithm with a computation in the Turing model". This is not a hard identification, and there have been proposals to expand the definition of algorithm. First, we'll discuss the standard definition of Turing-compliant algorithms, and I'll argue that they are unable to maintain sufficient complexity (degrees of freedom, if you will) to represent or reproduce human cognition. Next, we'll briefly explore what might be mean by analog algorithms. Finally, I'll suggest that, to realize Seaman and Rössler's "new paradigm of computation", we should discard algorithms as a fundamental component of computation. We need a new, more powerful description of computational activity to guide Neosentience research. While I am unable to provide such a description, I suggest that the paradigm of sensitive persistence is an ideal platform from which to launch such an activity.

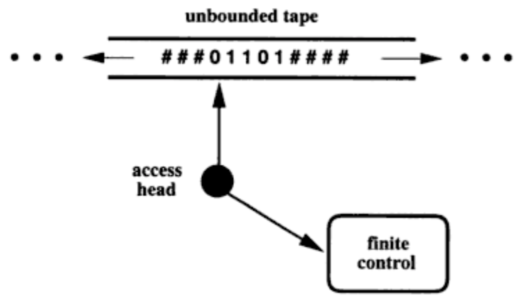


Diagram of a Turing machine, from Siegelmann 1999

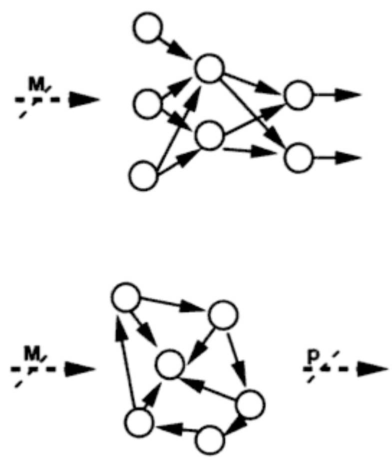


Diagram of sequential (top) and recurrent (bottom) neural networks, from Siegelmann 1999

Analog Algorithms

The meaning of the phrase ‘analog algorithm’ is itself under debate (see Bournez and Dershowitz 2010). Algorithms and the definition of computation are so intertwined that defining algorithms outside of the traditional domain of computing is much more complicated than it might at first seem. But the definition of an analog algorithm still must refer back to a logic of some form. Algorithms, executed on digital or analog computers, involve a program, a performance of logic. Neural networks (in their artificial sense) can exhibit these behaviors, as can actual neuronal systems.[18] But this programmability is a small subset of what actual brains can do, and, regardless of the outcome of the debate of analog algorithms and analog computation, Neosentience requires us to move beyond programmability, and, thereby, beyond algorithms.

Non-programmability

I believe, with Conrad, that digital computers are insufficiently complex to exhibit the functionality demanded of them by consciousness (Conrad 1989). Zeigler's 2002 summary of Conrad's views is instructive:

Aspects of (dis)analogy from theory of computation

Aspect of analogy	Digital computer	Conrad's conception of brain
Activity	Programmable from outside (passive)	Self-organizing (active tendency to learn and evolve)
Exploitation of interactions	Highly constrained (n interactions at most)	Highly interactive (tends toward $n^2/2$ interactions)
Structural programmability	Structurally programmable, low evolutionary adaptability, low computation efficiency	Structurally non-programmable, high evolutionary adaptability, high evolutionary efficiency

Table reproduced from Zeigler 2002.

Conrad's assertion that digital computers likely would never be able to pass the Turing test is in danger from the recent successes of 'Chatbots'.[19] But that assertion is not central to his argument in Conrad 1989. Rather, the central takeaway is that a digital computer should be viewed as in principle incapable of reproducing human cognitive processes not because it is

[18] I'll concede the term 'neural net' to computer science. This field's unfortunate and inaccurate usage is to indicate a system of abstracted nodes and connections, after McCollough and Pitts. To distinguish these simple, abstracted, and logical systems from the much more complex, embodied, and biological systems from which they were modeled, I'll use the term 'neuronal net' to refer to actual brains.

[19] A recent example is Shaikh, Strzalkowski, Taylor, and Webb 2010.

programmable (because small portions of it may be), but because that's *all* it is and it is unable to exhibit the deep adaptability of a neuronal network. Conrad's table from 1989 is reproduced in the Appendix.

So, in an electrochemical computer such as the brain, what might be a fruitful way of cognizing the computational processes at work? The system is non-programmable, self-organizing (shunts entropy), highly adaptable (adaptation = sensitivity + persistence), and highly interactive (intra- and inter- sensitive). The term algorithm is unsuitable for indicating the range of processes here. What would be suitable?

As noted above, I don't have an answer to offer at the moment. But I believe the answer will have to take Conrad's conception of the brain's characteristics into account, and that sensitive persistence will provide a clarifying vocabulary for this work.

Penrose, Non-Computability, and Sensitive Persistence

Roger Penrose's account of the non-computable nature of cognition is important to consider here. His book Shadows of the Mind is composed of two parts- Part I, where he attempts to show through rigorous mathematics that cognitive processes, especially 'understanding', are non-computable, and Part II, where he presents an account of how quantum processes within neuronal microtubules could provide the mechanism for this non-computable quality.

Regarding Penrose's Part I: 'Non-computability' in Penrose's sense refers to classical, Turing computation, and the work of Gödel, proving the incompleteness of this system. The arguments in my paper should be read as sympathetic to Penrose's, while of course less technical. But my main, largely terminological, disagreement with him is the abandonment of the word 'computation'. Instead, I suggest abandoning the term 'algorithmic' as descriptive of this previous paradigm Penrose uses to define his 'computational' phenomena. Hence 'post-algorithmic'.

Regarding Penrose's Part II: I will grant the plausibility of Penrose's suggested mechanism as a potential form of post-algorithmic computation. The objections raised to Penrose's theories of cognition should not apply, though, to sensitive persistence as a whole. If classical processes are sufficiently intra-sensitive to express the complexities of human cognition, then no resort to microtubules may be needed. But the paradigm does provide an important clarifying look at Penrose's work. First, the requirement of intra-sensitivity can guide explorations and scientific descriptions of neuroelectrodynamics[20]. Secondly, the element of persistence can guide inquiry to how these sensitivities- in Penrose's estimation, quantum coherence- is preserved. Penrose seems to be seeking a inter-synaptic coherence. But if the individual microtubules are able to preserve areas of quantum entanglement beyond the interval of a single synaptic firing, the sum effect of this could be to preserve the potential effects of past interfaces with other neurons. This could introduce billions of tiny quantum 'advice' functions (a la Wiederman and van Leewen 2001) into neuronal (and, thereby, inter-synaptic) computations, exponentially increasing the computational power of the system.

I'd like to emphasize here that Penrose's work is one potential method of satisfaction of the requirements implied by the sensitive persistence paradigm. Others may emerge with more plausible or precisely described mechanisms. But my contention is that they must provide a persistent inter- and intra-sensitivity to generate the insight we call cognition.[21]

Quantum Coherence & Insight?

Penrose seems to think that his microtubules need to preserve coherence across synapses, but I believe this may not be the case. One perspective on why quantum entanglement- even in small pieces- might be a source or contributor to cognitive experience can be provided by a discussion of its phase space. Phase space is a representation of the

---

[20]See Aur's Neuroelectrodynamics (2010)
[21]As an aside, I'd like to comment here that the unification of Penrose's theory with the Santiago Theory represents a promising area of future work.

total number of possibilities or degrees of freedom a system has. Losee writes, “The amount of information at a process’ output is proportional to the number of possible characteristics that are available.” Digital systems are prized for the constrained phase space of two outcomes per bit, which are stitched together to form the desired size phase space, offering programmability and reliability. Quantum systems have an infinite and continuous phase space, where a literally infinite number of characteristics is available, and where the probability density for every possible state is defined by the quantum wave function (albeit at infinitesimal values for most of the space). This unbounded phase space, even if playing a minor, ‘advice’-giving role, could be the source of Penrose’s ‘non-computational’ and my ‘post algorithmic’ computations. This work needs further, mathematical definition.

Inter- and Intra-sensitivity

Final support for my contention here might be provided by the inter- and intra-sensitive paradigm. If cognition is a process of inter- and intra-sensitivity (combining my terminology with the Santiago Theory), then digital computers’ lack of the requisite kinds and amounts of sensitivity present in the brain would be seen as the hard line I’m trying to portray it as. Partial confirmation and a partial threat to this conception comes from the recent success of IBM’s Watson. Though the supercomputer is not sentient, nor claimed to be so, it is able to achieve its astounding facility with natural language question-answering through a distributed processing architecture. The results of these distributed processes are then assigned numerical ‘confidence’ values so that they can be compared and the top value delivered. It is this last numerical reduction that collapses the potential of distributed processing and destroys any chance of Watson producing sentience. There is no chance of intra-action amongst the computations analogous to this potential in human brains. Though I believe it unlikely, it may be possible for a hybrid digital-analog computer, employing a distributed architecture, and highly unconstrained interfaces between the results of the distributed processes (e.g. supporting complex intra-action) to begin to express cognition. But, without crossing the boundary into electro- and/or bio-chemical computation, this system will likely remain too tightly constrained. Just as Siegelmann was able to show that powerful analog computers could be constrained to have exactly the computational ability of their more simple digital counterparts, but digital computers could not be shown to produce the complexity of her artificial recurrent neural networks, I contend that living systems (and, likely, other complex, sensitively persistent systems) can be constrained to replicate the computational outcomes of each of these more simple types, but also express a computation inexpressible by those means. This computation will be post-algorithmic; that is, we shall have to abandon an algorithmic conception of computation to describe it. As I see it, this is the great promise of Matutana and Varela's Santiago Theory of cognition, and, perhaps with the 'glue' of sensitivity, persistence, and insight, would provide the "new paradigm of computation" that Seaman and Rössler have identified as necessary to realizing the Neosentient.

Works Cited

Borko, H. (1968). Information Science: What is it?. *American Documentation*, *19*(1), 3-5.

Bournez, O., & Dershowitz, N. (2010). Foundations of Analog Algorithms. *Computer*, (250).

Brooks, D. R., & Wiley, E. O. (1988). *Evolution as Entropy* (Second Edi.). University of Chicago Press.

Collins, R. (2004). *Interaction Ritual Chains*. Princeton University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=3NirI6P8f5EC&oi=fnd&pg=PR9&dq=Interaction+ritual+chains&ots=fFUL6QsdL1&sig=2_wtd9ks-ba_-1H2upNF8MzR25U.

Collins, Randall. (1992). *Sociological Insight: An Introduction to Non-Obvious Sociology* (Second Edi.). New York, NY: Oxford University Press.

Collins, Randall. (1994). Why the Social Sciences Won t Become High-Consensus, Rapid-Discovery Science. *Sociological Forum*, *9*(2), 155-177.

Collins, Randall. (1998). *The Sociology of Philosophies: A Global Theory of Intellectual Change*. Cambridge, MA: Belknap/Harvard University Press.

Collins, Randall. (1999). *The Sociology of Philosophies: A Global Theory of Intellectual Change* (p. 1098). Harvard University Press. Retrieved from http://www.worldcat.org/title/sociology-of-philosophies-a-global-theory-of-intellectual-change/oclc/474674420&referer=brief_results.

Collins, Randall. (2004). *Interaction Ritual Chains*. Princeton, NJ: Princeton University Press.

Conrad, M. (1989). The brain-machine disanalogy. *Bio Systems*, *22*(3), 197-213. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2650754.

Eldredge, N., & Gould, S. J. (1972). Punctuated Equilibria: An Alternative to Phyletic Gradualism. In T. J. M. Schopf (Ed.), *Models in Paleobiology* (pp. 82-115). San Francisco, CA: Freeman, Cooper & Co.

Floridi, L. (2004). Open problems in the philosophy of information luciano floridi. *Metaphilosophy*, *35*(4).

Gould, S. J. (1977). *Ontogeny and Phylogeny*. Cambridge, MA: Belknap/Harvard University Press.

Isalan, M., Lemerle, C., Michalodimitrakis, K., Horn, C., Beltrao, P., Raineri, E., et al. (2008). Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, *452*(7189), 840-845. Nature Publishing Group. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18421347.

Israel, D. J. (1998). What is information?. *Artificial Intelligence*, *1*(4), 243-254. Center for the Study of Language and Information, Stanford University. Retrieved from http://www.sveiby.com/articles/Information.html.

Israel, D., & Perry, J. (1990). What is information?. In P. Hanson (Ed.), *Information, Language, and Cognition* (Vol. 1, pp. 1-19). Vancouver, BC: University of British Columbia Press.

Kowalski, R. (1979). Algorithm = logic + control. *Communications of the ACM*, *22*(7), 424-436. doi: 10.1145/359131.359136.

Kuhn, T. (1970). *The Structure of Scientific Revolutions* (Second Edi.). Chicago, IL: The University of Chicago Press, Ltd.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh*. New York, NY: Basic Books.

Lakoff, G., & Johnson, M. (n.d.). *Metaphors We Live By*. Chicago, IL: The University of Chicago Press, Ltd.

Leeuwen, J. V., & Wiedermann, J. (2001). Emergence of a Super-Turing Computational Potential in Artificial Living Systems. *Advances in Artificial Life: 6th European Conference, ECAL 2001*. Berlin, Germany: Springer-Verlag.

Lombardi, O. (2004). What is Information?. *Foundations of Science*, *9*(2), 105-134. doi: 10.1023/B:FODA.0000025034.53313.7c.

Maturana, H. R., & Varela, F. J. (1980). *Autopoesis and Cognition: the Realization of the Living*. Dordrecht, Holland: D. Reidel Publishing Company.

Penrose, R. (1994). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. New York, NY: Oxford University Press.

Rowley, J. (1998). What is Information?. *Information Services & Use*, *18*, 243-254.

Seaman, B. (2005). Pattern Flows: Hybrid Accretive Processes Informing Identity Construction. *Convergence: The International Journal of Research into New Media Technologies*, *11*(4), 12-31. doi: 10.1177//1354856505061050.

Seaman, Bill, & Rössler, O. (2011). *Neosentience: The Benevolence Engine*.

Shaikh, S., Strzalkowski, T., Taylor, S., & Webb, N. (2010). VCA : An Experiment With A Multiparty Virtual Chat Agent. *Computational Linguistics*, (July), 43-48.

Shannon, C. E. (1993). *Collected Papers*. (N. J. A. Sloane & A. D. Wyner, Eds.). New York, NY: IEEE Free Press.

Siegelmann, H T. (1995). Computation beyond the turing limit. *Science (New York, N.Y.)*, *268*(5210), 545-8. doi: 10.1126/science.268.5210.545.

Siegelmann, Hava T. (1999). *Neural Networks and Analog Computation: Beyond the Turing Limit*. Boston, MA: Birkhäuser.

Turing, A. M. (1937). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *2*(1), 230. Oxford University Press. Retrieved from http://plms.oxfordjournals.org/content/s2-42/1/230.full.pdf.

Wiedermann, J. (2010). On Super-Turing Computing Power and Hierarchies of Artificial General Intelligence Systems. *3rd Conference on Artificial General Intelligence* (Vol. March). Atlantis Press.

Wittgenstein, L. (1969). *On Certainty*. (G. E. M. Anscombe & G. H. von Wright, Eds.). New York, NY: Harper Torchbooks.

Related Works

Ben-Hur, a. (2004). Random matrix theory for the analysis of the performance of an analog computer: a scaling theory. *Physics Letters A*, *323*(3-4), 204-209. doi: 10.1016/j.physleta.2004.01.069.

Breazeal, C. (1998). A Motivational System for Regulating Human-Robot Interaction. *Proceedings of AAAI-98*.

Ji, S. (1999). The cell as the smallest DNA-based molecular computer. *Bio Systems*, *52*(1-3), 123-33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10636037.

Leeuwen, J. V. (2002). The emergent computational potential of evolving artificial living systems. *Ai Communications*, *15*, 205-215.

Olsen, M., Siegelmanndanieli, N., & Siegelmann, H. (2008). Robust artificial life via artificial programmed death. *Artificial Intelligence*, *172*(6-7), 884-898. doi: 10.1016/j.artint.2007.10.015.

Poeppel, D., Overath, T., Popper, A., & Fey, R. (n.d.). Towards a Theory of Information Processing in the Auditory Cortex, 1-39.

Siegelmann, H. (2008). Analog-symbolic memory that tracks via reconsolidation. *Physica D: Nonlinear Phenomena*, *237*(9), 1207-1214. doi: 10.1016/j.physd.2008.03.038.

Siegelmann, H. T., & Holzman, L. E. (2010). Neuronal integration of dynamic sources: Bayesian learning and Bayesian inference. *Chaos (Woodbury, N.Y.)*, *20*(3), 037112. doi: 10.1063/1.3491237.

Verbaan, P., & Leeuwen, J. V. (2004). Complexity of Evolving Interactive Systems. *Complexity*, 268-281.

Wiedermann, J. (1999). Simulating the mind: a gaunlet thrown to computer science. *ACM Computing Surveys*, *31*(3es), 16-es. doi: 10.1145/333580.333595.

Wiedermann, J. (2003). The computational limits to the cognitive power of the neuroidal tabula rasa. *Journal of Experimental & Theoretical Artificial Intelligence*, *15*(3), 267-279. doi: 10.1080/0952813021000047189.

Zenil, H., & Hernandez-quiroz, F. (2008). On the possible computational power of the human mind. *World*, 1-23.

Appendix: Table of Brain-Machine Disanalogies from Conrad 1989

Brain-machine disanalogies.

Digital computer	Brain
Programmed from outside (passive)	Self-organizing (active tendency to learn and evolve)
Structurally programmable, low evolutionary adaptability, low computational efficiency	Structurally non-programmable, high evolutionary adaptability, high evolutionary efficiency
Sequential use of resources (if run in programmable mode)	Massively parallel
Discrete dynamics (simple switches)	Discrete and continuous dynamics mixed
Highly constrained (n interactions at most)	Highly interactive (tends to $n^2/2$ interactions)
Macroscopic switching components	Switching processes at many scales (molecular, neuronal, phase transitions in neural nets)
Horizontal flow of information	Vertical information flow between macroscopic and microscopic forms
Physical realization of a formal system (equations of physics irrelevant to computing)	Highly dependent on material composition and physical interactions
String processing mode of operation	Tactile pattern matching at the molecular level
Classically picturable	Some controlling interactions not classically picturable
Addressable memory	Memory storage and retrieval probably involves synaptic facilitation and assemblies of neurons
Main information processing at network level	Main information processing at intraneuronal level; many specialized types of neurons capable of performing a variety of pattern processing tasks
Clocking mechanism discretizes time	Time coordination mediated by spatiotemporal pattern recognition capabilities of tactilizing neurons