Awareness Lies Outside Turing's Box[1]

Michael Manthey
manthey@acm.org

July 2017

Abstract. The idea of computation as formulated by Alan Turing in the 1920's dominates the contemporary discussion of the mechanism(s) underlying awareness. Unfortunately, the very semantics of computation, as standardly defined, seem namely to *exclude* awareness. Furthermore, this paradigm's descriptions of Nature in general are - for all their utility - conceptually barren in their fundamental sequentiality. We therefore replace Turing's automata-based semantic model with a multi-dimensional vector algebra, namely W.K. Clifford's *geometric algebra*. In our novel *automata-free* re-framing, it is obvious that Turing's semantics is inherently *time-like*, and that geometric algebra's *space-like* semantics provide a fertile foundation for the phenomenon of awareness. The new computational model of distributed systems has the global mathematical form $U(1) \times SU(2) \times SU(3) \times SO(4)$, ie. the Standard Model of physics augmented with 3+1d.

1. Introduction

The advent of computers has given us the opportunity to study *processes* to unprecedented and - as I will show - unexpected depths. Alan Turing, who invented an eponymous abstract universal computer, showed that its *sequential* computations (ie. *processes*) possess great power, but also tantalizing limitations.

As an example of the latter, he proved that multiple "Turing machines", working together in parallel, even non-deterministically, have the same ultimate computational power as a single machine working alone! This result - sequence is sufficient - has profound and far-reaching implications: it enshrines a functional point-of-view, it blesses strong reductionism, it is fundamentally classical (Isaac & Albert) in its reach, and it has stubbornly resisted attempts to convincingly describe the *multi*-process systems that comprise Physics, Biology, and the study of Mind.

Rather, the resulting models seem always to be imitative, often strained, and conceptually thin. The "artificial intelligence" model offered by so-called neural nets, powered by Bayesian statistics, is an excellent example: all the hype (and money) cannot disguise how far it strays from the the real thing - it needs *tens of thousands* of examples to learn a non-trivial category, it is unable to construct

[1] To appear in Proc. 11^{th} Int'l Whitehead Conference, Section "Whitehead, Mathematics and Logic". Cambridge Scholars Publishing (European Process Thought Series), 2018.

non-trivial causal sequences ("planning"), and it clearly does not understand anything that it says or does. Worst of all, it offers no apparent explanation of, or path to, the property of self-awareness. In other words, we get Zombies ... the lights are on but there's no one there. Thus have matters stood for some time, despite decades of effort.

The present claim is that the way out of this impasse is to replace the strait-jacket of sequential processes with the power of multi-dimensional vector algebra. In particular, choose the *geometric algebra* $\mathcal{G}$ invented by W.K. Clifford in 1860, but [unusually] operating over the arithmetic of $\mathbb{Z}_3 = \{0, 1, -1\}$, that is, the usual base $3 = \{0, 1, 2\}$ shifted one slot to the left, whence $1 + 1 = -1$ and zero $=$ *Void*, whence the underlying logic is that of exclusive-or, *xor*: same *vs.* different. The result is a general, *automata-free semantics* of all possible computations.

In the following, we will first show how the basic conditional act of *if-then-else* appears in this new algebraic language, which form reveals a close relationship with the quantum mechanical concept of *measurement*. Next comes the consideration of the *synchronization* of processes, which is necessary to both capture resource ownership and producer-consumer relationships, and to enforce determinism where needed. The synchronization primitives *wait(event)* and *signal(event)* also capture the concepts of memory and causality [Manthey 2013].

These two quintessential sequential functionalities - *if-then-else* and *wait/signal* - turn out to consume a *trivial* fraction of geometric algebra's semantic power. In particular, we will show how geometric algebra easily exceeds Turing's formulations and limitations via its ability to express the *indistinguishability* of events. A cornucopia of insight then reveals itself.

What we find is not of Turing's time-like sequential semantics at all, but rather of its opposite, *space-like* computation.

Here is a helpful conceptual analogy for what is about to be described: consider the distinction between the Operating System (OS) of a computer - nearly invisible in systems like smart phones, but more apparent on larger systems - *and* the "user processes" nowadays called "apps", plus resources like memory and Input/Output. The OS is the *manager*, ensuring that eg. a given screen window is protected from meddling by other processes, or that a given chunk of memory that was initially allocated to process A is now also visible to process B, so they can communicate.

Then there are the user processes A, B, ... themselves. The OS swaps them in and out of memory when they're idle, puts them on queues for needed resources, finds their files, etc. The OS has *no* idea what these processes are about, and oppositely, the processes have *no* inkling of the existence of the OS. The OS's job is to be aware of the *fact* that multiple processes are *simultaneously present* and competing for access to various resources - re-usable resources like memory and devices, and consumable resources like input data streams - to make sure that processes can proceed and do not step on each others' toes. *And*, most importantly, to be completely invisible to the user processes while so doing.

The promised helpful conceptual analogy is then this: the OS corresponds to quantum mechanics, and the user process and resource spaces correspond to the structure of relativistic 3+1d space. Indeed, one could imagine two apps, Experimenter and Observer, who cooperate in trying to figure out what's going on behind the OS's facade. This analogy, while *not!* saying the universe is a big operating system, is nevertheless deeply apt: W^m James' "thinnest of veils" refers to this same distinction, as does the term *axis mundi* [the axis of the world], the place where mind intersects matter.

2. If-then-else

A sequential program, unrolled into its future, forms a system consisting of a single process, namely itself - there is no talk of other processes: even if they're present, any synchronization is transparent, and any interference oblique and unrecognized. The single most important property of a process is that it is a *sequence*: the order in which its events take place is crucial, *defining* in effect what the process does. As will be seen, it is similarly crucial not to confuse the three concepts of ordering/sequence, determinism, and causality, as was done in the early years of quantum mechanics.

Suppose now that X, Y, Z are arbitrary expressions in the algebra, and consider the process XYZ, which states the process "do Z, then do Y, then do X", that is, we *always* operate multiplicatively on the left. If any of X, Y, or Z has an inverse, we could algebraically manipulate the product XYZ to produce some other order. This will not do!

Consider therefore the common view of a computer program - when it is executing - as a sequence of discrete operations

$$()()()()()()()()()()()()()()$$

where each parenthesis-pair stands for a single X, Y, Z-type operation. Such a sequence is called a *process,* and the following is a gloss on [Manthey 2007], to which the reader is referred for a more detailed exposition. This process-level of computational description refers not so much to entities themselves as to their interaction, and the sequence of states this produces. In our model, *everything* is a process, or an object built out of processes.

As just noted, the key property of a process is the exact *order* in which its component operations take place. To capture this ordering property algebraically we will require that each operation "( )" in the above sequence - now viewed as a product - be <u>irreversible</u> (ie. no multiplicative inverse). This prevents algebraic manipulations from changing the effective order.

In physical terms, this means that processes are irreversible and *time-like*, and we will intend these two terms interchangeably, as well as their opposites: possessing an inverse = reversible (which allows wave-like activity) = *space-like*. [This is *not* physical 3-D space, just space-like rotations.]

Thus the generic sequential program DoX; DoY; DoZ translates to the $\mathcal{G}$ product sequence *(-1+DoZ)(-1+DoY)(-1+DoX)*, where it is assumed that *Do?* is unitary, $(Do?)^2 = 1$, whence *(-1+Do?)* is idempotent. A theorem of $\mathcal{G}$ informs us that an expression $E \in \mathcal{G}$ is irreversible *iff* E contains an idempotent as a factor.

We now use this algebraic representation of computation to analyze the if-then-else construction. We will write ***if*** V ***then*** X ***else*** Y, where V, X, Y are arbitrary expressions representing arbitrary computations. For simplicity and with no loss of generality, take $V = a$, a 1-vector ("sensor").

"***if*** a" implies a probing of the current state of a: is it $+1$ (so do X), or is it -1 (so do Y).

Given that the only relevant states of a are ± 1, the next question is how to ascertain which of these obtains? Clearly, said ascertaining requires *measuring* a, where again idempotent operators play the central role. Consider the following identities:[2]

$$\begin{array}{lll} (1+a) = (1+a)(a) & (-1+a) = (-1+a)(-a) & (-1+a) = (1-a)(1-a) \\ (1-a) = (1-a)(-a) & (-1-a) = (-1-a)(a) & (-1-a) = (1+a)(1+a) \end{array}$$

Taking $P = (1+a) = (1+a)(a)$ as an example, multiply P's rhs out to get $a + aa$, whence we see that the $+1$ in the lhs can be seen as the product of a with itself.[3] It follows, and this is the key point, that if the a we have in hand - in the rhs's "$(1+a)$" factor - has the same sign as the a we probe - the rhs's "(a)" factor - then the sign of the scalar will be $+1$, whereas if the a we probe is actually $-a$, then the sign of the scalar will be -1. This also applies if P, oppositely, specifies "$(-a)$" and we find "$-a$" ($\Rightarrow +1$), or we find "+a" ($\Rightarrow$ -1). Finally, take *just* the scalar value from $(-1 \pm a)(\pm a)$ to complete the measurement (one can only *actually* measure scalars ... like a meter reading).[4]

This is the basic act of measurement. Because $(1+a)$ has no inverse, the act of measurement is irreversible, in accordance with contemporary understanding of the equivalence of energy and (Shannon) information. Furthermore, successive measurements using the idempotent form yield no new information, as required, in that $PP = P$. [5]

[2] All algebraic calculations have been done with a purpose-built Z3 calculator written by Douglas Matzke. Those wishing to verify our calculations on other symbolic algebra platforms are warned that floating point (but supposedly "really integer") arithmetic can easily produce errors.

[3] If the lhs scalar were -1, a would be inverted to $-a$ as well.

[4] Ie. the scalar or "dot" product of $(-1+a)$ and a, ie. $(1+a)\cdot a$ To avoid clutter, we will suppress the 'dot' til the end of the derivation.

[5] Actually, $(1+a)$ is the square root ("sqert") of an idempotent, cf. column 3 above, but this is unimportant for our present purposes.

So now we know how to do "***if*** a": we will write $(1{+}a)(a)$ or suchlike, depending. The next issue is to choose the correct continuation depending on what the measurement on a produces.

The basic idea now is to arrange for the conjugate forms $(1{+}a)$ and $(1{-}a)$, whose product is zero, to collide on the unwanted branch of the ***if***, thus eliminating that continuation. A zero means the computation's future is empty, ie. it does not occur; generating a zero to eliminate an unwanted continuation is a key tool in the following.

Therefore, write the test in the ***if*** as a probe: $1{+}a$ or $1{-}a$, acting on the actual a, which can be plus or minus. The ***then*** and ***else*** branches apply respectively $1{+}a$ or $1{-}a$ to the result of the test, whence one of them should yield 0 (because conjugate) and the other the correct continuation based on the observed value of a. There are four possibilities (the | marks off visually (only) the shared ***if***-probe, rightmost because it occurs first): [6]

if	*probe*	***then***	*left branch* | *probe*	***else***	*right branch* | *probe*
1	$(1+a)(+a)$		$X(1+a) \mid (1+a)(a)$ $= -X(1+a)$ ***yes***		$Y(1-a) \mid (1+a)(a)$ $= 0$ ***yes***
2	$(1+a)(-a)$		$X(1+a) \mid (1+a)(-a)$ $= X(1+a)$ ***no***		$Y(1-a) \mid (1+a)(-a)$ $= 0$ ***yes***
3	$(1-a)(+a)$		$X(1+a) \mid (1-a)(a)$ $= 0$ ***yes***		$Y(1-a) \mid (1-a)(a)$ $= Y(1-a)$ ***no***
4	$(1-a)(-a)$		$X(1+a) \mid (1-a)(-a)$ $= 0$ ***yes***		$Y(1-a) \mid (1-a)(-a)$ $= -Y(1-a)$ ***yes***

In situation 1 above, we probe for $+a$ with $(1{+}a)$, and a is in fact $+a$; situation 2 has the same probe, but discovers $-a$; situation 3 probes for $-a$ but discovers $+a$; and situation 4 probes for $-a$ and discovers $-a$. Notice that if we consider all four possibilities concurrently (ie. Left + Right, 1 thru 4), we get zero: this situation (namely, a having both values simultaneously) cannot occur. So instead, combine 1&2 and 3&4 by subtraction to get the desired terms to double instead of cancel: 1-2 $= +X(1+a)$; 4-3 $= +Y(1-a)$, and move the |-cue to the right, eliminating the common probe-preface of the previous version:

1 minus 2: $\quad -X(1+a) \mid (\pm a)$
4 minus 3: $\quad -Y(1-a) \mid (\pm a)$

Finally, run 1-2 and 4-3 concurrently (ie. add), and factor out $(\pm a)$:

$$\begin{aligned} &-X(1+a) \mid (\pm a) \;-\; Y(1-a) \mid (\pm a) \\ &= [-X(1+a) - Y(1-a)] \cdot (\pm a) \end{aligned}$$

If $a = +1$ then the Y term drops out leaving $+X$; and if $a = -1$ then the X term drops out, leaving $+Y$. Just as we wanted! Push the minus-signs on X, Y into the parentheses:

[6] The ***yes*** and ***no*** indicate desired (or not) outcomes.

$$= [X(-1-a) + Y(-1+a)] \cdot (\pm a)$$

and we see that doing *if-then-else* necessarily invokes observation, ie. idempotents, consistent with thermodynamic and quantum measurement theory. The form also makes good computational sense when multiplied out:

$$= X(-1-a)\cdot(\pm a) + Y(-1+a)\cdot(\pm a)$$

which transparently describes two independent processes X and Y, each independently and concurrently testing for its own condition, only one of which will succeed.

NB: if one tries simultaneously to measure with $1+a$ *and* $1-a$, one gets (summing) an inversion (1+1 = -1), but no knowledge of a, in accordance with quantum measurement theory: if one is to get information, one must specify *exactly* what it is one is looking for ... $+a$ or -a, and this *cannot* be finessed.

3. Wait and Signal

Having warmed up with *if-then-else*, we now tackle synchronization's *wait* and *signal*.

A primitive synchronizer T consists of a notional internal binary flag - *Open* or *Closed* - that can be changed by two operations: *wait* and *signal*, denoted hereafter by W and S. The restriction to binary behavior implies no loss of generality. A synchronizer must supply the following behavior:

$$\begin{array}{c} S_{out} \\ \uparrow \\ W_{in}\rightarrow \text{T} \rightarrow W_{out} \\ \uparrow \\ S_{in} \end{array}$$

a. A *signal* sets T to *Open*, and passes the *signalling* process;
b. Successive *Ssignals* are the same as a single *signal*;
c. A *Wait* on *Closed* T fails, ie. the *Wait*er is not passed thru;
d. A *Wait* on *Open* T sets T to *Closed*, and passes the *Waiter*;
e. Simultaneous *Waits* on the same $T \rightsquigarrow$ max 1 *Waiter* passes;
f. Simultaneous *signals* on the same T = a single *signal*.

In the above diagram, *waits* enter from the left and exit to the right; similarly, *signals* enter from the bottom and exit at the top. The exclusion of processes over (say) a printer is realized by placing the use of the printer on the W_{out} leg, and thereafter directing the process to perform a corresponding S_{in} before exiting entirely; this arrangement guarantees that processes will use the printer serially (otherwise, output from different processes would be meaninglessly interleaved on the paper record, which is why synchronization is necessary in the first place). More complex examples can be found in any good operating system textbook.

Implicit in such arrangements is the requirement that synchronization be *transparent* to the participating processes: it would be unacceptable for the correct operation of a program to be dependent on whether it "really" waited to acquire some resource because some other process(es) happened to be present. Hence, no information in the Shannon sense is conveyed between two processes via the act of synchronization. Rather, synchronization induces/enforces a phase shift at the inter-process level. This phase shift is expressed in the non-deterministic ordering of the processes as they pass through the synchronizer.

Items e and f refer to situations where there is competition between multiple *Waiters* and/or *Signallers*; see [Manthey 2007].

The first step comes from item b, which in effect says $SS = S$, ie. S must be *idempotent*.

Item d says that WT must succeed if T is *Open*. Therefore initialize T to *Open*, which we can do via item a by setting $T = S$. Item d then reads $WT = WS$, which must be non-zero to succeed.[7]

Item c in effect says (together with item a) that successive *Waits* without an intervening *signal* must fail. That is, $WW = 0$, so W must be nilpotent. So now we know the shapes of both W and S, and very specific ones at that.

Thus a sequence like $SSWSWSST = SWSWST = SWSWS$, and any sequence with consecutive W's yields zero, eg. $WWSWST = 0$.

Process-wise (see figure just below), there is process P_1, which after a sequence of arbitrary irreversible operations X issues the *signal* S, creating a so-called 'synchronization token'; and then there is process P_2 which after a sequence of Y's consumes this token by waiting on it, whereafter P_2 continues, executing Z's (read right-to-left: things begin on the right!):

$$\begin{array}{lll} P_1: & ...X\,X\,X\,S\,X\,X\,X...\hookleftarrow & \\ & \quad\quad\;\downarrow & X, Y, Z \;\text{ are arbitrary irreversible actions} \\ P_2: & ...Z\,Z\,Z\,W\,Y\,Y\,Y...\hookleftarrow & \end{array}$$

Despite the visually implicit timeline in the above two sequences, the $Wait$ in P_2 can occur any time 'before', 'simultaneously with', or 'after' the *signal* in P_1, but unless the *wait* occurs 'after' the *signal*, process P_2 is logically halted at the W. Whichever of these circumstances obtains, the ultimate result is a logically and physically seamless transition from P_1's $SXXX$ to P_2's $ZZZW$. *This sequence too is a process*, process P_3:

$$P_3: \quad ...ZZZWSXXX...$$

The fact that W *must* be nilpotent means that 'whenever' the WS mating actually occurs, it is just as though P_3 occurred seamlessly. An example: when one absorbs a photon in the retina, at that very instant one is exactly connected with the state that generated the S - even if the star that generated the photon has 'long since' disappeared. [8]

P_1 and P_2 are *classical*, in that we imagine them to be deterministic - good old-fashioned Newtonian / Einsteinian processes. [We might think of the state

[7] Initializing T to W (ie. T is initially *Closed*) doesn't work: $WT = WW = 0$, whence SWT also yields zero, which it shouldn't. Initializing T to 1 (which is idempotent) is indiscriminate - *any* W will succeed.

[8] It's pretty limited time travel tho - you only get the single bit of information that the photon carries ... not much of a view!

preparations preceding an actual quantum experiment, which are classical.] P_3, on the other hand, is non-deterministic, because it was precisely $>> P_2 <<$'s *wait* that succeeded, leading to the Z's. *If* however it had happened that some P_4's *wait* occurred ahead of P_2's, P_3's continuation would be entirely different.

This *emergent* non-determinism is old news in computer science, though it is most often noted in the form of unwanted *values* (cf. the interleaved printer output example), rather than the entirely proper non-deterministic *ordering* induced by the serialization as just described.[9] In both cases - *order* or *value* non-determinism - the root is the *asynchrony* of the interaction of two independent processes. Said a bit differently, *if* one is to use *process* as a conceptual primitive, *then* one necesarily must accept into the bargain the consequent, unavoidable emergent non-determinism born of the asynchronous interaction of these same processes.[10] Both non-deterministic values and non-deterministic order are produced by asynchrony. This asynchrony is the very source of QM's non-determinism, and the next Section elucidates its source.

Order-non-determinism forms the *coarse-grained* skeleton of physical non-determinism. Suppose now that one has guaranteed that only a *particular* Wait-continuation will match a given *signal*, so *order* is out of the picture. One still doesn't know what one will get from the measurement, cf. *if-then-else*'s measurement earlier. So within the *order*-skeleton is a second, finer-grained source of non-determinism, *value* non-determinism, induced by the measurements encapsulated in the Signals. For example, the idempotent $-1 + xy + xz$ expresses a value-changing intrusion into the entity $xy + xz$, which in principle "lives its own (reversible) life" both prior to and subsequent to the measurement.

Popping up conceptually, imagine now P_3's form as it evolves into its future. Its sequence of Z's is just shorthand for an arbitrary sequence of idempotents, for example $(1 + a)(1 + b)...(1 + r)$. Being idempotents, each of them can act as a *signal* to some matching *Wait* 'out there'. [It is important that they be idempotents, because this means that the event that the *Wait* is dependent on has actually physically occurred.] Ultimately, if every idempotent in P_3 triggers a *Wait*, and all those *Waits*' continuations do the same, the universe will be populated entirely by utterly non-deterministic processes that look like $(WS)(WS)(WS)...(WS)$ - these W's and S's being notionally distinct. In fact, we see that our classical view of P_1 and P_2 as deterministic processes puts them in an improbable and miniscule minority - namely that minority inhabiting/defining classical 3+1 space-time, *plus* all ordinary sequential computer programs ... which (despite appearances) includes the Internet.

Putting all this together, a sequential process - aka. a measurement sequence - looks like

[9] Both are the source of the most difficult bugs, because they are namely not repeatable; cf. Ullman's fine novel, "The Bug".

[10] It is the *necessity* for exclusion, at *every* step, that dictates that processes be discrete, cf. Planck's constant.

$$(-1+X_n)(-1+X_{n-1})...(-1+X_1) = \prod_n \hat{X}_i, \qquad X_i^2 = 1$$

This is probably all more or less familiar to physicists. But the *computational* reading of the algebra takes the correspondence much further. In this reading [Manthey 2007], the idempotent form $-1 + X$ is identified as the primitive synchronization operation signal(X), understood to mean "*signal* the occurrence of the event/state X".

Signal's complementary primitive is wait(X), ie. wait for the occurrence (signal) of event X. It is critical to understand that this waiting is not *polling*, ie. that the waiting process is constantly and actively checking to see if X has occurred yet, aka. *busy waiting*. Busy-waiting turns out to be a quite untenable view in an asynchronously concurrent universe because everyone spends most of their "time" polling each other, so something subtler is necessary. A careful analysis [Manthey 2007] reveals that the computational concept of wait(X) must be mapped, speaking now algebraically, to some nilpotent $\omega \in \mathcal{G}$, $\omega^2 = 0$.

Furthermore, quoting [Manthey 2007], "No information (in the strict Shannon sense) is conveyed between two processes via the act of synchronization. Rather, synchronization induces/enforces a phase shift at the inter-process level. This phase shift is expressed in the non-deterministic ordering of the processes as they pass through the synchronizer." So synchronization *per se* is memoryless - exactly what happened, exactly what will happen, is spread out in the phase structure of the entire concurrent computation. It is also, indirectly, the means by which non-determinism enters into our model.

In physics, nilpotents supply the causal - and energy conserving - connection between discrete physical events. Wait's play the corresponding role in the synchronizational context - causal connection and conserving information between computational events. Nilpotents are also irreversible, and since any irreversible $E \in \mathcal{G}$ *must* have an idempotent factor, we can reverse the theorem and derive our ω's from our idempotents.

We can derive ω's form by considering two consecutive events $\hat{U};\hat{V}$, forming the process $\hat{V}\hat{U}$. We will insist, now speaking computationally, that $\hat{V}$ *never* occur before $\hat{U}$, ie. the actual process must specify that $\hat{V}$ must always *wait* (ω) for $\hat{U}$. That is, we want $\hat{V}\hat{U} = \hat{V}\omega\hat{U}$. Rewriting the lhs as $\hat{V}\hat{U} = \hat{V}\hat{V}\hat{U}$ and expanding,

$$(-1+V)(-1+U) = (-1+V)(-1+V)(-U)(-1+U)$$

we find that $\hat{V}\hat{U} = \hat{V}(U - VU)\hat{U}$, and indeed $\omega = U + UV$ is nilpotent so long as U and V anti-commute.[11] Computationally speaking, anti-commutativity means "independent of each other", as in the practice of orthogonal software

[11] Note that we could instead have written $\hat{V}\hat{U} = \hat{V}\hat{U}\hat{U}$, which leads to $\omega = -V - UV$. This corresponds to the so-called advanced solution, whereas $\hat{V}\hat{V}\hat{U}$ corresponds to the retarded solution.

design, which focuses on ensuring that changes to one module do not affect another; or as in "asynchronously concurrent"; or both, as here.

Processes like $\hat{V}\hat{U}$ are exactly the processes covered by Turing's model of computation, and since entities like $\hat{U}$,$\hat{V}$ are the *projectors* of U, V respectively (so-called *measurement operators*), they are also the observational bedrock of quantum mechanics. The key property of such processes - irreversible sequentiality - makes them purely *time-like* processes. It is ultimately this time-like property that allows Penrose to conclude [Penrose 1989] that computational processes cannot capture all the phenomena that quantum mechanics has to offer, among which is entanglement, which is fundamentally space-like.

Summing up, we have seen that Turing's sequential model of computation maps to *products* of idempotents (*signals*) and nilpotents (*waits*) in geometric algebra. Products are inherently sequential structures[12] so the fit between Turing's view and that of algebraic products is exact and tight. That leaves sums ... Sums express true concurrency! Since true concurrency lies outside of any sequential theory of computation, this marks the escape path out of Turing's Box. We must ourselves erect a theory for it, and *the resulting general theory of distributed systems turns out to be a computational theory of awareness.*

4. The Coin Demonstration

By *awareness* I mean the *experience* that one is a coherent entity, even whilst in profound appreciation and contact with one's surround. *Self*-awareness is generally larger and mostly unconscious relative to our normal waking (ie. ego-) consciousness. Awareness, *sans* the unitary feeling, is said to accompany certain very deep meditative states. The mechanisms presented in the following allow all of these. [13]

From a "systems" point of view, awareness, being un-localized, looks like some kind of *distributed* computation. A distributed computation consists of many more or less independent processes that *together*, with little to no centralized control, nevertheless produce globally coherent behavior. Examples abound in Nature, from beehives and anthills to ecologies, and from molecules and crystals to the quark structure of protons. Other favorites are the schooling behavior of fish and flocks of birds turning *en masse*.

[12] This includes parenthesis trees, which translate directly to a partial order semantics, which "interleave" separately-sequential process events. Thus the potentially concurrent $f(g+h)$ is mapped to $f(g,h)$, where f, g, h are sequential processes. This approach, which successfully maintains Turing's truths to practical effect, is endemic in the formal theory of concurrent computational systems.

[13] Note that "awareness of awareness" has two interpretations, depending on whether the second awareness is internal or external. Both interpretations are valid, and the *Topsy Test for awareness* [Manthey 2007] requires both. Note also that this classification of "awareness" and "consciousness", while workable, is *very* crude compared to eg. Tibetan and Vedantic observations, whose often flowery language is actually *very* precise, but also often defines differing schools of interpretation.

However, the difference between these systems and *awareness* is that awareness is not material - it has no *substance* - and yet it nevertheless seems to possess agency, even though its coherence is ineffable.

The only general purpose concept (that I can think of) that matches this description is a *wave*. A wave, to be a wave, is an extended affair. I like to say that, so to speak, a wave is *everywhere*. The flip side of the wave concept is that, even though it is everywhere, it is also - simultaneously - *nowhere in particular*. In a system that works like a wave, "nowhere in particular" translates to the myriad local micro-changes that *together* make up the wave, just as H_2O molecules' motions (mostly vertical) make up water waves. Awareness *per se* can then inhere in a multi-dimensional wave-like spectrum (is my claim). So, so far so good: *awareness* is wave-like.

Mathematically, to be in a world of waves is to be in the world of Joseph Fourier, who in 1805 proved that (very nearly) *any* function can be *exactly* replaced by a suitable sum of sines and cosines. This was an astounding discovery, and even though it capped several decades of general interest in doing such a thing, his result nevertheless attracted much controversy in its day. Today, it is a ubiquitous - because enormously useful - piece of mathematical and technological furniture.

More to our purpose, however, is the closely related *Parseval's Identity* of 1799, which states that the projection of a function $\mathcal{F}$ onto an n-dimensional orthogonal space *is* the Fourier decomposition of $\mathcal{F}$. Parseval's Identity is a generalization of the Pythagorean theorem to n dimensions. In the n-dimensional coordinate system, $\mathcal{F}$'s current value corresponds to [the tip of] a hyper-hypoteneuse in an n-dimensional hyper-volume, and the projection breaks that hyper-hypoteneuse down into the various pieces along each of the dimensions that go into its construction.
To construct an n-dimensional volume, begin with an ordinary plane right triangle with sides a and b. Reflect this triangle on its hypoteneuse, forming a rectangle with sides a and b, area ab, and diagonal $d = \sqrt{a^2 + b^2}$. Next, lift this rectangle c units vertically to make a rectangular volume abc. Its diagonal is $d = \sqrt{a^2 + b^2 + c^2}$ and this sum-of-squares symmetry continues as we make a 4d cuboid, then 5d, etc.

At the same time, going back to the starting right triangle, we can also express the sides a and b as $a = cos\,\theta$ and $b = sin\,\theta$, where θ is the angle between a and the hypoteneuse. And now all becomes clear: substituting these sine and cosine equivalents for $a, b, c, ...$ up through the dimensions will yield, for the n-dimensional hypoteneuse (= the current value of the function $\mathcal{F}$, whose projection we began with), a big sum of ... sines and cosines, ie. Fourier's world.

So the world of waves and the world of orthogonal coordinate systems are the *same* world. It is in the latter that we will connect to computation. The connection is this: let each dimension correspond to the state of some process, where all these processes $a, b, c, \ldots, ab, ac \ldots, abc, \ldots$ are notionally independent (think *orthogonal*), though interacting otherwise freely and concurrently; these will be interpreted as elements of a graded vector algebra. [14]

Looking at the ongoing Heracletian flurry of process-state evolution in such a system, the high frequency Fourier bands correspond to short-term, fine-grained details, and low frequency bands to long-term symmetries and global developments. These cross-summed Fourier bands constitute the contextual world of *qualia* - the *feeling* of (eg.) redness *vs.* the optical frequencies detected by individual retinal cells.

And so we see that "distributed" behavior - ie. processes $a, b, c, \ldots, ab, ac \ldots, abc, \ldots$ all running (quasi-)independently - corresponds to wave-like behavior. In the physical world, various constraints (eg. conservation laws, entropy) rule out certain process behaviors as impossible or meaningless, and give form to the free-for-all that is the remainder. In addition, as the system self-organizes, it will - if it has sufficient complexity - learn ways to make itself transparent or reflective to those waveforms that are harmful to it; and complementarily, ways to absorb information and to promote its own further existence via energy-consuming reaction (in this connection, see [England 2014]).

The result is the regularities - short, medium, and long term oscillations - that we, and *any* awareness, will (indeed, must) experience. As a corollary, it is very likely that awareness is not possible if the surround is too unstable [2,3]. This is often seen in visualizations of chaotic systems, where there will be a stable oscillatory behavior for a while, which then suddenly disappears, to be replaced by state transitions with no apparent pattern at all.

We see also that *awareness*, being a wave, is an emergent, collective phenomenon, with nothing scientifically mysterious about it. The mystery is in the experience of it. The following ***Coin Demonstration*** clarifies.

Act I. *A man stands in front of you with both hands behind his back. He shows you one hand containing a coin, and then returns the hand and the coin behind his back. After a brief pause, he again shows you the same hand with what appears to be an identical coin. He again hides it, and then asks, "How many coins do I have?"*

Understand first that this is not a trick question, nor some clever play on words - we are simply describing a particular and straightforward situation. The best answer at this point then is that the man has "at least one coin", which implicitly seeks *one bit* of information: two possible but mutually exclusive states: *state1* = "one coin", and *state2* = "more than one coin".

[14] In the geometric (Clifford) algebra over $\mathbb{Z}_3 = \{0, 1, -1\}$ that I use, 1-vectors like a, b are processes with one bit of state, ± 1, whence an m-vector has m bits of state. For concurrent processes a, b write $a + b$; when a, b interact write ab; ab too is a process with external appearance ± 1 (aka. *spin*). And so on. NB: $ab = -ba \cong \sqrt{-1}$.

One is now at a decision point - *if* one coin *then* doX *else* doY - and exactly one bit of information can resolve the situation. Said differently, when one is able to make this decision, one has *ipso facto* received one bit of information.

Act II. *The man now extends his hand and it contains two identical coins.*

Stipulating that the two coins are in every relevant respect identical to the coins we saw earlier, we now know that there are *two* coins, that is, *we have received one bit of information*, in that the ambiguity is resolved. We have now arrived at the demonstration's dramatic peak:

Act III. *The man asks, "Where did that bit of information come from?"*
Indeed, where *did* it come from?! [15]

The bit originates in the *simultaneous presence* of the two coins - their ***co-occurrence*** - and encodes the now-observed *fact* that the two *processes,* whose states are the two coins, respectively, do not exclude each other's existence when in said states. [16]

Thus, there is information in (and about) the environment that *cannot* be acquired sequentially, and true concurrency therefore *cannot* be simulated by a Turing machine. Can a given state of process a exist simultaneously with a given state of process b, or do they exclude each other's existence? In concurrent systems, *this* is the fundamental distinction.

More formally, we can by definition write $a + \tilde{a} = 0$ and $b + \tilde{b} = 0$ $[\tilde{} = not = minus]$ meaning that (process state) a excludes (process state) $\tilde{a}$, and similarly (process state) b excludes (process state) $\tilde{b}$. [17] Their *concurrent* existence can be captured by adding these two equations, and associativity gives two ways to view the result. The first is

$$(a + \tilde{b}) + (\tilde{a} + b) = 0$$

which is the usual excluded middle: if it's not the one (eg. that's $+$) then it's the other. This arrangement is convenient to our usual way of thinking, and easily encodes the traditional *one/zero* (or $1/\tilde{1}$) distinction. [18] The second view is

$$(a + b) + (\tilde{a} + \tilde{b}) = 0$$

---

[15] [Think about it! Where *did* that bit come from? Thin air?]

[16] Cf. Leibniz's indistinguishables, and their being the germ of the concept of space: simultaneous states, like the presence of the two coins, are namely indistinguishable in time. Co-occurrences are *bosons* in physics-speak (whence *sequential* computation is fermionic).

[17] This is the logical bottom, and so there are no superpositions of $a/\tilde{a}$ and $b/\tilde{b}$: they are 1d exclusionary distinctions . Superposition first emerges at level 2 with ab via the distinction *exclude* vs. *co-occur*.

[18] Since $\tilde{x}$ is not the same as $0x$, an occurrence $\tilde{x}$ is meaningful; in terms of sensors, $x/\tilde{x}$ is a *sensing* of an externality of x, not x itself. That $\tilde{x} \neq 0x$ was also Dirac's innovation, leading to the concept of anti-matter.

which are the two superposition states: either both or neither.

The Coin Demonstration shows that *by its very existence*, a 2-co-occurrence like $a+b$ contains one bit of information. Co-occurrence relationships are *structural*, ie. *space-like*, by their very nature. This *space*-like information (vs. Shannon's *time*-like information) ultimately forms the structure and content of the Fourier bands, eg. the set {all 2-vectors}. See [Manthey 2013] for the mathematics.

Sets of m-vectors - $\{xy\}, \{xyz\}, \{wxyz\}, \ldots$ - are successively lower *under*tones of the concurrent flux at the sensory boundary $x+y+z+\ldots$, and constitute a simultaneous structural *and* functional decomposition of that flux into a graded hierarchy of stable and meta-stable processes. The lower the frequency, the longer-term its influence. [19]
But where do these m-vectors come from?

Act IV. *The man holds both hands out in front of him. One hand is empty, but there is a coin in the other. He closes his hands and puts them behind his back. Then he holds them out again, and we see that the coin has changed hands. He asks, "Did anything happen?"*

This is a rather harder question to answer. [20] To the above two *co*ncurrent *exclusion*ary processes we now apply the *co-exclusion inference*, whose opening statement is: a excludes $\tilde{a}$, *and* b excludes $\tilde{b}$, *whence* $a+\tilde{b}$ excludes $\tilde{a}+b$ and conjugately, $a+b$ excludes $\tilde{a}+\tilde{b}$. This we have just derived.

The inference's conclusion is: *Therefore, ab exists.* The reasoning is that we can logically replace the two *one*-bit-of-state processes a, b with one *two*-bits-of-state process ab, since what counts in processes is sequentiality, not state size, and exclusion births sequence (here, in the form of *alternation* between the two complementary states). That is, the existence of the two co-exclusions $(a+\tilde{b}) \mid (\tilde{a}+b)$ and $(a+b) \mid (\tilde{a}+\tilde{b})$ contains sufficient information for ab to be able

[19] Christopher T. Kello , Brandon C. Beltz , John G. Holden , Guy C. Van Orden: *The Emergent Coordination of Cognitive Function (2007)*. "Abstract: $1/f$ scaling has been observed throughout human physiology and behavior, but its origins and meaning remain a matter of debate. Some argue that it is a byproduct of ongoing processes in the brain or body and therefore of limited relevance to psychological theory. Others argue that $1/f$ scaling reflects a fundamental aspect of all physiological and cognitive functions, namely, that they emerge in the balance of independent versus interdependent component activities. In 4 experiments, series of key-press responses were used to test between these 2 alternative explanations. The critical design feature was to take 2 measures of each key-press response: reaction time and key-contact duration. These measures resulted in 2 parallel series of intrinsic fluctuations for each series of key-press responses. Intrinsic fluctuations exhibited $1/f$ scaling in both reaction times and key-contact durations, yet the 2 measures were uncorrelated with each other and separately perturbable. These and other findings indicate that $1/f$ scaling is too pervasive to be idiosyncratic and of limited relevance. It is instead argued that $1/f$ scaling reflects the coordinative, metastable basis of cognitive function." See Riemann Fever for the undertone calculation.

[20] What makes it tricky is that if at the same time as the man hides the coin he has shown you, you walk around to his back side (be careful how you do it), then it would look to you like nothing happened at all, *vis a vis* the coin, when he shows it again: it's still in the same place relative to you.

to encode them, and therefore, logically and computationally speaking, ab can rightfully be instantiated.

We write $\delta(a+\tilde{b}) = ab = -\delta(\tilde{a}+b)$ and $\delta(a+b) = ab = -\delta(\tilde{a}+\tilde{b})$, where δ is a co-boundary operator (analogous to integration in calculus); derivatives ∂ (defined as eigen-forms) do the opposite, $ab \overset{\partial}{\to} a+b$. A fully realized ab is, we see, comprised of two *conjugate* co-exclusions, a sine/cosine-type relationship. Higher grade operators abc, $abcd, \ldots$ are constructed similarly: $\delta(ab+c) = abc$, $\delta(ab+cd) = abcd$, etc. This hierarchical construction, which is thermodynamically favored, is the subject of the next section; see [4, §8 "The Bit Bang"] for further details.

We can now answer the man's question, *Did anything happen*? We can answer, "Yes, when the coin changed hands, the state of the system rotated 180^o: $ab(a+\tilde{b})ba = \tilde{a}+b$." We see that one bit of information ("something happened") results from the alternation of the two mutually exclusive states. [The transition $a+b \overset{\delta}{\longrightarrow} ab$ is in fact the basic act of perception, called the *first perception*, subsequent meta-perceptions being derivative.]

The occurrence here of 180^o is worth noting: it corresponds to an *inversion* of the state, ie. a reversible change, a *space-like reaction*. The physical analogy is a wave reflecting off a wall, eg. a tethered string. The instigating change from the boundary thus "reflects" off the top of the graded action hierarchy. Untethered, the reflection is 90^o, which corresponds to inverting just one of the boundaries, and is a *time-like reaction*.

With the co-exclusion concept in hand, we can now add a refinement to the idea of co-occurrence. Let S be the space of all imaginable expressions in our algebra $\mathcal{G}$. Thinking now computationally, this means that they are all "there" at the same time. That is, *S is the space of superpositions*, of all *imaginable* co-occurrences of elements of our algebra *$\mathcal{G}$ all at the same time*. Let then G be the space of *actually occurring* (but still space-like) entities, which means no co-exclusionary states allowed. When things move from S to G, superposition is everywhere replaced by reversible alternation *a la* $\pm ab$, ie. G is a sub-space of S..

In less abstract terms, we could say that (wave-world) S corresponds to *imagination*, that (wave-world) G corresponds to the *awareness* of actual possibilities *vis a vis* the surround - and finally, that an *Awareness's* reaction to the surround, via its changes to the boundary $a+b+c+\ldots$, projects G's possibilities (the "causal potential" Ψ) down into grounded action in external, *material* reality.

The result is a flurry of irreversible actions at the sensory boundary, constituting the time-like aspect of the Awareness' reaction. *Crucially*, these action are *guaranteed* to occur in their correct spatial context because they have all been filtered through a hierarchy of tauquernions that constitutes the very structure in which all events occur.

Speaking loosely, intuition and learning are captured by δ, and thought and action by ∂. The various specialized modules of the brain reflect different particular organizations of the functionalities described, and there is good reason to believe that alpha waves correlate with space-like cognition, and hence that the more rapid theta waves correlate with time-like or sequential cognition - speech, planning, etc.

Returning to Parseval's Identity, we see that the key (to being able to invoke it, thus getting wave-particle duality, and thus capturing the dual un/localized nature of awareness) is to organize the flux of changes at the boundary using the distinction *co-occur vs. exclude,* because in so doing, we can then use co-exclusion (= co-boundary operator δ) to perform a hierarchical lift/abstraction, which abstraction is again orthogonal to its components. The orthogonal space so formed allows the application of the Identity. The resulting (novelty-generated) increase in the dimensionality of the orthogonal space increases the complexity and temporal reach of subsequent responses to the surround, and simultaneously the scope of the *Awareness* itself, which inheres in the adaptive self-resonant wave aspect/*experience* of S and G.

5. Hierarchy as Computation

The concept of *hierarchy* - the controlled reversible hiding of information that supplies both context and locality - is of central interest in Computer Science because it is the key to achieving and maintaining conceptual control over very complex software creations. An example of this is the success of object-oriented programming languages (C++, Java, etc.), which offer the programmer sophisticated tools for building complex hierarchies of software objects, which hierarchies then implicitly funnel higher level functionality down to more detailed levels. Contemporary computer systems would be impossible to implement and maintain without these tools.

However, if one looks "under the hood" at what is actually going on at run-time, all of the programmers' fancy hierarchical constructions have been utterly flattened - by the compiler - into long sequences of nested function calls. The sophisticated object-oriented hierarchical structures at the programming language level are actually just syntactic sugar, disguising the fact that the only *real* hierarchy concept in contemporary software thinking is good old, tried-and-true function composition: $y = f(x)$ and $z = g(y)$ combine into $z = g(f(x))$, which says first do f, then do g.

Or instead of "objects", look at "remote procedure call", or "agents": in the end, virtually *everything* is made out of function composition, with maybe a little memory on the side. Thus, as a hierarchy concept, *function composition* *is fundamentally sequential*, that is, it really isn't hierarchical at all. Rather, it's fundamentally flat: we ourselves *design* the composition sequence, and in general we *impose* it on what is (according to Turing completeness) always ultimately a *sequence* in the real world [and not forgetting relativity!].

This state of affairs has *greatly* hindered the construction of highly concurrent, physically and conceptually distributed systems in software, whilst Nature's ability to blithely do this all the time has remained a mystery. The computational hierarchy concept offered in this paper - described here using homology theory's boundary / co-boundary operators ∂/δ - exposes how Nature might do this, and thus represents a major advance in the sciences of computation and information. For example, computations using it are space-like rather than time-like in their behavior, they are inherently self-organizing, and they can even be self-aware, cf. the *Topsy Test* for awareness.

As just sketched above, higher grade operators abc, $abcd$,... are constructed from lower grade operators via the co-boundary operator δ: $abc = \delta(ab + c)$, $abcd = \delta(ab + cd)$, etc. Clearly this process, which can variously describe thermodynamic evolution, organic growth, and conceptual learning, can continue indefinitely, creating increasingly complex global operators. Unfortunately, their very complexity, both individual and collectively intertwined, means that the we quickly lose control of the system's semantics ... What will it do in novel situation X? Can it accomplish goal Y? This puts a very real lid on the useful height of the growth hierarchy.

Fortunately, geometric algebra's very structure supplies an elegant solution to this problem: its semantics loop $mod\,4$. That is, $1^2 = +1$, $a^2 = +1$, $(ab)^2 = -1$, $(abc)^2 = -1$, $(abcd)^2 = +1$, $(abcde)^2 = +1$, ... and one sees that the sign pattern $++--++--$ ... is that of the powers of $i = \sqrt{-1}$. The table below shows one way[21] to then map potential higher grades to grades zero, one, and two, for then to repeat the δ-build-up process:

pairs	δ*(pair)*	*new level*		
$3\,mod\,4 + 3\,mod\,4$	$\rightsquigarrow$	6	$= 2\,mod\,4$	$\searrow$
$2\,mod\,4 + 3\,mod\,4$	$\rightsquigarrow$	5	$= 1\,mod\,4$	$\searrow$
$2\,mod\,4 + 2\,mod\,4$	$\rightsquigarrow$	4	$= 0\,mod\,4$	$\searrow$
$1\,mod\,4 + 2\,mod\,4$	$\rightsquigarrow$	3	*charge*	. $\downarrow$
$1\,mod\,4 + 1\,mod\,4$	$\rightsquigarrow$	2	*spin*	$\xleftarrow{6}$ $\downarrow$
$0\,mod\,4$ & $1\,mod\,4$	$\rightsquigarrow$	1	*existence*	$\xleftarrow{4,5}$ $\downarrow$

This semantic folding can be elaborated to produce arbitrary growth patterns using a new [TLinda programming language] construct called a *corm*, which, like its botanical namesake, is a self-replicating "root bundle" that specifies the growth to come.

[21] The algebra contains many such semantic symmetries, eg. with $mod\,8$, the two $mod\,4$ halves generate slightly different spaces; also octonions and other vector-algebra variants; see Wikipedia: Classification of Clifford algebras. It is very possible that various specialized brain modules implement such "non-standard" symmetries, these being particularly well-suited to their function. The overall structure via the table is an iterated $U(1) \times SU(2) \times SU(3) \times SO(4)$.

Speaking of growth, one should recognize the rapidly overwhelming $O(2^{2^n})$ growth in the potential size of the node space that is the price of recording every experience (eg. in the form of this hierarchy). There are four saving graces for this. The first is that a given co-exclusion is recorded *only once*. The second is that if nothing else, this hierarchical structure is, from an information storage and retrieval point of view, minimal.[22] The third is that even though *any* experience can be captured, a given system cannot possibly *actually* experience them *all* - this would take many universe-lifetimes! So only a tiny tiny fraction of the space will ever be used. The fourth is that one can always prune the exfoliation, eg. as plants, brains, nervous systems etc. do; or as statistics and/or other considerations indicate.

It should be noticed that the actual hierarchical structure at run-time is - ignoring the "wrap-around" at the topmost nodes, described below - a rooted acyclic lattice. That is, most nodes will have several parents, and the sensory bubble-up is roots to leaves, and the reaction's trickle-down is leaves to roots. Since all nodes are combinations of the grade 1 sensors at the system boundary (the roots), the latter are the ultimate parents of all later-created child nodes. So, unlike most tree-like structures in computer science, which are drawn downwards with the root at the top, the present co-exclusion based hierarchy grows from the bottom up. We refer to this lattice structure as "the hierarchy".

The algebraic prescription of execution behavior can be taken further using a form called an inner auto-morphism:

$$AXA^{-1} = X' = A^{-1}XA$$

The examples to follow use the form $A^{-1}XA$ for expository reasons, wherein the rightmost A is treated as the current state of the node in question, to be modified by the up-bubbling sensory X by multiplication on the left. An upgraded X is bubbled further up and eventually inverted, namely via some A^{-1}, again via multiplication the left, whence $X \mapsto X'$. The evolving state X' then trickles back down, meeting leftmost A^{-1}s, each of which reduces X's grade and trickles it further down. Eventually X' will meet an effector (ie. grade 1, at the sensory boundary), therewith completing the system's reaction to the input.

An example of an inner auto-morphism is

$$ba(a+b)ab = -a - b$$

[which is the same as

$$ab(a+b)ba = -a - b$$

[22]Modulo clever compression schemes.

because for simple m-vectors like $ab, abc, ...$, their inverses equal their reverses:

$$AA^{-1} \quad = \quad AA^{\dagger} = aa = abba = baab = abccba = cbaabc = abcddcba \ldots = +1\,]$$

The hierarchy is built out of inner auto-morphisms $A^{-1}XA = X'$ such that $A = \delta X$ for the bubble-up, and $X = \partial A$ on the trickle-down:

$$\begin{aligned}
ba(a+b)ab &= -a-b \\
cba(a+bc)abc &= a+bc \\
dcba(ab+cd)abcd &= ab+cd \\
dcba(a+bcd)abcd &= -a-bcd \\
edcba(ab+cde)abcde &= ab+cde \\
edcba(a+bcde)abcde &= a+bcde
\end{aligned}$$

These transformations hold for all variants of the X-form, eg. $cba(b-ac)abc = b - ac$. The $mod\,4$ grade-cycling means that these transformations cover all eventualities.

The first and third lines in the table suggest trying input $(a+b)+(c+d)$ on the hierarchy $abcd = \delta(ab+cd) = \delta(\delta(a+b)+\delta(c+d))$.

The first thing that happens is that the state ab will be operated upon by *new* sensory information, namely the bubble $a+b$ (whence it follows that the immediately prior state was $-a-b$). For this bubble:node collision - multiplying, as prescribed, on the left - we write $(a+b)ab$, and similarly for $c+d$.

Suppose as a preliminary example that ab is the top of the hierarchy. It will then (by default) simply turn the bubbled-up change into a goal to change it back, ie. maintain equilibrium by completing the oscillation. So write therefore $ba(a+b)ab$ whose result $-a-b$ becomes the two goals, $+a \rightarrow -a$ and $+b \rightarrow -b$. Equilibrium has been maintained.[23]

Now expand the example so that $abcd = \delta(ab+cd)$ is the top of the hierarchy. Then the algebra tracks what happens as follows:

1. Two new state bubbles $(x+y)$ operate on their parents xy, making new internal xy states: $(+a+b)ab\ +\ (+c+d)cd = (-a+b)\ +\ (-c+d)$; and forming the grade 2 bubble $-ab-cd$.
2. The bubble from (1), $-ab-cd$, reaches $abcd$: $(-ab-cd)abcd$.

[23] One can also, of course, simply ignore the bubble. But given the fundamental wave-like nature of the paradigm, inverting the change seems the obvious best ***default*** choice. Clearly, some top-most nodes will require human approval rather than defaulting.

3. Reflect from top $\Rightarrow$ bubbles become droplets: $dcba(-ab - cd))abcd = ba + dc$. [Recall that $-ab = ba$.]

4. The two droplets from (3) hit their respective xy's: $ba(-a+b)+dc(-c+d)$ & then trickle further

5. down, having now become $-a - b - c - d$, which, being on the sensory boundary, are effector commands, ie. $+a \to -a$, $+b \to -b$, $+c \to -c$, $+d \to -d$.

Thus once again the change is complemented, the oscillation completed, and equilibrium maintained.

The hierarchical execution regime we have just described is a pure space-like computation. It reacts to its environment with utter immediacy - what *can* happen is what *does* happen ... and *only* that - in an uninterrupted, purely entropic and subjectively timeless *Now*.

The computation's only "purpose" is to maintain equilibrium with its surround, all the while growing new structure based on its sensory experience. But the joker here is that while the experiences have been faithfully catalogued in hierarchical growth, the environment in which they occurred is constantly changing. So an action-reaction cycle as dilineated above that once worked may well not work the next time.

As a result, there will inevitably be a backlog of uncompleted equilibrium-attempts which will continue to drive behavior on an opportunistic basis. In fact, without more structure, *What* the computation does, *how* it actually attains its cthonic goals[24] will be purely dependent on happenstance, on the surround just happening to be in a favorable configuration. To be able to pursue a novel goal, various *pieces* of these earlier experiences must be strung together, on the fly, to cope with novel variants of previously experienced situations.

This organizing, this stringing together demands further structure, which structure together with the hierarchical dynamic described above are the principle content of the patent "Space-like Computation for Computational Engines" [ ], which issued in the USA in April, 2020. What is namely needed is a way to recruit relevant actions that, taken in appropriate order, will accomplish the goal with (generally novel) sequence(s) of actions.

It works like this. We have seen how the bubble-up of an impulse is matched by a subsequent complementary trickle-down. When an action, say xy, receives a sensory bubble from below, it does two things. The first is to ascertain if it can satisfy the bubble by itself; if so, the xform is carried out by trickling the appropriate sub-goals down as usual.

If however it cannot satisfy on its own, it can find help via a technique called *back-chaining*. The idea with back-chaining is to realize that instead of planning from current state *forward* to the desired state, there are so many possibilities

[24] *Cthonic* [Wikipedia]: "In analytical psychology, the term chthonic was often used to describe the spirit of nature within".

that it's much more efficient to plan from the goal state *backwards* to whatever elements of the current state can begin the solution.

For example, suppose we have an action xy, which wishes to transform to $\tilde{x}y$ (ie. the bubble contains $\tilde{x}$, which means that we want $\tilde{x} \to x$). So we want to find an action that inverts x, say cx, whence $(cx)(xy) = cy$. But maybe cx needs help to invert c (cuz for any action, *both* components must flip), leading via backchaining to e.g. bc, so now we have (bc)(cx)(xy). But maybe bc needs help too, leading to e.g. ab which can complete w/o help, so now we have (ab)(bc)(cx)(xy) = a(bb)(cc)(xx)y = ay. [If also y needs help to flip, the same backchaining occurs simultaneously for it.

The TLinda code for this is almost simpler than the explanation: ——————————————-

Genuine *purpose* - the *conscious* pursuit of a goal - demands further structure [Manthey 2016].[25]

Nevertheless, some local scalar mechanisms can still be applied, especially by living systems, to optimize their responses. The most obvious of these is to count uses of nodes, and use these counts to choose from the possibilities that have been served up by the hierarchy. This clearly works, since our own nervous systems do this, and it has long been the basis for learning in neural nets, cf. Bayesian learning.

The disadvantage is that as the counts build up, what is initially a behavior by choice can become first a habit and then eventually a rigidity, while at the same time ignoring other possibilities, which are still being offered by the hierarchy. The result is a kind of over-specialization or tunnel vision. Thus, while this kind of innovation makes evolutionary sense, it is unnecessary when there is no speed penalty associated with always considering all possibilities.

Opposite to specialization is generalization, where this term is used to mean the passage from a particular given state - in the form of some set of hierarchy nodes that are simultaneously active - to another, larger space of which the original particular state is a sub-space, and this sub-space is highly correlated with the other sub-spaces that altogether make up the larger space.

Recalling that Parseval's Identity applies to our algebra, it is easy to imagine how the phenomenon of generalization is accomplished, namely via resonance. That is, the states that resonate with the original state are those that are most similar, the most closely correlated in the frequency domain. Frequency in turn translates to a vector's *grade*. That is, a particular state - say $a+c+ab+acdef$ - will resonate, possibly over many $mod\,4$ levels, with all states containing $u+v+uw+uvxyz$. These latter, individually and together, form the generalization. Indeed, one can imagine a purely wave-based implementation of the hierarchy; and one might well find working examples in cells or insects. Analogy and metaphor are linguistic examples.

[25]The TLinda source code for everything (seven pages!) will be revealed when the the patent issues, ?*primo* 2018.

As the hierarchy grows, the time for a bubble to reach the top, and the corresponding reaction to trickle all the way back down, is proportional to the number of hierarchical levels traversed, ie. the reaction time is generally logarithmic in the overall size of the hierarchy. So the execution time of a hierarchical space-like computation scales very well indeed. It's also fast, because the *vast* majority of nodes is always logically (and automatically, via the co-ex basis) blocked, and activity tends to be localized and coherent. The traditional AI "searching" exercise is done by the bubble-up and trickle-down processes, that is, via δ, search is inherent in the hierarchy's very structure, whence most of the "computation" is pointer-following, and the rest trivial.

To capture the *persistent* aspect of awareness - it's present whenever I am - I postulate that it is a *resonant state* - a self-maintaining and very complex oscillation - where the spectrum of this resonance will vary, eg. according to the properties of the surround wherein the awareness is emplaced. This resonant state rests on, and derives from, the brain's neural substrate, but *nevertheless*, the mathematical space in which the resonant state exists is *outside* of (and *much* larger than) the mathematical space defined by individual neural function, because it is a co-occurrence (ie. superposition) state. The resonant state deriving from the various individual oscillations is of size $\mathcal{O}(2^n)$, versus the default $\mathcal{O}(n \times n = n^2)$.

That is, algebraically, an EEG-type wave of brain activity is a *scalar* sum of neural activity, treating all neurons as being in the *same* dimension. But as the Coin Demonstration shows, a close analysis of co-occurring processes leads to the conclusion that the processes involved (eg. neurons) lie on *orthogonal* dimensions, which algebraically means that $ab = -ba$. Thus any argument that relies on globalizing the definition of an *individual* neuron's function is flawed. In other words, our aware experience is *usually* an on-going 3d projection of a *much* larger space, of which 3+1d is the *result*, and not the default one-and-only beginning-and-ending place.

So both the materialists (the Pythagorean side of Parseval) and the non-materialists (the Fourier side of Parseval) get their cake, and get to eat it too ... for the price of also being half wrong, ie. for claiming that their story was the *whole* story. From a discrete process and informational point of view, Parseval's *Identity* cements the argument that both stories are correct, simultaneously, all the time ... *all the way down* [Manthey 2013].

6. Awareness is outside of Turing's Box

The foregoing has introduced much novelty, ... much-needed novelty, one could though say. So perhaps a brief recapitulation will be helpful.

We began by showing that the essentials of Turing's concept of computation - *if-then-else* and *wait/signal* - are captured by idempotents and products thereof, all belonging to the geometric algebra $\mathcal{G}$ over $\mathbb{Z}_3 = \{0, 1, 2\} = \{0, 1, -1\}$. The reasoning is very direct, in that Products are inherently sequential, and sequentiality is the core of Turing's model.

We also saw how the processes formed from *wait/signal* connections are inherently non-deterministic. We can conclude that natural processes simply *are* non-deterministic ... the reason and the mechanism are clear and there is *no* mystery to be solved. Thus Everett's idea of explaining non-determinism using the device of ever-forking ("parallel") universes in the face of decisions is seen to be entirely superfluous and lacks empirical content.

We then showed how Sums in the algebra can be construed as co-occurrences of computational events, and, via the Coin Demo, that there is information inherent in co-occurrence, *per se*. The information in complementary co-occurrences can be combined, via the co-exclusion principle, to form a new node that encodes its constituents, creating successive meta-level descriptions of the computation at the level below.

Thus arises the Hierarchy. The nodes of the Hierarchy, ie. m-vector elements of $\mathcal{G}$, are mutually orthogonal, which is interpreted to mean that the processes they both symbolize and encapsulate are nominally independent and concurrent.

Processes being mutually orthogonal also allowed us to invoke Parseval's Identity, which equates every expression in $\mathcal{G}$ to the Fourier decomposition of a corresponding input. This in turn allows us to responsibly claim that the experience of complex, self-arising, self-adaptive resonances in the hierarchy is what we call awareness and awareness of awareness ("consciousness").

The phrase "awareness of awareness" can be understood to mean that there exist two simultaneous resonances, a "lower" and a "higher", in that the higher consciousness is at a higher hierarchical level and therefore oscillates more slowly than the lower. That is, the higher level via its lower frequency is experienced as relatively unchanging, whilst the lower level but higher-frequency processes evolve much faster and "things change". It's easy to see this structure as a sandwich made of *mod4* hierarchy segments, where the meat in the sandwich is the *mod4* phase transition that connects them.

From a subjective point of view, one can responsibly describe the higher level resonance, the *entity* so to speak, as *experiencing* the lower's progress.

This last statement is just one example of how space-like computation directly captures common phenomena that traditional sequential models struggle mightily with. The following introduces more such revealing captures.

We begin with the idea of $3d$ space, with three orthogonal dimensions/axes called a, b, c. Taking a, b, c as vectors, the product ab is the a, b-plane, the product ac is the a, c-plane and the product bc is the b, c-plane. Similarly, abc is a volume.

The set $\{ab, bc, ca\}$ is called a quaternion triple, and they form a mathematical *group*, namely the quaternion group. The sum $ab + bc + ca$, ie. the concurrent existence of these three planes, specifies the *same* point in $3d$ space as the sum $a+b+c$, in that the $3d$ rotation $abc(a+b+c) = ab+bc+ca$!

The quaternions were discovered by W^m Hamilton in 1843, and introduced mathematics to the surprising and liberating concept of anti-commutativity, $xy = -yx$ [ie. x, y are orthogonal = independent asynchronous processes]. Also, $\gamma = a+b+c$ identifies a photon, wherein we see the intimate connection between light and space first uncovered by Einstein.

However, in addition to Hamilton's quaternions, it turns out that there are exactly *two more* quaternion variants lurking deep in $\mathcal{G}$'s and mathematics' shadowlands: the triples $\{ab-cd, ac+bd, ad-bd\}$ and $\{ab-cde, ac+bde, bc-ade\}$. We have dubbed the former *tauquernions* [Manthey 2013] and the latter *tauquinions* [Manthey 2014]. What is special and unexpected about them is that they are simultaneously 3d space *and* irreversible operators!

How can this be? The answer is simply that the space-like rotations described/performed by the tauquernions and tauquinions increase entropy as they project the hierarchy space into $3+1d$, and in so doing, continually re-construct space-time on-the-fly. Algebraically, notice that $(ab+cd)$ can be factored to read $(-1+abcd)ba$, where $-1+abcd$ is idempotent, and $ab+cde$ similarly results in $-1+abcde$.[26]

Besides being representations of the quaternion group, both the tauquernions and the tauquinions also map to a group called *Spin{6}*:

$$\{ab, ac, ad, bc, bd, cd\} \mapsto \{ab-cd, ac+bd, ad-bc\} \mapsto SO(4)$$

$$\{ab+cde, ac-bde, ad+bce, -bc-ade, bd-ace, -cd-abe\} \mapsto SU(3)$$

Spin{6} can in turn map to either $SU(3)$ - the world of quarks, charge and electro-magnetism - or $SO(4)$, which is $4d$ orthogonal space, one version of which is our familiar relativistic $3+1d$. Thus our model of awareness has the basic mathematicsl structure $\mathcal{G}_2 \rightsquigarrow U(1)$, $\mathcal{G}_3 \rightsquigarrow SU(2)$, $\mathcal{G}_5 \rightsquigarrow SU(3)$, and $\mathcal{G}_4 \rightsquigarrow SO(4)$, ie. the Standard Model of physics augmented with *3+1d* space in the form of $SO(4)$!

Recalling the bubble-up/trickle-down example from the preceding section, we see that it was precisely the $2+2$ tauquernions and the $2+3$ tauquinions that were the bricks that formed the Hierarchy.[27] That is, $\mathcal{G}_2$ gives us the circle group and $\mathcal{G}_3$ the quaternion group, ditto $\mathcal{G}_4$ and the tauquernions, and $\mathcal{G}_5$ and the tauquinions: *everything* is geometric !

One can only conclude that *awareness is entirely space-like, and everything is made out of it*. Since the Hierarchy builds from the bottom, this ultimately means[28] that we and the universe are neither separate nor to be distinguished from each other.

[26]NB: The forms $-1 \pm abcd$ and $-1 \pm abcde$ play the role of $+1$ in their respective algebras; -1 is $+1 \mp abcd$ and $1 \mp abcde$ respectively.

[27]See [Manthey 2013] for these and other cases.

[28]...if the mod4 buildup has the gauge property, as is likely.

Furthermore, both the tauquernions and the tauquinions present as *fields*, not particles, so they are well-suited to capturing the more ineffable phenomena of feelings and emotions (respectively).

That is, the phenomenology of our [tauquinion based] emotional system is fundamentally electro-magnetic in character. Common slang supports this conclusion: sexually magnetic, lightning temper, sparks of rage, charged situation, etc. The Vedic chakra system agrees as well: "2" is the hara chakra, which deals with basic distinctions like inside/outside, left/right, up/down, male/female, life/death. And "3" is the solar plexus chakra, our social brain, the one that always and immediately understands and remembers who owes whom in every exchange; and also the seat of vision (cf. quaternions).

The phenomenology of "4", the heart chakra, derives from the tauquernion field and is very different. All of its elements, namely 2-vectors like ab and cd, and 4-vectors like $abcd$ and $efgh$ will, when operating on each other, commute, ie. $PQ = QP$. This is because vectors of even grade always commute, but the implication is that for the tauquernion field, there is ***no*** temporal order even though all "motion" is irreversible (because entropic). The *only* operating distinction in this ***acausal* regime** is co-occurrence. This is the ***Now***.

Add in the facts that (1) a tauquernion triple, being nilpotent, is a strong candidate for the name Higgs, and (2) that tauquernions are entanglement operators, and we have a description of gravity [Manthey 2013]. The gravitic glue that gathers the cosmos is thus identified with the most profound experiences a human is capable of - consciousness, selflessness, altruism, love, and awe. This self-gathering *gravitas*, this *amour* as Newton originally called it, is at the root of human cooperation and unity.

So yes, not only are *We* outside of Turing's sequential box, *We* are *way* out of that box ... *We* are namely made of pure space, and are entangled entirely with the rest of universe.

References & Links

1. Manthey, M. *Synchronization: the Font of Physical Structure*. Proc. ANPA 2007. See TauQuernions.org.
2. Penrose, R. "The Emperor's New Mind: Concerning Computers, Minds, and The Laws of Physics", 1989.
3. Manthey, M. *The Topsy Test for Awareness*. Proc. ANPA 2014. See RootsOfUnity.org.
4.Manthey, M. *Tauquernions* $\mathcal{T}_i$, $\mathcal{T}_j$, $\mathcal{T}_k$*: 3+1d Dissipative Space out of Quantum Mechanics*. Proc. ANPA 2013. See TauQuernions.org.
5. Manthey, M. *Riemann Fever*. Proc. ANPA 2015. See TauQuernions.org.
6.England, J. "A New Physics Theory of Life". https://www.quantamagazine.org/a-new-thermodynamics-theory-of-the-origin-of-life-20140122/
7. Manthey, M. *Space-like Computations in Computing Devices*. USPTO - #PCT/US2016/035885, and other patents pending.
8. Manthey, M. *Electro-magnetism is a Fundamentally Fractal Fibonacci Fourier Field*. Proc. ANPA 2014.