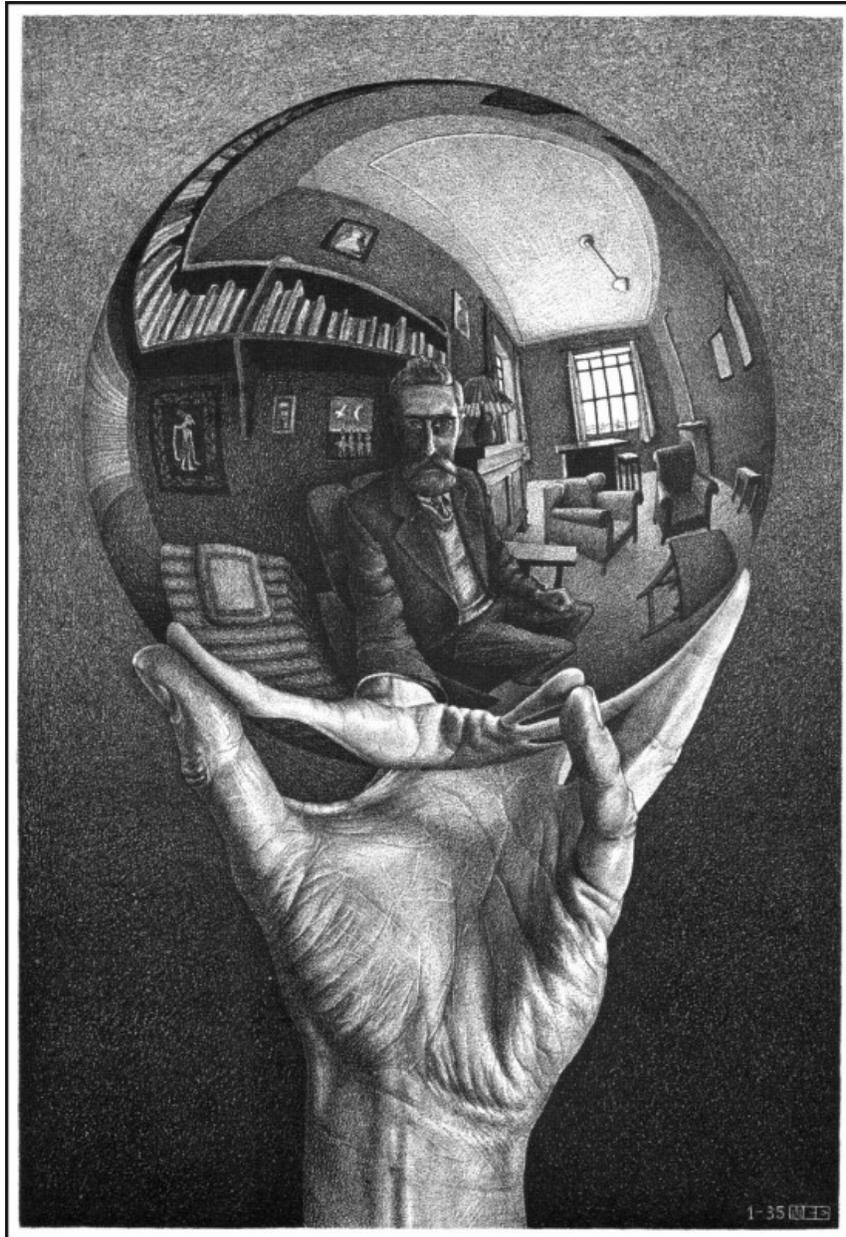# II.

Romanticism and Consciousness (1972)

It is most unfortunate, but the point of this story has been reached where a justification of the expression "Murphy's mind" has to be attempted.

— Samuel Beckett: *Murphy*.

I have long since lost the original, but recall a notebook which commenced with the melodramatic pronouncement:
"The *central* problem is the relation of the mind to Nature."

That wasn't bad, actually.

The two obvious answers are that one (take your pick) contains the other.

The less obvious answer is that both of those answers are correct at once.

(Novalis: "It is all one whether I posit the universe in myself or myself in the universe.")

Bad acid (1967)

(Must I say something about the drugs? would this otherwise be *Hamlet* without the Prince of Denmark? — More like omitting Polonius, I am thinking. In any case —)

To the epistemological adventurer, the hippie always seemed unforgivably bourgeois. He took drugs and acted weird and dressed funny and thought strange things, true. But he *needed* the drugs. That was the crucial difference. — Indeed the really amazing thing was how many people hadn't realized they *had* minds, before they started fucking with them. — Thus it was always obvious that, whatever this was, it couldn't last.

{...}

I recall an expedition to the Haight in the autumn following the Summer of Love: my comrades and I, a company of geeks, wandered around the neighborhood all day, gawking at the outward signs of the New World Order; and then, needing a place to crash, followed a couple of the natives who were trying to make up the rent back to their townhouse. — Which must have been costing them more than the few bucks we ponied up, but the economic mysteries of that era then as now defied rational explanation.

There we found ourselves in the company of the usual motley crew — something like the melting pot of the Hollywood warmovie foxhole: the speed freak with fresh blood on his tracks who explained to me that the essence of Dylan was that he would take any drug, he wouldn't even ask what it was, he didn't give a shit, he'd just shoot it, snort it, smoke it, jam it up his ass in a suppository, whatever, at which I nodded sagely [thinking "What is this bullshit existential bravado?"]

and said Wow man, far out — the couple humping furiously in the back room who barely nodded on their way out to panhandle[1] among the tourists — somebody bitching about Janis and her rowdy juicer friends running around the neighborhood — somebody who gave the definitive one-line review of the latest Airplane album:[2] "It's more of a *meth* sound... ." — They were all still wearing those swinging-London faux-Victorian-dandy outfits like the Stones did but everything was dirty, because nobody bathed, or washed or even changed their clothing, because they had nothing else to change into. It was clear there was already some kind of devolution going on from love and flowers and cosmic consciousness to closed minds and harder drugs, but it was not complete — indeed one of the guys crashing there among us, *mirabile dictu*, was one of Owsley's own distributors,[3] and was carrying an entire gram of LSD in a vial, which we all examined with the utmost curiosity. — So we talked to these characters, cracked wise in our customary fashion, and they all kept laughing at what we were saying. — Finally someone asked "Hey, what are you guys on?" — Here followed an embarrassed pause as we looked at one another. — I wanted to say "DNA", but restrained myself because these were our hosts and it would have been rude to put them down so bluntly, even if the insult would have sailed over their heads. — Instead we performed the geek ritual of self-effacement, mumbled the usual self-deprecating apologies for being brilliant and original, and doubtless missed the opportunity to make our fortunes selling the latest three-letter wonder drug to a witless mob of scientific illiterates.

---

[1] A friend of mine — a far more gifted hustler than ever I could aspire to be — reported that, as an energetic panhandler in the Haight in the Summer of Love, he could net about seven dollars an hour from the tourists; at the time, about five times the minimum wage. (Yes, this guy later sold out, went to law school, and became a Yuppie parasite. We all know the story of the Decline of the West.)

[2] This must have been *After Bathing at Baxter's*.

[3] Really, one look at his pupils and you knew he was telling the truth.

{...}

Though I personally had never had difficulty having uncanny experiences without the aid of any other chemicals than coffee and fatigue poisons, I was nonetheless an early adopter, if then a rapid renouncer, of the abuse of psychedelics. — I dropped acid a few times without serious incident, but then had a slight disagreement with one of my best friends; this somehow became amplified and distorted, and in consequence even though I understood perfectly well that none of it was really his fault or mine, that the drugs had opened the doors of misperception, I couldn't stand to be around him for a year or more afterward.

It seemed obvious that anything that could poison my relationships so easily was not to be trusted; the argument that you could undo it all by repeating the experiment a couple of days later was obviously ridiculous, like claiming your brains would not be permanently addled so long as you kept banging yourself on the head with the same brick. — So by and large I let that wave pass me by.

{...}

But what was philosophically interesting about the experience was that the argument people like Huxley and Watts tried to make from it was completely backwards. They said the hallucinations that came from taking psychedelics proved that quotidian reality was an illusion, and the brain was a crude filter that dulled your senses and prevented you from perceiving things are they really were; from seeing the light that shone forth from the radiant center of things.

But clearly it was just the opposite, a simple matter of cause and effect. The brain was the determinant, and simply twiddling a few knobs on the back of the chassis to adjust some excitation thresholds could make the mental reality that depended on it entirely different.

In the comic books mad scientists had invented pills and potions that could turn you into a giant or make you invisible; Albert Hofman had invented one that could make you see God. — Nothing more succinctly proved the chemical basis of consciousness, the ultimate identity of the mental and the material. — That there was a speculative, a metaphysical chemistry. — A chemistry of the soul.

{...}

Drugs at any rate foregrounded the problem of consciousness: there *was* something you were altering, so what was it?

One thing was that the difference, as it were, between acting in the movie and watching it could suddenly be manifest: you seemed to step outside yourself, to realize that the point of conscious life was something like you were playing yourself, it was a role; you were a simulation of yourself.

There was some dim realization of this going around, but some inappropriate interpretation got hung on it, maybe it was the instinct to moralize — it came off as: what you think you are is phony; it's just you playing a role, it isn't the real you. — The real you, presumably, was something you had yet to discover, and a host of "spiritual guides" (aka bullshit artists) popped up like psilocybin mushrooms to make their fortunes explaining it.

But it was always obvious to me that the ultimate fact was that you *were* this simulation — this act, the game — and that the fact you could perceive this was also part of the game; was wired into it; was intrinsic to it, in fact. This didn't mean that all foundations vanished and reality was illusion — if reality was an illusion, then the distinction between reality and illusion was itself an illusion, any child should have been able to see that — it was just that reality was much more complicated, and therefore interesting, than anticipated. The Self might be an illusion, but it was easy to see that meant the universe was as well, a

complex of them, and that *this made no difference*.[4] — What is, still is, and the point is still to try to understand it.

The converse (to separation) was also true: the difference between watching or fantasizing or listening to a story and living in it could suddenly vanish, so that you couldn't tell which was which. Listening to a song might trigger a series of hallucinations consistent with your being somehow *in* it (clichéd though it must seem, my first such experience came listening to "White Rabbit", and the chessmen really did get up and tell me where to go), taking part rather than merely witnessing. — When Dylan sitting in his loft abruptly begins a verse "The peddler now speaks/To the countess who's pretending to care for him," it's because there's a movie on the television in the background (or he thinks there is), and he's suddenly not watching it from without but from within, he's entered into it. — Among other classics of Stoner Literature, in Pynchon's *Gravity's Rainbow* there is Slothrop's extended fantasy about disappearing down the toilet bowl — brought rather too vividly to the screen by Danny Boyle in *Trainspotting* [1995].

{...}

Here pro forma I should insert the Uncanny Experience anecdote that will be found in all such memoirs — the occasion when drugs did do — *something* — though I never figured out exactly what......[5]

{...}

What would it be like if telepathy were really possible? if you could enter another person's head? — Like this, you had to think: completely different though topologically equivalent, a distorted map,

———————————

[4] Bloom: "Shakespeare invents in *Hamlet* the destruction of any boundary between being oneself and playing oneself." Or recognizes that there never was any such distinction.

[5] Can't bring myself to do it. Too stupid.

the Self viewed in a funhouse mirror. It would be just like taking drugs.

{...}

Spinoza somewhere defines consciousness as the idea of ideas, which is close to what I eventually decided, but this wasn't the proximate cause for the inspiration. — Here I must digress: when I was a freshman I went through one of those phases in which the world seemed ripe to be remade anew — myself along with it — and briefly contemplated correcting my posture, which then as now was atrocious. I made a few vain stabs at trying to stand up straight, posing upright in front of the mirror — predictably this got me nowhere — but then more sensibly thought to compare myself to my natural peer group, the residents of Fleming House. — And then after a brief survey realized, to my considerable amusement, that the three worst slouches in the place belonged to myself and a couple of upperclassmen — call them Chris and Mitch[6] — whom I knew to be the smartest guys in the school. — "This slouch is aspirational!" I declared to myself — laughed —and never gave a thought to my posture again.

(It was here, incidentally, that I first began to realize why therapy so often turns people into assholes: because they don't really change; they only bullshit themselves into thinking they have. — It is like running a foreign emulation program: pretending to be something you are not. — I *might* have changed my posture, but not my nature.)

The two gentlemen in question were both notorious acidheads, and like all the rest of us at the time expended much effort theorizing about the drug experience.

---

[6] Both subsequently had distinguished academic careers, and why pretend I can cloud their reputations by naming them here. But still.

Most of the commonplaces in circulation were predictably worthless, but I recall Mitch telling me[7] that Chris had said any mathematical model of the mind would have to be isomorphic to the group of its own automorphisms.

He had based this on some particular type of acid experience with which I wasn't familiar, but the idea was so obviously beautiful that it didn't matter. Out of the countless hours of stoned silly bullshit I had endured it stood out with an electric clarity. Everything else I'd heard I had ignored. But this I stole at once.

{...}

Later it occurred to me that a formal equivalent closer to the mark was the apparently paradoxical property of the lambda calculus, which Scott[8] was able to explain, that any model would have to have the property (*prima facie* impossible in set theory) that it was isomorphic to the ensemble of functions from the model to itself.

{...}

This was not the best drug-inspired contribution to the philosophy of mind, however. That was due to my friend [BA],[9] who one evening came up the stairs into the hallway we shared in Fleming talking

---

[7] Chris I barely knew, but Mitch and I were, considering the disparity in our social positions, fairly tight. I must have made a favorable impression upon him, because I was told he said about me "That guy is the weirdest kid at Caltech," which I still figure to have been the finest compliment I ever received. — Of course it required no effort to live up to his expectations.

[8] Dana Scott, "Continuous Lattices." Oxford University Computing Laboratory Technical Monograph PRG-7, August, 1971. — Quoting from the abstract: "The main result of the paper is a proof that every topological space can be embedded in a continuous lattice which is homeomorphic (and isomorphic) to its own function space. The function algebra of such spaces provides mathematical models for the Church-Curry lambda-calculus."

[9] Now deceased, alas, and never shy about crediting his drug abuse for making him a computer millionaire. But let us not name names.

animatedly to another party, both of them stoned out of their gourds on acid and apparently arguing about the Nature of Television; as he passed my doorway he posed the question "Does the picture make the sound, or the sound make the picture?" — Of course this solved the mind-body problem at a single stroke.

{...}

And this, I suppose, has been a Flashback.

Stoned at King Soopers (1967)

An epiphany: entering the grocery, the door opening automatically for me — remembering Emerson, "All the thoughts of a turtle are turtle"; realizing that I had seen into the mental state of the photocell: pure anticipation.

The old principle that automata must have souls. — Really, this was there already in Leibniz.

Ghost in the shell (1971)[10]

There is supposed to be an argument from Gödel's theorem to show that the mind can't be a machine, but I've never understood it. Of course I have never thought that was my fault.[11]

Penrose, for one, made a book out of it; and though I didn't believe him either it was amusing that whole issues of the journals[12] were repurposed to try to refute him.

At any rate both sides of the argument are bullshit. It doesn't matter whether minds are machines or not. Even machines aren't machines. This can be seen in at least two ways:

---

[10] The basic argument (here reconstructed) has not changed a great deal since it first occurred to me, though it has obviously been revised and amended to reflect the march of mathematical progress.

[11] There are (at least) a couple of good reasons to be skeptical. — First, Gödel himself always thought this was a consequence of his incompleteness theorem, and was said to have been working on a formal proof of the proposition; which, however, he never finished, and didn't publish. That seems suspicious. — Second, the idea that a human agent could find itself trapped in repetitive cycles of mechanical behavior is supposed to be prima facie absurd; nonetheless it is the fundamental thesis of psychoanalysis; and indeed the Freudian method looks a lot like teaching a Turing machine its Gödel sentence. (Compare Thomas Mann: "No man remains what he was once he has recognized himself.")

[12] The book [published in 1989] was *The Emperor's New Mind.* For expressions of outrage cf., e.g., *Behavioral and Brain Sciences* Vol. 13 #4 (1990), pp. 643-705, and Vol. 16 #3 (1993), pp. 611-622.

— First, "machine" in the sense of artificial intelligence never really means "Turing machine" anyway; rather one augmented by a (true)[13] random number generator — a source of randomness for nondeterministic algorithms; neural networks, for instance, fall under this description, as do Metropolis and genetic algorithms.

— Second, even deterministic machines aren't deterministic. — That is, though you have the picture (sharpened by formal models) of a system with a delimited[14] set of states, whose behavior is determined by a function which computes the next state (in a discrete series) from the current state, and assume that knowledge of the next-step function entails knowledge of all its iterates — that "deterministic" means "completely predictable", in other words — this is a kind of optical illusion. It doesn't really work that way.

{...}

It's more amusing to explain this anecdotally.

I saw Skinner lecture once, in Berkeley in the early Seventies. This was shortly after the publication of one of his numerous paeans to

---

[13] I.e., one referring to what theory terms "an Oracle", some external source of input like a Geiger counter recording radioactive decays. — Purely computational (pseudo) random number generators fake it, by producing sequences which are determined by such complicated rules that they "look" random (a literature has been expended trying to define the implicit oxymoron), i.e. take a long time to repeat, but on the other hand can be rapidly computed. It is truly amazing how often the naive employment of these mechanisms leads to mortifying blunders. Nearly every serious programming project I have undertaken has been almost immediately been sidetracked by an attempt to write a better random number generator than the one that has just fucked me in the nose.

[14] This is tricky: machine theory allows not simply for the case of a finite state set, but also for a finite "internal" state set augmented by potentially-infinite auxiliary storage, the tape of a Turing machine or the stack of a PDSA, e.g., which can only be accessed finitely, e.g. one item at a time. — In practice, of course, all machines are really finite, and immense ingenuity is expended to overcome limitations of time and space.

Mind Control:[15] he spoke in a large lecture hall, to a full house packed with an extremely hostile crowd, and though he couldn't win them over, he did at least earn their respect. — There is a certain naive pigheaded charm some nerds possess, and he had it in great measure. If nothing else, I admired his balls.

We were all jammed in like sardines, and I was sitting in the aisle a few feet downhill from my girlfriend, so as it turned out I couldn't talk to her until afterward and it wasn't obvious we were attached. Instead I found myself embedded among a covey of attractive female undergraduates. One of them was lecturing her friends on the nature and context of the debate we were participating in, and every time she hesitated because she didn't quite know how to continue, I finished her sentence for her. — This provided me with the standard anecdote I used in later years to describe what Berkeley was like, in the Golden Age: this was the first, last, and only time a girl wanted to go home with me because I knew Beckett wrote *Endgame*.

At any rate I was fascinated by Skinner's insistence on the predictability of human behavior; there was an echo of that Freudian certitude that had always seemed so maddening, but his explanatory apparatus was cleaner, much more austere. So what was wrong with it?

Part of it, obviously, when I read over theoretical behaviorism later[16] to find the basis for his claims, was that the most consistent version of his approach made it a point of dogma not simply that one *should* not but that one *could* not assign internal states to the organism; since simple thought experiments showed that removing the brain from the skull would produce a noticeable difference in behavior, at least

---

[15] Probably *Beyond Freedom and Dignity* [1971].

[16] Not that I wasn't familiar with it already from, e.g., Russell's synopses in *The Analysis of Mind* [1921], but it was instructive to read the modern literature and observe how little theory had progressed since Watson and Pavlov.

among people who hadn't voted for Nixon, that was obviously wrong. — Part of it was that the kinds of laboratory experiments to which behaviorists confined themselves made essentially meaningless measurements of a kind which could not, for instance, tell you anything about the functioning of even the simplest digital computer.[17]

But the main thing — what was instantly suspect — was his claim that behaviorist methods would suffice to explain even the "behavior" of mathematicians. For this seemed, after all, to be a bizarre assertion: were we seriously to think that from considerations of elementary physics — presumably by solving some system of differential equations — not that Skinner ever wrote any down, of course, but an explanatory framework based on the measurement of quantities expressed in real numbers — i.e. founded on physics envy — would inevitably (as any real physicist could instantly see) lead to such a theory — that we could tell whether a mathematician was going to be able to prove a theorem? How was one *complicated* mathematical problem — indeed all of them at once — supposed to reduce to another which seemed so much simpler?[18] — And *why* it seemed

---

[17] A technical refinement of the point, which Chomsky used to great polemical effect, was that though for the simplest class of finite-state automata internal states can in principle be defined away as equivalence classes of mappings from inputs to outputs, this [a] relies on the examination of infinite sets, and [b] the conditioned-reflex prescription applied to a finite training set of stimuli and responses only works for this simplest class, and cannot determine the behavior of machines that recognize more complex grammars. Since such machines already existed and even then were generating our utility bills, this seemed a fairly crushing objection.

[18] Actually it isn't impossible that a relatively simple differential equation, or system of them, could be universal in the sense of Turing; the solution of Hilbert's tenth problem showed something analogous for Diophantine problems, i.e. that there is an equation of the fourth degree in 14 variables that is universal: see Martin Davis, "Hilbert's Tenth Problem Is Unsolvable," *American Mathematical Monthly*, March 1973, 233-269. — One could conjecture, in other words, the existence of a universal *analog* computer. — But a simulation that modeled a universal Turing machine with a differential equation wouldn't be any *simpler*. The inherent difficulty of the problem is irreducible. So the picture you have of having found a solution is a kind of optical illusion. — "All I have to do is solve this equation, and..." — but how? In practice you have only replaced one intractable computation by another of equivalent difficulty.

bizarre wasn't difficult to figure out. For though if we asked the mathematician to prove, say, some statement in the predicate calculus it might seem unlikely on intuitive/romantic grounds that we'd be able to describe the necessary "creative leap", really it isn't necessary to appeal to this at all: one could simply ask the mathematician to attempt mechanically to construct a proof using some method like semantic tableaux; and then observe that whether this procedure terminates on arbitrary input is, in general, undecidable. — I.e. you needn't appeal to a magical black-box mechanism at all; even if you know the mechanism, even if the box is transparent, it makes no difference. — So the grand reductive gesture of pretending the box *has* no internal degrees of freedom is doubly pointless.

{...}

Put another way, one need not challenge Skinner with the problem of predicting whether, say, Gauss sitting at his worktable will be able to come up with a proof of, say, Goldbach's conjecture;[19] one can simply ask Skinner to tell us whether Gauss in performing the arithmetical check will find a counterexample to Goldbach's conjecture in finite time; and if so, *when*. Because this means that the behaviorist must then in effect be able to tell us in advance whether Goldbach's conjecture is true. (And decide this by solving some magic differential equation, or system of them.)[20] — True, we can, if we are faster, stay ahead of Gauss in the computation. But this is not an *effective procedure*;

---

[19] Communicated in a letter to Euler in the 18th century, the statement (based then on very flimsy empirical evidence, based now on dismayingly extensive tests) that every even number greater than 2 is the sum of two primes. A proof now does appear to be closer, but the feeling has generally been that if there really are "natural" elementary statements about the integers that are true but not provable, they would look like this. (Gödel himself referred to this possibility explicitly; see the notes to *Gödel 1972a* in his *Collected Works*, Volume 2.)

[20] Given the hypothetical universal system one might solve the equations "by computer", i.e. numerically, but then we simply have one machine emulating another of equivalent complexity; nothing is *reduced*, in other words.

we can't guarantee an answer to the question exists in advance.[21] — We can't say how the computation will come out. — And therefore, in the most significant sense, we cannot *predict* what Gauss is going to do, *even if he is emulating a machine*.[22]

$$\{...\}$$

There are various equivalents[23] that illustrate the case equally well, but the canonical question is the halting problem for Turing machines: suppose we give Gauss the description of a Turing machine, and an input tape — all this is finite — and then ask Skinner to tell us his prescription for deciding, in the general case, when/whether Gauss will finish computing the answer, and what the result will be. — To explain his behavior, i.e. — But he can't, because this is known to be impossible. — Conceivably Skinner might object that the proof of unsolvability assumes the validity of Church's thesis, an essentially metaphysical hypothesis[24] which he rejects — another myth which will dissolve in the acid bath of his scientific rationality; but then he's saying that he has some method of computation (an oracle, e.g.) that is

---

[21] I.e. though I may not know before I perform the computation that 16117667 times 16283543 is 262452723654181, I do know that there is an answer, and if I follow the rules for multiplication I will find it within a certain number of steps which can be bounded in advance. Not all computations come with such guarantees.

[22] I take it for granted that a human (like Gauss) can emulate any Turing machine; since after all the idea of the Turing machine is that it formalizes the abilities of a human calculator. — It is assumed, in other words, that the objection that Gauss might not have enough time or scratch paper is frivolous and irrelevant to the principle at issue. (This has nothing to do with his *behavior*.)

[23] The word problem for semigroups, e.g., which asks whether there's a general method for deciding whether two strings of symbols are equivalent under a given finite set of equational transformations, or the general Diophantine problem (Hilbert's Tenth), whether an mechanical procedure exists to determine whether a polynomial equation in a finite number of variables with integer coefficients has integer solutions.

[24] Fred Thompson was the first guy I heard call it that. He was certainly right.

more powerful than a Turing machine. — At which point we tell him to put up or shut up. And the rest is silence.

{...}

You can summarize the lesson of this gedankenexperiment as follows: since prediction is simply computation,[25] machines in general are not predictable; since people can emulate arbitrary machines,[26] the behavior of people is not predictable.

So behaviorism isn't completely useless; its refutation teaches us something valuable.

{...}

This doesn't explain why a mob of hippies showed up to howl for Skinner's head, of course. That had to do with the supposed conflict between the freedom of the will and determinism. But I think the real issue there is related, essentially psychological, the anxiety that you feel about the possibility not that your actions are "determined" in some complex and unknowable fashion, but that they can be *predicted*.

We all remember Dostoevsky's lengthy rant[27] in *Notes from the Underground*, the famous Crystal Palace passage about the conflict between the freedom of the will and mathematical certainty:

> ... then, you say, science itself will teach man ... that he never has really had any caprice or will of his own, and that he himself is

---

[25] This seems self-evident, but in the same way all propositions do that insinuate metaphysical hypotheses. (Here again Church's thesis.)

[26] By definition: when Turing refers to a "computer" in his original paper, he means a human following rules with pencil and paper; electronic computers did not yet exist.

[27] It would be anachronism to call it that, but this is a classic example of what is now called a flame.

something of the nature of a piano-key or the stop of an organ, and that there are, besides, things called the laws of nature; so that everything he does is not done by his willing it, but is done of itself, by the laws of nature. Consequently we have only to discover these laws of nature, and man will no longer have to answer for his actions and life will become exceedingly easy for him. All human actions will then, of course, be tabulated according to these laws, mathematically, like tables of logarithms up to 108,000, and entered in an index; or, better still, there would be published certain edifying works of the nature of encyclopaedic lexicons, in which everything will be so clearly calculated and explained that there will be no more incidents or adventures in the world.[28]

Or more succinctly:

Good heavens, gentlemen, what sort of free will is left when we come to tabulation and arithmetic, when it will all be a case of twice two make four? Twice two makes four without my will. As if free will meant that!

But though the existentialist antihero of the *Notes* thus insists perversely on behaving irrationally to express his defiance of soulless rationalism, he needn't have bothered. Arithmetic itself is perverse enough.

That is, though it is already difficult enough to understand the traditional problem — your *will* is still free even if what you *want* is determined, [29]and so what — the point is really that determinism appears to entail predictability, and prediction allows control: if people

---

[28] This is the Constance Garnett translation.

[29] I cheerfully admit that emotional responses are often predictable, at least for most people much of the time; else they would be more difficult to manipulate. But here again the claims of psychology are exaggerated.

are machines, then seemingly they can be *used* as machines; *that* is the terror of mechanism.

You have the oppressive sense that some puppet master like Skinner can look over your shoulder (with his "table of logarithms") and nod smugly at everything you do, because he has foreseen it all in advance; and since he knows what you will do when he pushes your buttons, *he can make you do whatever he likes*. — And Skinner of course endorses this interpretation at every turn, this is the plan for his Utopia. — That it might be determined[30] in advance but not known or even knowable — well, there is something that never occurred to the determinists; omniscient though they were supposed to be. In fact it doesn't seem to have occurred to anybody.

{...}

The anxiety is not unknown among physicists. There is a strong resemblance, e.g., between Eddington's argument (made nearly as soon as the uncertainty principle was invented)[31] that the indeterminacy of quantum mechanics permitted the freedom of the will, and Penrose's rather weird assertion (1989) that "microtubules" within the cell could turn the brain into some kind of quantum computer beyond the reach of Turing.[32] In both cases it is clearly the predictability of the mechanical that disturbs them. — Your will cannot be free if someone can know what you will do. — More than that, an artist or a musician or a mathematician cannot be truly creative, since whatever they produce is simply the result of a

---

[30] To return to the model of the system of differential equations, there are in general existence theorems that tell you they *have* solutions which are determined uniquely by their initial conditions. This doesn't mean you can say what the solutions are. (Or — the butterfly effect — that they are stable under infinitesimal perturbations, which is a necessary condition for computer simulation.)

[31] Cf. *The Nature of the Physical World* [1927].

[32] Pure science fiction, so far as anyone can tell.

mechanical process. One could simply write a Shakespeare emulation program and output *Hamlet*, without the intervention of the fifty million monkeys with typewriters. — This is a slightly more interesting problem, but the predictability issue is again key: one might in principle be able to program a (pseudo)machine to write something like *Hamlet*, but it would never turn out the same way twice, and given time and sufficiently many rewrites would turn into something else entirely. — Whether that would satisfy Penrose I don't know. But my credentials as an unreconstructed Romantic are unquestioned, and it satisfies me.

{...}

There is also an amusing functional equivalence between Skinner's implicit[33] assertion that he could predict the answer to any mathematical question from the laws governing the organism (the differential equations, or whatever) and Plato's insistence that all mathematical knowledge is something the soul obtained in a previous life/is engraved upon the Forms; it is accordingly suggestive that they envisioned similar Utopias. (And that they bore a suspicious resemblance to the Crystal Palace.) — Who were our behaviorist overlords going to be, but the new Guardians? — Moreover there are parallels with the apparent aims of the classical school of artificial intelligence, as exemplified by Minsky: if the brain was just a machine running a determined program, then those select few who could read the source code could make mere humans (aka "the lusers")[34] do whatever they wanted; traditional hacker culture was also based on fantasies of control, the domination of the programmers over the programmed.

---

[33] You have to say "implicit" because it is obvious he did not understood what was coming out of his mouth. Certainly he never understood Chomsky's critique.

[34] Traditionally the MIT school divided people who interacted with computers into two classes, programmers and users; the former were the master race, the latter, serfs and peons. — It is not an accident that, as Big Tech continues to conquer the world, more and more of it reverts to feudalism.

{...}

Another fantasy of determinism, indulged by the imaginative, is that one ought to be able to predict the course of history in advance. — This is not, precisely, the usual motivation of the self-styled grand theoreticians of history, who seem not to have advanced beyond Linnaean notions of classification — Spengler, e.g., goes on at great length in his philosophical preamble about Goethe, morphology, the incapacity of trivial concepts of causality to grasp the architecture of Destiny, etc.[35] — the game of hypothesis and prediction never caught on among the German idealists, obviously — but it is a fairly common speculation in science fiction. Isaac Asimov's *Foundation* novels are probably the most famous examples, and have been quite influential[36] in that respect: he imagines the decline and fall of a galactic empire on the pattern of Gibbon's Rome, and a dedicated cabal of monks, privy to detailed advance knowledge of the pattern history must follow, working to preserve civilization through the ensuing Dark Age, whose duration they will thus be able to minimize.[37]

---

[35] In the translation of Charles Francis Atkinson: "The means whereby to identify dead forms is Mathematical Law. The means whereby to understand living forms is Analogy." — "... there can be no question of taking spiritual-political events ... at their face value, and arranging them on a scheme of 'causes' or 'effects' ... ." — "That there is, besides a necessity of cause and effect — which I may call the logic of space — another necessity, an organic necessity in life, that of Destiny — the logic of time — is a fact of the deepest inward certainty... ." — "Mathematics and the principle of Causality lead to a naturalistic Chronology and the idea of Destiny to a historical ordering of the phenomenal world." — And so on. Of course all this is nonsense.

[36] Sometimes in unobvious ways: the Nobel laureate Paul Krugman, for example, is a science fiction fan, and has often remarked that Asimov's vision of a social science that could make rigorous predictions inspired him to study economics.

[37] This idea of a monastic order preserving knowledge through a Dark Age is another favorite theme of science fiction; see for instance Walter Miller's *A Canticle for Leibowitz*.

The superficially convincing argument for the possibility of such prescience is the analogy with statistical mechanics: you don't need to know how each individual gas molecule is moving to calculate the pressure on a cylinder. — The argument probably fails on appeal to the butterfly effect, since there are many examples e.g. of critical battles won or lost by accidents of timing, and (pace Tolstoy) great men (and women) do seem to appear fortuitously and decisively alter the course of events — this is a more complex dynamical problem than that posed by a gas, after all — still, though one can't predict the weather exactly, one *can* predict climate change; so one might guess that on a longer time scale the rolls of the human dice may even out.

Nonetheless something similar to Skinner/Gauss does apply: the future of industrial civilization as we have it right now, for example, depends at bottom on facts of physics and astronomy as yet unknown — whether room temperature superconductors exist, whether fusion reactors can ever be practical, what results may come from mining the asteroids, whether irreversible ecological collapse is really at hand — whether an undetected asteroid is going to run into the Earth and reprise the extinction of the dinosaurs — and you can't tell how human history will turn out without knowing the answers to these external questions. — As was the past so determined: the history of the modern world follows in large part from the contingent fact that when Columbus sailed west, there was an extra continent to bump into. — So the one kind of omniscience presupposes the other. Even economics, which involves measurable quantities and superficially seems more easily predictable, depends at bottom on the ways that we can extract free energy from our environment, and thus on unpredictable boundary conditions and undiscovered facts of mathematics and physics (and chemistry and biology and geography and ...) which cannot be known without — well, without being known.[38] How could an economist in 1950 have predicted that nuclear

_____

[38] Here I'm sure Heidegger would insert some rhapsody on the knowable knowingness of being-known, but — thank the gods who have not yet fled — I lack his gift for tautological obfuscation.

power based on fission reactors would turn out to be more trouble than it was worth, or foreseen the laser, the transistor, the photovoltaic cell, the microchip, or Moore's Law? — Von Neumann saw none of that coming, and he was as omniscient as anyone could have been at that time — for instance, he famously stated that four computers like his[39] primitive MANIAC[40] would suffice for all the computational needs of the world.[41]

{...}

A slightly weaker statement, whose relationship to undecidability is still not completely understood, is that a computation may not be impossible but nonetheless may be prohibitively difficult. This might seem like a frivolous objection were it not the case that relatively simple problems can be shown to be unsolvable within existing space and time.[42]

---

[39] Actually constructed by Nicholas Metropolis at Los Alamos following Von Neumann's IAS design, but why quibble. — Authorship of the acronym, which was meant to stamp out this reprehensible practice in its infancy and failed miserably, has been ascribed to both.

[40] Less powerful than a pocket calculator of the Seventies, and many orders of magnitude less powerful than the contemporary iPhone; which exceeds in computational power the fastest supercomputers of even the Eighties.

[41] As a final note, Lockheed Martin is supposed to be pitching a tool called the World-Wide Integrated Crisis Early Warning System (Google at your own risk), originally a project funded by DARPA, which is supposed to have had some success anticipating national and international crises. Apparently among other things it predicts the imminent collapse of the Russian government, surely a consummation devoutly to be wished. — The historian Peter Turchin, on the other hand, on the basis of mathematical analysis of a large data set measuring a variety of historical trends, finds parallels between previous periods of crisis and the current situation of the United States, and predicts the disintegration of civil society within the decade. — And he does indeed begin *War and Peace and War* [2006] by invoking the example of Asimov's hero Hari Seldon.

[42] David Ruelle ("Is Our Mathematics Natural?" *Bulletin of the AMS*, Vol. 19, Number 1, July 1988) mentions a suggestion of Pierre Cartier that the axioms of set theory might be inconsistent but a proof of this would be so long that it couldn't be performed in the physical universe.

One class of examples would include the travelling salesman problem, which scales exponentially in the number of cities;[43] a greater degree of difficulty may be found in problems like computing Ramsey numbers, or evaluating the Ackermann function, which is defined as follows:

$$A(x, 0) = 0$$
$$A(0,y) = 2y$$
$$A(x, 1) = 2$$

else

$$A(x, y) = A(x - 1, A(x, y - 1))$$

Then

$$A(1,n) = 2^n$$
$$A(n, 1) = 2$$
$$A(n, 2) = 4$$
$$A(2,3) = 16$$
$$A(2,4) = 65536$$

in general

$$A(2,n) = A(1,(A(2,n - 1)) = 2^{A(2,n-1)}$$
$$A(3,1) = 2$$
$$A(3,2) = 4$$
$$A(3,3) = 65536$$

and

$$A(3,4) = \ldots$$

---

[43] Given a planar map and the positions of n cities upon it, to construct a route of minimum length that visits each city exactly once; for n around 120 the number of possibilities that must be examined exceeds the number of cells of dimension the Planck length in the visible universe.

i.e., this is a recursion that will not terminate before the stars go out, and the answer couldn't be written down[44] if you used all the volumes in Borges' Library of Babel.

{...}

Regarding the Goldbach conjecture, subsequent developments have only confirmed Gödel's intuition. Certainly it is possible that there is some simple and elegant proof of this proposition, but it seems more likely there is not; and then there are curious questions about how complicated a proof, even if one does exist, might have to be. The proof of the celebrated four-color theorem,[45] for example, another result with an extremely simple statement[46] which defied demonstration for several generations, turned out not to involve (at least has not thus far) the elegant manipulation of powerful abstractions developed from mathematical theories of great scope and formal beauty — as did, for instance, the proof of the famous Weil conjectures (Deligne 1973), the proof of Mordell's conjecture (Faltings 1983), and the celebrated proof of the Taniyama conjecture (Wiles 1993/4), which entailed the last theorem of Fermat, for three centuries the most famous unsolved problem in the subject — but

---

[44] In decimal notation, at least. Of course in effect we have already specified the number with a small finite number of symbols.

[45] Appel and Haken (K. Appel and W. Haken, "Every planar map is four colorable, Part I: discharging," *Illinois Journal of Mathematics*, **21** (1977) 429-490; K. Appel, W. Haken, and J. Koch, "Every planar map is four colorable, Part II: reducibility," *Illinois Journal of Mathematics*, **21** (1977) 491-567) considered more than 1900 configurations and more than 300 so-called discharging rules; the proof was so complicated that no one could simplify or even check it for twenty years. Finally Robertson, Sanders, Seymour, and Thomas (N. Robertson, D. Sanders, P.D. Seymour, and R. Thomas, "The four-colour theorem," *Journal of Combinatorial Theory*, Series B **70** (1997), 2-44) reduced its complexity to 633 configurations and 32 discharging rules — a simplification which allowed a complete proof to be written out and verified by computer. — An executive summary is provided by B. Bollobás, *Modern Graph Theory*. Berlin: Springer-Verlag, 1998; pp. 159-161.

[46] Specifically: that any map in the plane can be colored with no more than four colors in such a way that no two contiguous regions have the same color.

rather the enumeration and systematic elimination of over a thousand separate cases, handled mechanically by a computer program and not, at least not immediately, understood directly by any human mathematician. This engendered a rather painful debate, and raised ugly questions: is there any guarantee a cleaner proof exists? are many unsolved propositions with simple statements destined to have similar resolutions? and so on. — Mathematics is supposed to be an elegant duel with light-sabers, not some kind of rude barbarian combat in which the victor clubs his opponent to death.
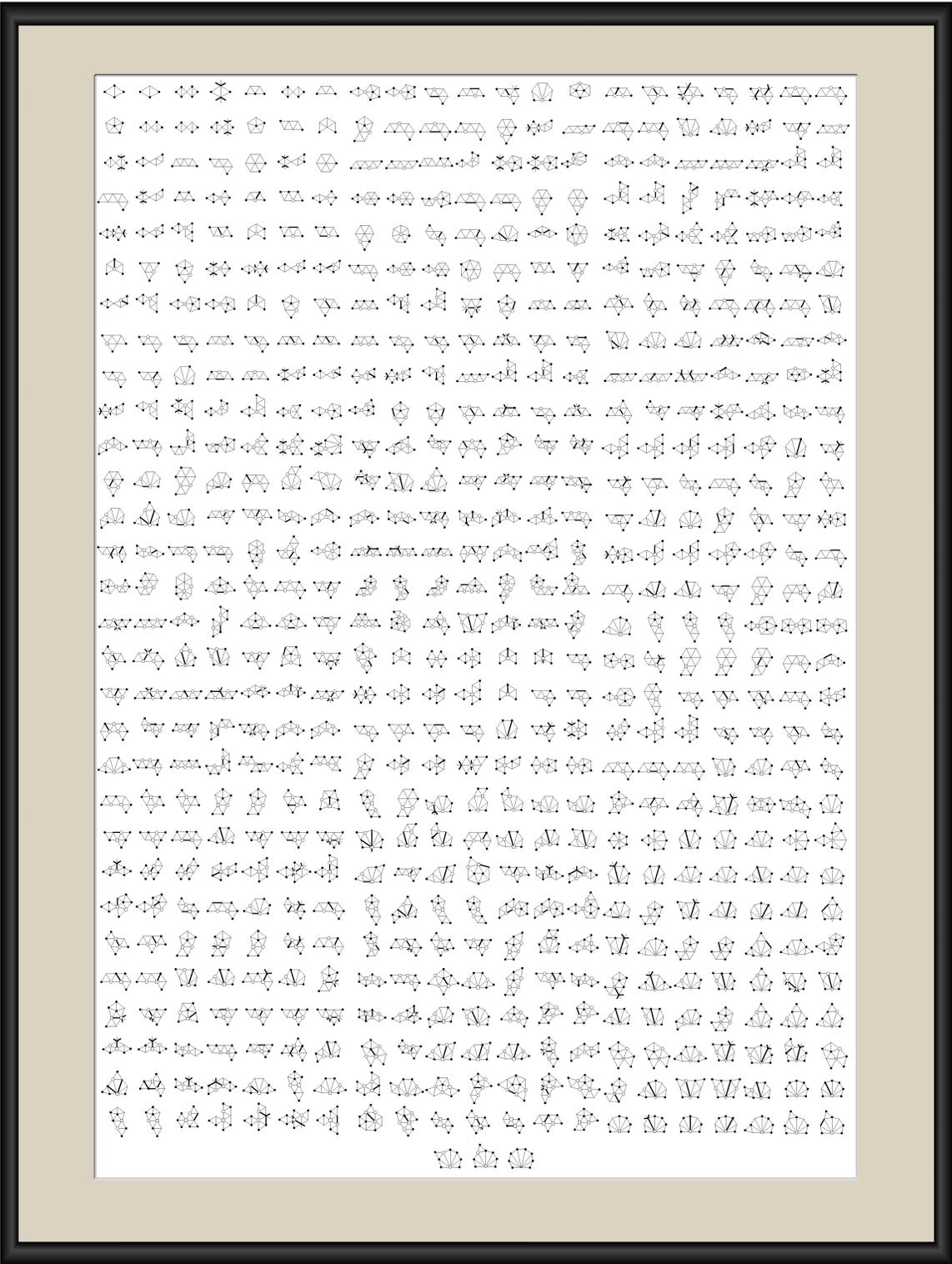
The 633 configurations of Robertson, Sanders, Seymour, and Thomas.

<center>{…}</center>

One might contrast the solution of the game of checkers, obtained by researchers at the University of Alberta; they examined 500,000,000,000,000,000,000,000 different configurations to show that there is a strategy for the game that does no worse than draw.[47] But there is nothing particularly shocking about this, because the rules of checkers are the product of a kind of caprice, and games of strategy in general have unbounded logical complexity;[48] in fact it's almost surprising it was this easy. — One would expect chess and Go to be solvable in similar fashion, though it is difficult to imagine that a computer could finish enumerating the cases before the heat death of the universe.

(It is instructive, incidentally, to consider the case of a human playing a machine at chess; the moves of the latter are completely determined by a set of algorithms; the moves of the former are not, and it is obvious no behaviorist ever considered the question of how they could be reduced to a finite set of conditioned reflexes[49] — this despite the fact that programming a computer to play chess was one of the first problems that occurred to the pioneers of artificial intelligence.)[50]

---

[47] Jonathan Schaeffer, Neil Burch, Yngvi Björnsson, Akihiro Kishimoto, Martin Müller, Robert Lake, Paul Lu, Steve Sutphen. "Checkers is solved." *Science* 14 September 2007: Vol. 317, Issue 5844, pp. 1518-1522. The program (Chinook) can be played online.

[48] As Ulam was fond of pointing out, questions about games of strategy nest quantifiers to arbitrary depth — the problem of chess, e.g., can be stated as whether for all opening moves by white there exists a move by black such that for all moves by white there exists a move by black such that, etc. — whereas in normal mathematics few definitions (Ulam's pet example was that of an almost periodic function) nest them more than four or five deep.

[49] Could operant conditioning be employed to teach a rat to play chess? — No? (Why not?) — What about tic-tac-toe? — If one rat can't be conditioned to play chess, can a roomful of them? Enquiring minds want to know.

[50] Turing himself wrote one of the first such programs; apparently to revenge himself upon his colleagues at Bletchley Park, who pissed him off by beating him so consistently.

But mathematics is supposed to be *necessary* truth. When a simple question has an enormously complicated answer[51] it *looks* like truth by accident.

In the case of Goldbach's conjecture results have been obtained which show that a related proposition holds for all numbers greater than an enormous lower bound; though thus far this lies far beyond the range of possible computation, it is conceivable that some combination of faster computers and improved lower bounds could make it possible to construct a complete proof by pasting together an analytic result (true for even numbers greater than some enormous N) and brute force enumeration of the rest of the cases (verified by explicit computation for even numbers less than or equal to N). If this were the case, it would present us with an example of a number-theoretic theorem about the integers, what we would like to think of as quintessential necessary truth, which would nonetheless have the appearance of being true only by accident. —Wittgenstein would have loved this, but no one else.[52]

(Obviously it is also disturbing that a proof based on a computer program depends on a proof that the program is correct; these in practice are practically impossible to provide, and, handwaving arguments about the probability of error being vanishingly small

---

[51] The usual situation goes the other way around — a complicated problem has a simple solution: the problem of the thirteen pennies, for example. [Not sure I can explain that without a diagram, and how are diagrams included in footnotes? Hmmmm...........]

[52] *Remarks on the Foundations of Mathematics*, III.42: "It might perhaps be said that the synthetic character of the propositions of mathematics appears most obviously in the unpredictable occurrence of the prime numbers. ... The distribution of primes would be an ideal example of what could be called synthetic a priori, for one can say that it is at any rate not discoverable by an analysis of the concept of a prime number." (Translated by G.E.M. Anscombe. Cambridge, M.I.T. Press, 1967.) — This sounds surprisingly Kantian, but there is something intuitively correct about it.

notwithstanding, it isn't immediately obvious that we haven't been presented with an infinite regress.)

$$\{...\}$$

Appended note:

The march of mathematical progress has now brought this scenario to fruition: the weak Goldbach conjecture, which states that that every odd number greater than 5 is the sum of three odd primes, had been proven by the refinement of analytical techniques due to Hardy, Littlewood, and Vinogradov, among others, to be true for all numbers greater than a bound C; a series of attempts to lower C had [2002] reduced it to about 10^1350, still far beyond the reach of computer verification. Recently, however, Helfgott[53] has lowered C to 10^27 and since computational efforts[54] have extended numerical verification nearly to 10^31, the proof-theoretic chimera has now been stitched together. — The question remains whether further refinements of these techniques can gradually reduce C to some value more satisfying to intuition: 10 certainly would work, but 100? 1000? 1000000? — Where to draw the line? — In the meantime, though the weak Goldbach conjecture is now known to be true, it falls into a kind of uncanny valley[55] between the analytic and the synthetic.

$$\{...\}$$

---

[53] H.A. Helfgott, "The ternary Goldbach conjecture is true"; arXiv:1312.7748v2, 17 January. 2014.

[54] These too rely on (partial) empirical verification of another open question, the Riemann hypothesis, for which enough zeroes have been computed to bound the gap between successive primes sufficiently well up to 10^27 that an odd prime can be subtracted from the triple to yield an even number less than the limit to which the even Goldbach conjecture has been verified, of the order of 10^18. Not to take anything away from Helfgott's remarkable achievement, this argument is a ridiculous kludge.

[55] A term used in computer graphics to designate the disturbing gap between the obviously phony and the photorealistic. Thus synthesized faces possess an unsettling quality.

A similar simple proposition about the primes no one has any idea how to prove is the twin prime conjecture: that there are an infinite number of pairs (p, p+2) which are both primes.[56] — About this Cohen after expressing skepticism regarding the ability of axiomatic frameworks to capture the properties of the mathematical objects they describe asks "Is it not very likely that, simply as a random set of numbers, the primes do satisfy the hypothesis, but there is no logical law that implies this?"[57]

In fact the natural metamathematical conjecture is that almost all[58] conjectures that appear to be true on probabilistic grounds are true but unprovable; i.e. that these two senses of "true, probably" and "probably true" are equivalent.[59]

The Goldbach example suggests that there may be many conjectures with relatively simple statements whose probability of truth is unity (since they can be shown to be true on the complement of a finite set) but which then are true or false globally by a kind of contingency, in that the proof can only be filled in by case by case enumeration. Indeed this situation may be typical.

{...}

To state one moral, then: philosophical intuition is not completely worthless, but like any other kind of intuition it is based on a kind of

---

[56] See (xix).2003.7.8, "Minor triumphs".

[57] Paul J. Cohen, "Skolem and pessimism about proof in mathematics", *Phil. Trans. R. Soc. A* (2005) **363**, 2407-2418. (12 September 2005.)

[58] "Almost all" has the technical definition "except on a set of measure zero" and doesn't really mean anything unless such a measure can be defined. Here it can.

[59] For reasons that may be obvious this occurred to me while meditating gloomily on a lecture about the abc conjecture.

experience; and it should not, therefore, be a surprise that its conclusions evolve when that experience broadens.

The analytic/synthetic distinction was introduced by Kant; was almost immediately questioned by Gauss, who had already understood the possibility of a non-Euclidean geometry; and then revised after radical extensions of the idea of entailment to include inferences like "7 + 5 = 12" and "a straight line is the shortest distance between two points", even though (as Kant pointed out) neither falls under the traditional definition of a conclusion being included in the premises.

Now, it becomes clear, it may be less a black and white distinction than a grayscale continuum, resolving under closer examination into an arbitrarily ramified hierarchies of the kind with which we have lately become familiar in complexity theory. — The more extensive our experience of what constitutes proof, the more baroque may our intuition of necessity become.

{...}

Fundamental misconceptions about mathematics and the nature of prediction notwithstanding, there was a larger fallacy involved in behaviorism: it was based upon an artificially limited, indeed an essentially invalid idea of what constituted science.

You could see it in the polemics Skinner's partisans wrote against Chomsky — here was a real theory of language at last, or at least a piece of one, and it was attacked as unscientific because it was (in Eddington's phrase) physics and not stamp collecting; because they not only did not recognize theory when they saw it, they did not understand its necessity — because they had trapped themselves in the most limited possible conception of empiricism, almost a throwback to Bacon, one in which scientific endeavor consisted entirely in the blind accumulation of disconnected "facts"; the reduction of the philosophy of nature to making statements in an

observation language — which, of course, they didn't even see was ill-defined.

I suppose this was natural. Psychology had spun its wheels from Hume to William James trying to found itself in introspection. A radical break seemed to be called for, what more comprehensive revolt against subjectivism than to deny the existence of the subjective entirely, and in so doing why not banish all "metaphysical" statements altogether? this was the spirit of the age, after all.

{...}

There was, in other words, a desperate anxiety among psychologists that what they were doing was not "science". And quite understandably they sought to make what they were doing "scientific" by imitating what they saw their intellectual elders doing: performing experiments in laboratories and making measurements that produced copious amounts of numerical "data" — publishing "results" in "papers" in "journals", filling them with graphs, tables, charts, and statistical analyses — *going through the motions* — hoping that, by performing the same ritual abasements as (real) biologists, chemists, and experimental physicists, psychologists could acquire their mojo. — There is a name for this, and it is not "scientific thinking".

{...}

Freud gives as examples of what Frazer called imitative or homeopathic magic the following:

> Rain is produced magically by imitating it or the clouds and storms which give rise to it, by 'playing at rain', one might almost say. In Japan, for instance, 'a party of Ainos will scatter water by means of sieves, while others will take a porringer, fit it up with sails and oars as if it were a boat, and then push or draw it about

the village and gardens'. In the same way, the fertility of the earth is magically promoted by a dramatic representation of human intercourse..." and summarizes the principle as follows: "If I wish it to rain, I have only to do something that looks like rain or is reminiscent of rain.[60]

Later Feynman described the practice as follows:

> In the South Seas there is a cargo cult of people. During the war they saw airplanes land with lots of good materials, and they want the same thing to happen now. So they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas — he's the controller — and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn't work. No airplanes land. So I call these things cargo cult science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land.[61]

He did not, however, recognize that the cargo-cult phenomenon extends beyond pseudoscience into what is supposed to be "science" itself. — Behaviorism had a theatrical run of a couple of generations. But the planes never landed.

---

[60] Sigmund Freud, *Totem and Taboo*, transl. James Strachey, London: Routledge Classics, 2001. Chapter 3, "Animism, Magic, and the Omnipotence of Thoughts".

[61] *Surely You're Joking, Mr. Feynman!* New York: W.W. Norton, 1985.

{...}

So that is one way of putting it: the cargo cult imitated the "behavior" of the operators perfectly, but didn't look inside the radios to see what made them work. There must be a moral there.

Another way of putting it is that no one ever said what "behavior" was. There was some vague appeal to observable physical states of the organism, but "observable" and "state" and "organism" weren't defined, and the "state" per se wasn't what was referenced in any case, rather some notion of "action", presumably definable in terms of a (short, finite) temporal sequence of states — though this wasn't defined either, of course. So for all anyone could tell "behavior" might include the response of the subject to cold in the form of goosebumps, or to ultraviolet light in the form of sunburn — note that the response in this case varies dramatically from one subject to another, and that no kind of input-output table relating insolation to degree of burning will say anything about the chemistry of melanin, the real causative factor — or the humidity of expelled breath, or height and weight, or for that matter what the subject said in response to the question "What are you thinking about?" — Behavior could have been *anything*, until it was defined. In fact simply by declaring it to have meant the microstructure of brain activity, to which real scientists more sensibly have turned their attention, the program could now be pronounced a success.

Again: a rigorous definition of "behavior" would entail definitions of "stimulus" and "response", and those in turn would require an enumeration of possible inputs and outputs. — Implicitly, as Chomsky pointed out, the mathematical model behind the smoke and mirrors here is that of a finite-state automaton, which takes as inputs strings of symbols selected from a finite alphabet; each one inducing a state transition for which, in turn, a string of symbols from a finite output

alphabet is produced; this may be pictured, e.g., as a state transition graph with edges labelled by inputs, for instance:[62]
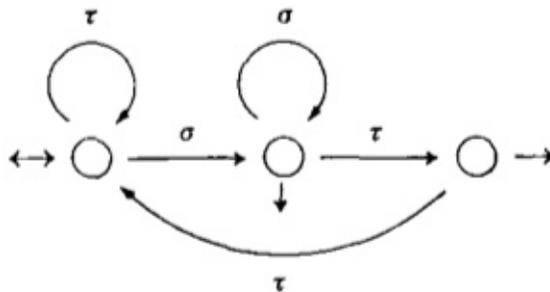
It would then be easy to define away the internal states of the machine as equivalence classes of maps from inputs to outputs, and a kind of behaviorist program can be said to have succeeded.

But where do the input and output alphabets *come from*? Some kind of language is presupposed to specify just what "observable behavior" is, and in practice there is that familiar philosophical bait-and-switch, the appeal to self-evidence, and a host of unexamined assumptions are insinuated by inclusion and omission. And so we have ring bell/ salivate, shout/wince, electric shocks and bits of cheese, and not, say, observations of the form "I rebuked him, and observed that he took offense at the harshness of my manner of expression" — though why not, no one will ever bother to tell you. — "Measure something" — but why measure *this* and not *that*? — And the answer, of course, is that what is and is not relevant has been decided by an implicit appeal to an unstated theory, something beyond the reach of scrutiny. — Elsewhere this is styled "metaphysics".

---

[62] Example 2.6 of Samuel Eilenberg, *Automata, Languages, and Machines*, Volume A. New York: Academic Press, 1974. — I'm cheating here, this is a slightly different type of automaton, but the principle is the same.

Moreover in practice you have only a small subset of the input-output mapping, and have to guess the rest — another version of the problem of induction — and the most natural theoretical device employed to model it is, guess what, an internal state space whose transformations are induced by inputs — in the language of the electrical engineering lab, the wiring of the black box; for a given set of pairs {(input, output)} there will be an ensemble of possible wirings, some kind of maximum-entropy probability distribution imposed upon it, and (ideally) an optimal set of yes/no experiments that will, in the limit, identify the correct internal configuration and thus determine the mapping.

But obviously this is too complicated for a psychologist to appreciate. It may be better to let them keep playing with the knobs on their empty boxes.

{...}

At any rate the simplest objection is still the most powerful: if the brain is a relatively trivial mechanism programmed by the conditioned reflex, then it shouldn't be difficult to reverse-engineer it, and build a model of one.[63] — So: Mechanical Turk; put up or shut up. — Of course this has turned out to be harder than it looked. And though admittedly the training procedure for neural networks bears a family resemblance to the process of conditioning, it is precisely that which to date has rendered it so incredibly inefficient.[64]

{...}

---

[63] Though of course: just because you can build it doesn't mean you can predict what it will do.

[64] Neural networks are trained on sample sets which number in millions, billions, or even trillions. Human infancy does not last a thousand years. Therefore, etc.

By way of general conclusion: though self-reproducing[65] living organisms are composed of cells, biochemical factories which contain a large but finite number[66] of kinds of molecular machines which function according to the laws of physics and chemistry — and one can, in principle, write down equations of motion for the dynamical system this ensemble represents — even in simplified form these would involve millions of variables, there is no sensible way in which one can suppose they could be solved, and in any case they are, strictly speaking, quantum-mechanical in nature and thus indeterministic; not that intrinsic thermal jiggle does not render the classical problem stochastic anyway. In consequence even when some kind of recognizably mechanical procedure is being implemented, in the operation of an enzyme, e.g., or the reproduction of a strand of DNA, nothing ever works the same way twice; not even the fabrication of the machines themselves. — Moreover this is not some kind of regrettable design flaw which would be eliminated in a more perfect world — as designed by Plato/Skinner/Minsky/... — this is precisely what made life possible in the first place. (It is also what renders biological design so robust.)

So in the sense that disturbs us — that mechanism is something which does the same thing the same way every time that it functions — biological machinery is not machinery at all. — Indeed to think that it is, or even that it ought to be, is simply insane. — Life is the product of evolution, and evolution consists precisely in making up rules in order to break them.

Perhaps we should call this the paradox of vitalism, then: that despite being wrong about everything it ends up winning most of the arguments anyway.

---

[65] Viruses are simpler, but must hijack the machinery of cellular organisms to reproduce.

[66] Counting genes, I would guess between ten and a hundred thousand. This is probably low.

So even though there is a philosophical moral to be found here, as usual it looks like a joke.

"The complicated thing about the British scene was that Faithfull's song ['As Tears Go By'] was not written by a coffeehouse folksinger or even by a commercial songwriter counterfeiting the coffeehouse style—it was by Mick Jagger and Keith Richards, and they were simultaneously rockers and Bohemians. When Faithfull eventually hooked up with Jagger in a more intimate way, they went to his room and discussed Roman Polanski, 'disturbed states of mind,' and Dylan, then took a walk in the park, where she quizzed him about King Arthur and the Holy Grail. 'It was quite ludicrous,' she recalls, 'but that's how we were then. You would ask your date, "Do you know Genet? Have you read *À rebours*?" And if he said yes, you fucked him.'" [Elijah Wald, *Dylan Goes Electric*.]

Meno (1972)

In the approximation of classical relativistic theory the creation of an electron pair (electron A, positron B) might be represented by the start of two world lines from the point of creation, 1. The world line of the positron will then continue until it annihilates another electron, C, at a world point 2. Between the times t1 and t2 there are then three world lines, before and after only one, However, the world lines of C, B, and A together form one continuous line albeit the "positron part" B of this continuous line is directed backwards in time. Following the charge rather than the particles corresponds to considering this continuous world line as a whole rather than breaking it into its parts.[67]

Plato in his *Meno* argues that knowledge is reminiscence: Socrates summons a slave boy ignorant of mathematics and starts cross-examining him about a geometrical demonstration; when the kid begins to evince understanding he claims this is evidence that what is

---

[67] R. P. Feynman, "The Theory of Positrons." *Physical Review* Volume 76, Number 6. September 15, 1949, pp. 749-759. — In his Nobel lecture [1965] he cheerfully admits to having stolen this idea from Wheeler.

known, or knowledge of abstractions, at least, is already resident in the soul; which must therefore be immortal.[68]

It seems to me that this argument is more or less correct; though I don't think it says what Plato thought it did. — Certainly it has aged well; what Chomsky said about grammar and innate ideas was not very different.

Presumably Plato thought of this because learning something feels[69] more or less the same as remembering: you don't understand, and then you do. There is an abruptness to it, which he remarks; something that feels like the transition from lost to found.

Of course you wonder then about imagination — is this like remembering something that didn't happen? — but the real curiosity is invention, because this is exactly like remembering something you have yet to learn.

———————————

[68] I think that insofar as Plato's argument works, here as always it actually shows that the soul lies outside of, is independent of time, that it is Being and not Becoming, and that if you pitch this in contemporary language most people would still buy it. Not that this has anything to do with playing harps in the Celestial choir; Plato is as usual prone to what appear to us to be purely verbal confusions like the conflation of "timeless" with "immortal." But I would point out, e.g., that though Windows 3.1 — no, too dreadful, I shouldn't say that, I should say the System 7 Mac OS — may not be running on anything at the moment, this doesn't mean it doesn't *exist*. — Though it hasn't *gone* anywhere either.— Compare also the Pythagorean doctrine of metempsychosis with (software) "installation": they are not precisely isomorphic, and indeed it's strange no one thought of the idea that more than one person may have been who you were in a previous life; many distinct persons have the same physical ancestors, after all. (And do all bacteria have the same soul? really, we can do this all night —)

[69] It is amusing to try to come up with an explanation for this subjective feeling: learning involves an inductive computation which is much longer and more complex than it appears to the conscious mind; many processes go on in parallel to try to piece together a solution to the problem; when the result is presented to consciousness, the fragments of understanding are, by this time, things which are "already known," and the process of retrieving them to explain the whole is isomorphic to remembering; so when the conscious ego sees the answer it is, indeed, something already resident in the soul, albeit in a part not easily accessible to conscious inspection; the effect is functionally not dissimilar to retrieving memories from a past life.

So you could with equal justice say this: if knowledge is reminiscence, then invention is remembering backwards in time.

{...}

Because Plato's argument doesn't have any direction to it. If you believed it, it would apply as easily to something no one has learned yet — to any possible result of mathematics, if not to any sort of contingent matter of fact: I could recover the memory of the proof of the Birch/Swinnerton-Dyer conjecture, for instance; though presumably not of how I'd spend the prize money once I published it. — If the soul[70] is boundless and immortal and swims in the sea of eternal truth, time cannot apply to it.

(Leibniz remarks "every mind is of unbounded duration.")[71]

Admittedly this creates a problem with why you understand premises before conclusions: what kind of ordering is introduced by the arrow of logical inference? Is it the same as the arrow of time? — it doesn't appear to be, at least, since in a proof many premises may precede a conclusion, and the order among them is somewhat arbitrary, not necessarily linear. (Linearity is an artifact of exposition.)

This has something to do with the P/NP distinction, about which — ha! — more anon, but for the moment note that one way of stating that[72] is to say there is an inherent difference in difficulty between finding a proof and verifying one. — The latter is straightforward and

---

[70] "The soul" has several different meanings, and the one I take seriously (Aristotle's) is rather different from Plato's; let alone from what Catholic theology derived from it. But we're playing by Plato's rules here.

[71] Loemker, p. 160.

[72] See Jan Mycielski, "The meaning of the conjecture P≠NP for mathematical logic." *The American Mathematical Monthly*, **90** [1983], 129-130.

leads from a leaf of the tree back to the root, which is linear; the former is a search, involves tracing a path from the root of the tree to the leaf that holds the solution, and is in general exponential.

In this sense the distinction between remembering and inventing is just which computation is harder. Time reversal is not a symmetry of the problem.

(Unlike quantum field theory: obviously this argument only occurred to me because I knew that Feynman identified positrons with electrons running backward in time, but the situations are not isomorphic.)

I.e. in following a proof we have a series of applications of modus ponens: A, A —> B, therefore B. We write this down and it looks linear, but if we turn it upside-down the illusion evaporates: from B there are an arbitrarily large set of pairs B <— A, A to derive it from. Even when we can bound the number, as in a procedure like the construction of a semantic tableaux, the complexity of the search grows exponentially in length.

It is for some reason like this that you can remember where you came from but not usually where you are going. — Penrose has an elaborate argument about Fourier decompositions and the wave equation, but Patti Smith is more succinct: "I don't fuck much with the past but I fuck plenty with the future."

Personal immortality

The argument of the *Meno* is that the soul is independent of time. It says both that you always have existed and that you always will.

Young children seem to believe this instinctively. A Pythagorean belief in metempsychosis is as natural as primitive animism.

When I was a child my sister and I would address one another at the breakfast table: "When I was a bird, I used to go like this [making swooping motions with our forks]." — Of course isn't that a peculiar use of "when"? It seems to point not so much to past or future as to some location elsewhere in the manifold of possibility; somewhere sideways in time.

The mental and the physical

It is not a *natural* isomorphism, in exactly the categorical sense: the functor that relates the mental and the physical is contravariant, and reverses all arrows.[73] Thus Kierkegaard said the tragedy of life was that it is lived forward but understood backward.[74]

$$\{\ldots\}$$

This can be illustrated by the example with which Eilenberg/MacLane motivate the discussion in their original paper on category theory:[75] consider the class of vector spaces over a fixed base field F, together with the natural mappings between them, i.e. linear transformations

$$V \underset{f}{\to} W$$

which are closed under composition. For any vector space V, there is a dual space V* consisting of the linear mappings

$$V \underset{\phi}{\to} F$$

which is isomorphic to V, albeit in a way which depends on a choice of basis.

---

[73] This is not very different in spirit from what Russell said about the mental and the physical simply being different schemes for grouping relations, though it reflects changes in mathematical fashion.

[74] According to Susan Sontag, Walter Benjamin referred to memory as "reading oneself backward". Same thing. — The (already classic) cinematic expression of the thesis is Christopher Nolan's *Memento* [2000].

[75] Samuel Eilenberg and Saunders MacLane, "General Theory of Natural Equivalences", *Transactions of the American Mathematical Society*, Vol. 58, No. 2 (Sep., 1945), pp. 231- 294.

Then for any mapping

$$V \xrightarrow[f]{} W$$

there is an induced mapping

$$W^* \xrightarrow[f^*]{} V^*$$

defined by

$$f^*\big(\phi : W \to F\big) = \phi \bullet f : V \to F$$

A similar principle is at work in the relation of the mental and the physical: there is a mapping from the state of the world to the state of the mind at any moment; the natural development in the physical world is the dynamical evolution map from the state of the world at one time to the state of the world at a later time (in quantum mechanics this is, literally, a linear [unitary] operator mapping the Hilbert space of physical states onto itself); the corresponding induced mapping on mental states is the map from the perception of that state at the later time to the perception at the earlier time — i.e., memory.

$$\{\ldots\}$$

Put another way: given two states of the world s, s′, with s′ later than s, causality provides a mapping from s to s′. If you are *examining* the states, however, the arrow goes the other way: given s′, you *explain* what brought it about by reference to (remembered) s.

The problem of the relationship of mind to body is there already in the question of how a physical process can represent a computation; and *that* — really — is quite hard enough.

(Why it never ceases to amaze me that we can *build* computers. That these can be represented within physical systems.)

Awareness

Leibniz, "A Fragment on Dreams":[76]

> Sleep differs from waking in that when we are awake everything is directed, at least implicitly, toward an ultimate goal. But in dreaming there is no relation to the whole of things. Hence to wake up is nothing but to recollect oneself ... to begin to connect your present state to the rest of your life or with you yourself. Hence we have this criterion for distinguishing the experience of dreaming from that of being awake — we are certain of being awake only when we remember why we have come to our present position and condition and see the fitting connection of the things which are appearing to us, to each other, and to those that preceded. …
>
> [He remarks on visions in dreams, how they are superior to waking life in that respect, and concludes:]
>
> [Such visions] are sought by the waker; they offer themselves to the sleeper. There must therefore necessarily be some architectural and harmonious principle, I know not what, in our mind, which, when freed from separating ideas by judgment, turns to compounding them. A reason must be given why we do not remember waking experiences in a dream but do remember the dream when awake.

Here compare separating/compounding to serial/parallel.

---

[76] Loemker pp. 113-115.

When you awaken, your mind wanders in a kind of superposition. And then — abruptly — it observes its own state. The wave packet reduces. You are conscious.

I have to wonder whether Von Neumann had this in the back of his mind when he described the measurement process.[77]

---

[77] Cf. Chapter VI of *Mathematical Foundations of Quantum Mechanics*; Princeton: Princeton University Press, 1955.

Nietzsche says that consciousness is a recent development of the organic and thus unfinished;[78] again, a kind of work in progress. — This however should be read as a gloss on the Socratic precept, to know yourself — to be able to clearly perceive and judge yourself — not simply to act without reflection; to be able to process feedback. — About which nonetheless Nietzsche, the champion of the development of the instincts — the prophet, as it were, of the unconscious — is extremely skeptical.[79]

(Oddly enough this reminds me of Whitehead's remark about inventions like algebraic notation, that advances in science largely consist of eliminating the necessity for thought, of reducing it to the automatic employment of technique; that this allows you to accomplish more with the expenditure of less mental energy.)

---

[78] Cf. *The Gay Science* #11.

[79] Thus for instance the extended rhapsody in *Ecce Homo* about the composition of *Zarathustra*, how it had come to him in a burst of inspiration; as if dictated by a *daemon*. — "If one had the slightest residue of superstition left in one's system, one could hardly reject altogether the idea that one is merely incarnation, merely mouthpiece, merely a medium of overpowering forces." — Compare, of course, Rimbaud, "I is something else," but also Henri Poincaré [below].

Augustine in trying to resolve the paradox of the Trinity proposed an analogy with the tripartite nature of human identity:

> There are three things, all found in man himself, which I should like men to consider. They are far different from the Trinity, but I suggest them as a subject for mental exercise … . The three things are existence, knowledge, and will, for I can say that I am, I know, and I will. I am a being which knows and wills; I know both that I am and that I will; and I will both to be and to know. In these three — being, knowledge, and will — there is one inseparable life, one life, one mind, one essence; and therefore, although they are distinct from one another, the distinction does not separate them. This must be plain to anyone who has the ability to understand it. In fact he need not look beyond himself. Let him examine himself closely, take stock, and tell me what he finds.
>
> [*Confessions* XIII.11]

There is something to that, though in trying to relate mental phenomena to their evolutionary origins we might skip over Augustine and go straight back to Aristotle: any living thing has an identity, what he called the vegetable soul; some clearly have some kind of awareness, the animal soul; humans are self-conscious, and possess a rational soul.

This is not a continuum, exactly, nor is it a simple progression from simplicity to complexity; there are apparent phase transitions, something like the Chomsky hierarchy, but the ordering (this is so obvious that it has become cliché) is treelike, not linear. — One reason why any generalization about biology is probably wrong. —

There are universal principles which are manifested throughout, but this just means that pieces of functionality are scattered all about the animal and vegetable kingdoms. Plants, e.g., are supposed to be passive and insentient, animals capable of simple problem-solving but mainly relying on instinct, not reason. Nonetheless bread molds can solve mazes,[80] some birds really do have linguistic capabilities, and many animals not only have the ability to count but even seem to have a conception of zero as the cardinal of the empty set.

At any rate these appear to be distinct puzzles.

The question of identity applies to anything above the level of the individual cell. How is the organism defined? How do you draw a line around it? What is the predicate demanded by the comprehension principle?

On the one hand I want to say there is some kind of Brouwerian fixed-point theorem involved, in that a suitably-constrained system in which information is exchanged will be found to revolve, as it were, around some central axis; on the other I want to say, duh, the whole thing is grown from a strand of DNA, there is a code, a nucleus, defined physically, and the logical aspect follows from that, but it isn't simply genetic, there is the fact that, e.g., a plant can grow back from a root system, animals have central nervous systems to which signals are referred and from which they emanate. — A colony of cells is

---

[80] See Merlin Sheldrake, *Entangled Life: How Fungi Make Our Worlds, Change Our Minds and Shape Our Futures*. London: Bodley Head, 2020. — The mechanisms are not completely understood, but fungi form extensive networks which can transfer information from one place to another, permitting a form of computation which is remarkably efficient: the mycologist Lynne Boddy made a model of Britain from soil, seeded fungi at the points corresponding to major cities, and grew a connecting mycelial network that looked just like the highway system; others have used slime molds to reproduce the subway system of Tokyo, suggesting novel applications of biological computation. — About this I remarked to a correspondent "Unfortunately fungi despite the advantages of massive parallelism have a slow clock and long cycle time, and probably can't be harnessed to crack RSA encryption. Too bad, I would love to see NSA having to learn how to cultivate slime molds."

distributed, a multicellular organism has some kind of central processing. The line between them is fuzzy (Portuguese men-of-war) but whatever is emergent here, emerges rapidly; probably with the application of the principle of division of labor and its genetic correlate, the regulatory network. — At any rate there is a problem here.

The question of awareness is perplexing since it relates to the distinction between mental and physical. The old German-idealist way of looking at this invoked the distinction between noumenon and phenomenon, the difference between looking at things from the inside and from the outside. Bohr's analysis in terms of complementarity descends from that directly.

Russell's attempt to explain the distinction as one between different ways of grouping relations is similar, but has the added advantage that he saw clearly the need to formulate a kind of mathematical duality between the two viewpoints; noumenon and phenomenon should be something like Fourier transforms of one another. (Or, more generally, the set of all phenomena and the set of all noumena should be equivalent.)

Consciousness, however, is a distinct problem, and mainly has to do with the logic of memory. About that, Augustine again had the first and in some respects the definitive word.

Augustine

If there was a literary invention of self-consciousness, it was in the *Confessions* of Saint Augustine.

Who possessed an amazing philosophical acuity, really no one can touch him between the Greeks and Descartes; in the *Soliloquies*[81] he anticipated the Cartesian c*ogito* and the question of whether to be was to be perceived (he answered in the negative); in *De Trinitate*[82] he raises the problem of other minds, and gives the usual modern solution (analogy).

The philosophical passages in the *Confessions* are mainly in Books X (on memory) and XI (on the Creation and the problem of time). In brief they run as follows:

Animals have awareness [X.6] and process sensory input, but do not perceive meaning. Man, however, can *question* nature. — It is necessary, then [X.8] to go beyond the faculties of sense, the doorways into the soul, to consider memory,[83] the repository of what the senses have admitted. And this is rather mysterious:

> The power of the memory is prodigious.… It is a vast, immeasurable sanctuary. Who can plumb its depths? And yet it is a faculty of my soul. Although it is part of my nature, I cannot understand all that I am. This means, then, that the mind is too

---

[81] In *Augustine: Earlier Writings;* trans. J.H.S. Burleigh, Louisville: Westminster John Knox Press, 2006.

[82] A*ugustine: On the Trinity*, transl. Stephen McKenna, Cambridge: Cambridge University Press, 2002.

[83] Augustine has an expansive interpretation of "memory", and later says [X.14] "the mind and the memory are one and the same."

narrow to contain itself entirely. But where is that part of it which it does not itself contain? Is it somewhere outside itself and not within it? How, then, can it be part of it, if it is not contained in it?

Which states the problem of the unconscious. (Unsurprisingly, he does not resolve it.)

Also whenever he says anything about memory he always qualifies it by mentioning forgetfulness: the *fallibility* of memory puzzles him.

He draws the distinction between memory of events/sensations and functional memory, knowledge of how to do things; he notes [X.9] that besides the data of the senses memory receives and stores things like grammar. Trying to understand where the knowledge of conceptual principles comes from, he reinvents the theory of innate ideas [X.10]:

> How, then, did these facts get into my memory? Where did they come from? I do not know. When I learned them, I did not believe them with another man's mind. It was my own mind which recognized them and admitted that they were true. I entrusted them to my own mind as though it were a place of storage from which I could produce them at will. Therefore they must have been in my mind even before I learned them, though not present to my memory. Then whereabouts in my mind were they? How was it that I recognized them when they were mentioned and agreed that they were true? It must have been that they were already in my memory, hidden away in its deeper recesses, in so remote a part of it that I might not have been able to think of them at all, if some other person had not brought them to the fore by teaching me about them.

He notes among the properties of memory the ability to hold facts which are not present [X.11] but can nonetheless be retrieved; he

gives number [X.12] as an example of knowledge that does not originate in the senses. (No empiricist he.) — He remarks the oddity [X.13] that it is possible to retain both correct and incorrect logical arguments, and to review and compare them, and notes that one can remember a feeling, though this is not the same as experiencing it. (He seems here to have a keener appreciation here of the problematic character of the impression/idea distinction than Hume did.)

When he speaks of the sun [X.15], it is the sun's image he recalls, not the image of an image, but when he speaks of memory — ?! — how does that work? How can memory be present in itself except as an image of itself? — He is also [X.16] somewhat amazed that he can remember forgetfulness; this seems as though it should be self-negating. — He sees, i.e., the self-referential nature of the memory, he perceives this is the central problem. — More, he realizes "I have become a problem to myself" —

> I am not now investigating the tracts of the heavens, or measuring the distance of the stars, or trying to discover how the earth hangs in space. I am investigating myself, my memory, my mind. There is nothing strange in the fact that whatever is not myself is far from me. But what could be nearer to me than myself? Yet I do not understand the power of memory that is in myself, although without it I could not even speak of myself. What am I to say, when I am quite certain that I can remember forgetfulness? Am I to say that what 1 remember is not in my memory? Or am I to say that the reason why forgetfulness is in my memory is to prevent me from forgetting?

but rhapsodizes [X.17]

> The wide plains of my memory and its innumerable caverns and hollows are full beyond compute of countless things of all kinds. Material things are there by means of their images; knowledge is there of itself; emotions are there in the form of ideas or impressions of some kind, for the memory retains them even

while the mind does not experience them, although whatever is in the memory must also be in the mind. My mind has the freedom of them all. I can glide from one to the other. I can probe deep into them and never find the end of them. This is the power of memory! This is the great force of life in living man… .

But then after wondering [X.19] how you can recover memories temporarily misplaced — what is the mechanism here? — he proceeds into a catalogue of the various ways the senses can admit temptation into the soul, displays predictable guilt over erotic dreams, etc., etc., and indulges in the absurdly amplified self-flagellation for which many generations of the heinously repressed have celebrated him.

Book XI is then concerned with the problem of creation *ex nihilo*: he dismisses the question of what God was doing *before* the Creation as a pseudoproblem, because it is only *within* the cosmos that time can be said to exist; rejects the Platonic picture of an artisan employing tools to shape the world from preexisting materials; and [XI.5] concludes (following the Gospels) that "you spoke and they were made. In your Word alone you created them." — Which is a very provocative suggestion of ontological sleight of hand.

Speaking, however, no matter whether any voice was heard, involves some kind of *motion*, which brings him back to the problem of time. — And here [XI.7] he makes remarks which seem to me to bear an uncanny resemblance to current cosmological speculation, though why, precisely, should be the subject of another postcard. — But settles on the question: how does the Word, which is timeless, somehow set in motion the engine of Becoming: "Yet the things which you create … do not all come into being at one and the same time, nor are they eternal."

God to Augustine has something like the absolute perspective we associate with Minkowski space [XI.13]: "Your years neither go nor come … [they] are completely present to you all at once, because they

are at a permanent standstill. They do not move on …they never pass at all. … Your today is eternity." — So [XI.14] what *is* time? and what has it got to do with the cosmos? The past is gone, the future isn't here yet, the present certainly exists but is always lapsing from existence. "In other words, we cannot rightly say that time *is*, except by reason of its impending state of *not being*."

Here [XI.15-16] he tries to figure out how time can be measured, a thankless chore before the invention of the clock.[84] Retreating from this question in confusion, he wonders again how the past and future can exist. The latter has been seen by prophets, the former by anyone who can describe it by examining the contents of his mind/memory. — Therefore, etc., QED? — No [X.20]:

> …it is abundantly clear that neither the future nor the past exist, and therefore it is not strictly correct to say that there are three times, past, present, and future. It might be correct to say that there are three times, a present of past things, a present of present things, and a present of future things. Some such different times do exist in the mind, but nowhere else that I can see. The present of past things is the memory; the present of present things is direct perception; and the present of future things is expectation.

Returning to the problem of measurement, he rejects the notion that time can be defined in terms of movement — quite the contrary, he suggests — in fact [XI.23] refutes a kind of argument from Mach's Principle, and concludes [XI.24] "Time … is not the movement of a body." He conjectures that one might attempt to measure duration by

---

[84] Compare, though it cannot be taken so seriously, the difficulty faced by a disciple of Lacan in defining self-consciousness before the invention of the mirror.

the lengths of syllables,[85] but gives up finally, and concedes [XI.26] "It seems to me, then, that time is merely an extension, though of what … I do not know. I begin to wonder whether it is an extension of the mind itself." And after further labor concludes

> It is in my own mind, then, that I measure time. I must not allow my mind to insist that time is something objective. I must not let it thwart me because of all the different notions and impressions that are lodged in it. I say that I measure time in my mind. For everything which happens leaves an impression on it, and this impression remains after the thing itself has ceased to be. It is the impression that I measure, since it is still present, not the thing itself, which makes the impression as it passes and then moves into the past. When I measure time it is this impression that I measure. Either, then, this is what time is, or else I do not measure time at all.

Finally he suggests that what moves through time (if anything does) is "the faculty of attention", and provides this vivid summary:

> Suppose that I am going to recite a psalm …. Before I begin, my faculty of expectation is engaged by the whole of it. But once I have begun, as much of the psalm as I have removed from the province of expectation and relegated to the past now engages my memory, and the scope of the action which I am performing is divided between the two faculties of memory and expectation, the one looking back to the part which I have already recited, the other looking forward to the part which I have still to recite. But my faculty of attention is present all the while, and through it passes what was the future in the process of becoming the past. As the process continues, the province of memory is extended in

---

85 This remained a problem well into the seventeenth century. Galileo, of course, referred to his pulse, but it is conjectured his musical training (he was a composer for the lute) gave him an enhanced sense of minute intervals, derived from the necessity of playing sixteenth and thirty-second notes in time.

proportion as that of expectation is reduced, until the whole of my expectation is absorbed. This happens when I have finished my recitation and it has all passed into the province of memory.

{…}

To summarize: consciousness presupposes memory; memory is paradoxically self-referential; time involves similar paradoxes; the flow of time is an artifact of the function of memory.

All this is completely correct.

The intention here may not be obvious. This isn't a hermeneutic exercise. I'm not trying to *interpret* Augustine — learn his language, explicate his precise meaning, understand him on his own terms. What I'm trying to make clear is that if he spoke *my* language and knew what *I* know, he'd be saying the same things that I am.

Only better, of course. And in Latin.

So: Augustine identified the problem of consciousness with that of memory, and related both to the problem of the nature of time.

Obviously he was right about everything, and the appropriate place to begin, accordingly, is with an idea of Thomas Gold.[86]

Gold wished to understand how temporal succession arises in theoretical physics, and suggested that one might look at it as follows: suppose you have a set of index cards, and on each you write a description of the instantaneous state of the world. — Then you shuffle them. — How do you put them back in order? Where in the data on the cards do you find the arrow of time?

In physics, at least, it is more natural to think of this as a movie: you cut it into individual frames, snapshots of an action; scramble them; and then try to reassemble them in the proper order. — Of course this is the problem of montage, the Russian school addressed this with the famous Kuleshov experiments,[87] and taken as a question of cinema there is not a single unambiguous answer: different orders have different meanings, and the arrangement is a matter of art. — Taken as a question of physics, it is problematic, but less ambiguous — in the usual example of frying an egg, there's no question what order the snapshots should assume, the difficulty is in explaining why that order is correct. Nonetheless under the usual laws of mechanics the problem is solvable.

---

[86] See *The Nature of Time*, Thomas Gold, ed. Ithaca: Cornell University Press, 1967. This volume summarizes the proceedings of a symposium organized by Gold, and contains not only papers presented at the meeting but the very animated discussions which followed them. Since Gold did his best to invite all of the world's leading theoreticians, these discussions are very interesting indeed.

[87] See (for instance) Chapter 6, "Soviet Cinema in the 1920s", in Kristin Thompson and David Bordwell, *Film History: An Introduction*. [New York: McGraw-Hill, 1994.]

Taken as a question of conscious experience, however, it is trivial. Mental states *refer* to one another, and that defines a natural temporal direction, provided by memory.[88] (Which is, physically, based on irreversible processes which have the thermodynamic arrow of time built into them; allowing us to cheat, and skip over the difficulties of the problem as it presented itself to Gold.)

---

[88] Of course if memory *cannot* be relied upon, the questions of order and meaning again become problematic. This is exactly the problem faced by Guy Pearce with his Polaroids in *Memento* [Christopher Nolan, 1999]. As the authors were undoubtedly aware. — See also the collected works of Philip K. Dick.

I wouldn't say that consciousness is prior to temporal order, or vice versa; rather that the two ideas are inseparable.

I don't mean to say that consciousness is the missing link that determines the direction of time; or not that exactly. — What I mean, mainly, is that though we can imagine physical worlds in which time does not have a direction, these are not worlds in which consciousness is possible. Consciousness requires memory, and memory requires [provides] direction.

How it is that when you remember two things, you remember — or take it for granted you *ought* to be able to remember — which one came first. Because if memory were perfect, you would remember that when one occurred you remembered the other had.

Here as always we are talking about an ideal mathematical model of the phenomenon.

So the most elementary conception of the way that consciousness registers/sorts/imposes temporal order is exemplified by the ordinal numbers.

Von Neumann defined an ordinal as the set of all preceding ordinals, grounding the series in the empty set.

Literally, the sequence is generated by bracketing:[89]

$$\emptyset, \{\emptyset\}, \{\emptyset,\{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset,\{\emptyset\}\}, \ldots$$

so that **m** < **n** if and only if **m** is contained in **n** if and only if **m** is an element of **n**, and each ordinal *contains* its predecessors. — This is it exactly.

$$\{\ldots\}$$

The infinite descending epsilon-chain is a major no-no of set theory, of course, and though it's obvious how we avoid this in the case of the ordinals, in the case of conscious states it requires postulating a first memory.

The problem with that is not that it doesn't exist, but that it is usually an isolated incident which is not an element of a connected progression, as the model requires. — Memory functions erratically in early childhood, consciousness in consequence takes a while to turn itself on.

---

[89] This may, by some coincidence, be more or less the sense in which Husserl used the notion, but of course I'll never bother to find out. Let's simply give credit where credit is due: consciousness is, indeed, consciousness-of (and thus self-consciousness is consciousness-of consciousness itself); and "bracketing" isn't a bad way to describe the elementary mental operation.

So this is the other thing we have to say about this model: it represents an ideal limit, something that is only approximated by the phenomenon as we observe it in the physical world, and thus in the mental world that we inhabit. Consciousness as we possess it is fragmentary and imperfect, because memory is fragmentary and imperfect.

Perhaps the best representation of the idea in the cinema is the famous shot in *Citizen Kane* of Kane/Welles walking down a corridor in his mansion and passing between two long mirrors, which produce an infinite series of reflections of reflections of a myriad of Kanes, stretching off into obscurity.[90] — In reality, with mirrors which are not perfectly reflective, the images will grow fainter as they recede in the apparent distance. So also memory dims — though not, really, in the same way, it is not an analog copying process, like making a copy of a tape; it involves some kind of digital representation, as if there were some sort of sentence recorded at every instant of awareness that contains a summary not only of what is transpiring at that instant but also a thumbnail description of what preceded that — or rather: of *what you thought* about what preceded that. — Else the past would fade into an undifferentiated blur. Instead we have a sort of album, with photographs and jottings and guitar licks and (Proust) odors, in my case probably of farts.

But in the moment, in the experience of awareness, what connects one moment to what preceded it is just that: the near-perfect retention of the previous thought, the present thought, and the perception that the latter *contains* the former; *is about* the former; *refers to* the former. — At bottom the relation is that of metalanguage to language; there is a

---

[90] It is also a metaphor for the multiple-perspective insect-eye view the film creates of Kane, an intimation of the now-ubiquitous idea of the multiverse, and much else besides. But I'm doing this one first.

process of *evaluation* going on.[91] — To represent it naively, supposing I have just formed the internal sentence "The coffee cup is on the table to the left of the laptop," the present sentence is something like "I am reaching for the coffee cup with my left hand, having asserted that 'The coffee cup is on the table to the left of the laptop.'"

(With which I do so. — Cold. — Into the microwave again.)

{…}

In reality the mechanisms that create and store memories are (out of physical necessity) complex, and employ representations whose nature and functioning no one as yet understands. — This has to be the case, because, for instance, if we take the ordinal representation literally then the amount of storage required to keep your experiences straight would grow exponentially with time. This isn't possible, of course. — In fact even a linear growth is impractical; this is the point of Borges' story "Funes el memorioso". One cannot absorb the past if it is so vividly present that you are still living in it.[92] — A back-of-the-envelope calculation shows that the brain would already have popped like a balloon in infancy if it retained everything; a ruthless process of data compression which proceeds continuously is an engineering requirement. This entails the necessity of generalization — you have

---

[91] Another purely logical way to provide an arrow for time is the one implicit in Kripke's semantics for intuitionistic logic:, which can be interpreted as referring to a set of states of knowledge: you picture a set of elementary propositions, and a series of partial valuations whose domains are ordered by inclusion. The problem of the sea-battle, for instance, is resolved by noting that the truth or falsehood of the statement belongs to an extension which has "not yet" been performed, because (the tautological formulation) "the future" is simply what you don't know about — that to which a valuation has not yet been assigned. — Or: "past" = "that to which a valuation has been assigned"; "future" = "that to which a valuation has not been assigned"; "present" = the process of valuation. Another expression of the principle that information *by definition* always propagates forward in time.

[92] The problem would be solved if the longer you lived, the faster you thought, i.e. if at a minimum the speed of thought increased linearly as a function of time. Of course if anything the opposite is the case.

no room for the heap of facts, you must summarize them succinctly or forget them entirely — thus induction — and the sketchy and imperfect character of any recollection.[93]

But that's real life, and this is more like mathematics. The question to be addressed by speculation is whether the ideal of consciousness is attainable; whether it can be physically realized. — What is the end point of mental evolution? Because it obviously isn't me.

In fact that's the most obvious thing about consciousness: that you haven't attained it. That you are always stumbling around half-awake, never really aware of who you are and what you are doing. That it is all out of control, that you have no grip on it.

---

[93] Of course this may be advantageous, for the same reasons that the imperfection of DNA transcription is a fundamentally creative principle. — Nietzsche: "The artist needs the infidelity of memory in order not to copy but to transform nature."

The feeling of not being on top of things, of reeling out of control — what Heidegger called *Geworfenheit*, the sense of having been *thrown* into the world[94] — is the feeling of never being able to catch up; the incompleteness of consciousness is at the root of that. — If it were possible to retain everything and perceive it all at once — if we were fully conscious — complete understanding might be possible. — But it isn't, and it isn't. At best we perceive things in flashes, and perceive only fragments.

Schopenhauer said philosophy begins upon a minor chord; in truth the notes sound indistinctly, and never all at once.

The imperfection of memory guarantees that you are never really yourself.

The imperfection of memory: this is why you always have the feeling that you aren't all quite there; that you are out of control, somehow, off balance and reeling forward against your will — *losing track* of yourself. You want to say, *get a grip*, but you never can: the present is *there*, but slips away and fades into the past before you can grasp it. Because there is this mathematical ideal, that is what *ought* to define you, and you are only an imperfect realization of it.

It is a kind of broken symmetry. You never attain the Platonic ideal.

---

94 Oddly enough — or not — this aspect of the human condition found its perfect cinematic expression in a silly scifi action movie — *Predators* [Nimród Antal, 2010] — which commences abruptly *in medias res* with an abducted mercenary warrior (Adrien Brody) regaining his senses, confused and disoriented, to find himself plummeting through the atmosphere of an unknown planet, toward a surface where he is about to be hunted by — well, you guessed it —

A shape not seen directly; its shadow only.

Lived experience is not like a movie that you can pause, and step forward and backward. But it *ought* to be.

The principle is no different (in kind, anyway) from the limited realization of the Turing machine: a physical computer does not have an infinite supply of scratch paper. — But it *ought* to, and we reason about it as if it did. — This is one of those instances in which the difference between large finite models and the infinite limit is an annoying distraction from the real principles at issue.

(We do have to leave open the possibility, however, that there's some subtlety here that can only be addressed by taking these limitations into account; as the theory of computation develops another dimension when you consider constraints of time and space.)

The definition of an ordinal is inductive, i.e.

(i)   Ø is an ordinal.

(ii)   If **x** is an ordinal, then the union of **x** and {**x**} is an ordinal.

Which constructs the ordinals as Peano did the integers, from the bottom up. One can invert the procedure, and work from the top down; this is the variant which in programming is called recursion.

At first glance it looks like a flagrant violation of the vicious circle principle, since it defines a concept in terms of itself.

I.e. the factorial function[95] on a positive integer argument[96]

$$n! = n*(n\text{-}1)*(n\text{-}2)* \ldots * 1$$

so that, e.g.,

$$5! = 120$$

can instead be specified (in Lisp notation) as

```
(defun factorial (n) (if (= n 1) 1 (* n (factorial (- n 1)))))
```

---

[95] Dana Scott somewhere remarks that the factorial is the most overdefined function of all time.

[96] Euler, no disciple of Wittgenstein, extended the definition from positive integers to any real or complex argument; in this reinvention it is referred to as the gamma function.

which can be read as "the factorial is defined as the function of n such that if n equals 1, it is 1, otherwise it equals n times the value of the factorial for n minus 1."[97]

Similarly the Fibonacci sequence

$$1,1,2,3,5,8,13,21,34,\ldots$$

can be defined as the function

$$f(1) = 1$$
$$f(2) = 1$$
$$f(n) = f(n - 1) + f(n - 2)$$

---

[97] It is assumed that the arguments are positive integers.

Each conscious moment *contains* the one that preceded it.

Linguistically, each statement is *about* the last. The tower of inclusion is metalinguistic. — Viewed subjectively "now" is the assignment of a truth value.

(Why fucking with temporal order leads you straight into the Cretan liar.)

Consciousness and recursion.

(Here as everywhere we owe the first recognition of the importance of the idea to Kurt Gödel.)

Since I lacked access to computers until I could actually buy one, I came to coding as a pastime relatively late in life, and thus wrote my first recursive function call when I was trying to draw a picture of a (self-similar) tree.

For a simple binary tree, the algorithm is as follows: you define a function which takes as inputs a length, a minimum length, a multiplier less than one, a starting-point, a direction, and an angle; if the length is less than the minimum length, it returns without drawing anything (this is the cutoff that prevents the program from running forever); otherwise it draws a line segment from the starting point at an angle left by the given angle from the input direction and calls itself with new inputs the length times the multiplier, the minimum again, the multiplier again, the starting-point the endpoint of the segment just drawn, the direction that of the segment just drawn, and the given angle; when the function called to the left returns from execution it repeats the procedure on the right. Thus the function calls itself not once but twice, and the result of its execution is to draw a tree which (modulo the asymmetry introduced by the finite cutoff) draws a tree which consists of a line segment to which are attached two isomorphic copies of itself, one angled to the left by a fixed amount, one angled to the right. — The ungainliness of this verbal description explains the usefulness of programming languages, which are simultaneously simpler and more precise than English prose.

At any rate the first time I did this it was peculiarly difficult (the fact that I was trying to do it in a mediocre implementation of a bad

language[98] didn't make it any easier), and when I reached the stage in the code where it committed the sin of self-reference and *called itself* — not once but twice — I felt a peculiar vertigo, as if my eyeballs were trying to turn around and look back into my head.

And of course as a veteran of philosophical perplexity I recognized this instantly! as precisely the disorientation I had always felt when I tried to think about the problem of consciousness, to think about thinking, exacerbated enormously when later I began to worry about the foundations of mathematics — can you reason about logic? make mathematical models of mathematics?

Because *it is all the same problem*.

What is my argument here? — *It gave me the same headache*. — Oddly enough though it is more than a trifle silly it is the most powerful argument of all.

_____

[98] Microsoft Basic, on my first Mac.

Consciousness is consciousness-of; self-consciousness is consciousness of being conscious.

Descartes based his existence not so much on the fact that he thought, but on the fact he was *aware* that he was thinking.

The sense of self is that of the subject as object.

The degree of consciousness is indexed by the efficiency of the implementation of recursion.

(Depth of stack is one determinant. — How far short we fall of the "potentially infinite".)

There is (the ideal limit) some perfect mathematical model, which would at the least require perfect eidetic memory. And then the actual realization of the idea, which is erratic, and works by fits and starts.

Something in this captures the *difficulty* of the mental: there is a dual foundation for the world, but it has been obscured by a kind of broken symmetry.

(Of course the material also has its limitations — in effect in quantization we posit the Hamiltonian picture as an unrealizable ideal as well.)

My old idea was "partial representation". Like the realization of the Platonic Idea of the circle.

Not unrelated: the denial by Hume and Nietzsche, among others, of the reality of the Ego. The argument is valid as far as it goes — there

is reference to an ideal object which does not "really" exist — but does not grasp the simplification made possible for, say, a geometry, when ideal points are adjoined. — It is something like what happens when you stare too closely at a set of colored squares — they are just pixels, after all — but if you move back and take the set in as a whole, it resolves into a picture. You see a human face.

In principle you leave a physical record behind you, in the form of a series of patterns of neuronal excitation. (Say.) But all this is like the print on paper, there is something it points to; something that it means.

What is conscious is what enters into memory. — That's simple enough. That is what Augustine said.

A bit *too* simple, as the overwhelming evidence for ongoing unconscious labor shows. But something like this must be true nonetheless.

There is a kind of SIMP[99] model that naturally imposes itself. A function is being evaluated. It depends on many variables, each the result of a functional evaluation. — And so on, recursively. — The separate subevaluations can be performed in parallel; and surely are, all the evidence suggests it.

But what defines the fixed point, the center? — The traditional objections to the Ego don't address this problem at all, why the mass of thoughts should even *appear* to be directed toward a center. — Why there is one voice in your head that is the loudest.

E.g. Nietzsche (*Late Notebooks* 34[123]): "man is a multiplicity of forces which stand in an order of rank… . All these living beings must be related in kind, otherwise they could not serve and obey one another … . The concept of the 'individual' is false. … the center of gravity is something changeable; the continual generation of cells, etc., produces a continual change in the number of these beings. And mere addition is no use at all. Our arithmetic is too crude for these relations, and is only an arithmetic of single elements." — Which actually contradicts itself: the equilibrium may not be static, but there is some kind of "order of rank" defined by the logic of the dynamics. — But he is quite acute on these questions: [24] "The logic of our

---

[99] Gaah, always a mistake to introduce an acronym. — I think it's "Single Instruction Multiple Processor" but don't quote me. — The *idea* is that one processor delegates execution of a set of separate and logically independent subprocesses to a set of processors which execute their computations in parallel, and return the results to the calling program without mutual interference.

conscious thinking is only a crude and facilitated form of the thinking needed by our organism, indeed by the particular organs of our organism. For example, a thinking-at-the-same-time is needed of which we have hardly an inkling."

Later he refers to "the development of consciousness as an apparatus of government"; almost too cute —

$$\{\ldots\}$$

But really the arguments against personal identity are the same as the ones that would try to persuade you that the rotation group in three dimensions does not "exist" either; for all that you need to study its representations to understand the quantum mechanics of angular momentum.

Still, why is there only one center of attraction? It seems like a fixed point theorem, something like the reason when you stir the cream in a coffee cup there's a single point at the center of the vortex that isn't turning.

A bundle of functions which are calling one another. The arrows of direction converging on a single one to which all the results are returned.

A related mystery: the evolution of arithmetic.

Where does counting come from? how did it arise?

The operation of abstraction. In the lambda calculus this is simply function abstraction, lambda: to take the functional form as an object in itself.

Something like this is made possible by the reduction of memories to data; by being able to examine these like sense data.

Perceiving the similarity of the five cows to the five trees. Perceiving the one-to-one correspondence.

But how would you *invent* counting? There must be some evolutionary pathway that leads to it. How could you discover it?

The imperfection of consciousness is the imperfection of memory. — In which one must include the imperfection of the mechanisms that retrieve and examine memories, but — I have listened to the Haydn symphonies dozens of times, so that anything I hear sounds familiar, but I have nothing like the memory of Wittgenstein, who could whistle through his favorite passages from Beethoven and analyze them as he went; and Mozart was undoubtedly better, more completely the master of everything he had ever heard. — But could either do calculations in his head, like Euler? and though my usual example is Von Neumann, I am usually thinking "if only I could keep track of everything like that" but he wasn't particularly musical, and even his mathematical intelligence was curiously limited, he had very little geometric intuition, for instance. — So it seems very obvious that there has never been a fully conscious human individual. — Save in the imagination of Borges, who could see the way that this would be an affliction.

Matters of degree: my dogs are conscious. (How do I know this? because they can be *unconscious*; I have often watched them dreaming.) But do they inspect their internal states in the same way I do? — Yes, but not as well. I have a larger brain, a better memory, and (in the use of language) employ a better shorthand for summarizing previous experience. — If there is a phase transition here, it is in whatever makes possible the use of language — in symbolic representation. — But basically our feeling is that *Natura non facit saltus*; and even when she does, she approximates the jumps by continuous functions.

Though dogs are not as intelligent as birds in some respects — and so on — and so on. — The picture (Dawkins?) of a multiplicity of functionalities, developed and distributed among a multitude of species.

$$\{\dots\}$$

Some Cartesian I am if I don't revert to paranoia, however. — Given the dependence of consciousness on the chain of memory, then — what if it's fake? — Of course this is the entire oeuvre of Philip K. Dick.

What you have to wonder here is whether it really makes a difference. Superficially you have a kind of grue paradox applying to the past, at every step in the historical record you can insert a discontinuity, but isn't it Dick's point that fiction here has the same ontological weight as reality? that this is where the two interpenetrate? — The replicants in *Blade Runner* with their fetishistic reliance on their yellowing collections of photographs. — Rachael remembers a mama spider eaten by her babies, and this is "really" a memory of Tyrell's niece. — But is it? doesn't it define Rachael just as well as it did its originator? Doesn't the shared memory entail a shared identity? an overlap

between two persons? Isn't the consciousness of the one simply a clone of the consciousness of the other?

I don't feel like chasing Dick down this particular rabbit-hole at the moment — for that you do need the drugs — but my basic feeling is that the anxiety you are supposed to feel here is unfounded.

The deck of cards (again)

A natural way to look at it is to think of a strip of film, representing a series of snapshots taken along a four-dimensional world line. (Life is one long tracking shot.) — Consciousness (assumed to be external) is then a light illuminating the individual frame (the present) — the bulb in the projector, what throws the picture onto the screen: "Attention is *here*."

This is incorrect. The light — the act of projection — is superfluous. The idea of order emerges from the way the frames refer to one another; from the way they fit into one another. This putting-in-order is *itself* consciousness. — Which is not outside the picture, but part of it.

In the spirit of recalling all my favorite episodes of cognitive dissonance, here I must interpolate the observation that, during the period when I was first thinking about all this, I was flabbergasted to discover an exposition of the essential idea in a ridiculous scifi movie I happened across on the late show — namely *Escape from the Planet of the Apes* [Don Taylor, 1971]. — In this, after it is realized that the ape astronauts who have landed on the Earth didn't come from another planet, but rather our own future, an industry-standard Scientist With A German Accent ("Dr. Otto Hasslein") is dispatched to the television studio to explain the nature of time to the national audience. — "I think that time can only be fully understood by an observer with the godlike gift of infinite regression," he says. And offers as an illustration a painter at work upon a landscape, who realizes the picture is not complete unless he includes himself in it; painting a picture of himself painting a picture of, etc. — As an explanation of how consciousness constructs the temporal ordering, this is essentially correct.[100]

---

[100] This episode in the Apes franchise was written by Paul Dehn, noted mainly for screen adaptations like *Goldfinger* and *The Spy Who Came In From The Cold*. How inspiration possessed him on this occasion one can only guess. Maybe he was a neighbor of Fred Hoyle's.

This isn't quite it, of course. The fundamental paradox is one that Francis Crick used to emphasize, one exhibited in another (better) silly movie, Woody Allen's *Everything You Always Wanted to Know About Sex But Were Afraid To Ask* [1972]. In this there is a scene in which, while a guy is trying to achieve success in a sexual encounter, a crew of operators in something that looks suspiciously like NASA Mission Control frantically shout orders and throw levers and switches in the control room of his brain. — This, Crick remarks, is a very natural picture, one which is automatically assumed: there is a kind of nerve center to which all the messages are relayed, and someone is sitting there watching all the monitors — the master of the (inverse) Panopticon, see Fritz Lang in *The 1000 Eyes of Dr. Mabuse*, *The Matrix*, etc. — issuing commands. — Yes, asks Crick; but *who* is sitting there? — Well, *you* are; there *is* a regress. But in the nested chain of control rooms (the inclusion relation runs the other way in this picture) each one contains the *next* one, the one corresponding to the next instant; just as the definition of the factorial refers to its value at the preceding integer. A new one is created, as it were, at every step; (n + 1) within n.

It's possible to play games with a video camera and a monitor that illustrate the paradox directly, as Hofstadter pointed out.[101] — Still photographs don't really do it justice: when you aim the camera at the screen that is displaying its output, you get the infinite descending chain of pictures within pictures, sure, but they can be skewed with respect to one another, which is disorienting, and the dynamical effects produced by *turning* the camera are uncanny. — Most disconcerting of all, I discovered, was to employ the automatic zoom. The coordinated creep of the entire visual field is the only thing I have ever seen, cold

---

[101] See Douglas Hofstadter, *Gödel, Escher, Bach*. [New York: Basic Books, 1979.]

sober, that reproduces the effects of LSD: the famous Hitchcock trick[102] on steroids.

---

[102] The effect produced by simultaneously moving the camera toward an object while zooming out to compensate, thus optically amplifying the sense of depth in the field of view. First used in *Vertigo*, and imitated a thousand times thereafter.

The consistency of visual metaphor

The representation of time as a spiral, or a concentric set of circles, like tree rings [*Vertigo*] — is the self-similar: that which can be mapped into itself one-to-one.

(In three dimensions the onion and its layers.)

Are other models possible? Yes, certainly. You can see the hint of one in the final confrontation in *Fight Club* [David Fincher, 1999], when Ed Norton shouts "You're just a voice in my head!" and Brad Pitt replies calmly, "And you're just a voice in mine."[103] Because there is no *logical* objection to the idea of two control centers, each nested within the other.

It is easy to write something like this in Lisp, e.g.:

```
(defun even (n)
    (if (= n 0) 0
      (+ 1 (odd n)))))

(defun odd (n)
    (+ 1 (even (- n 1)))))
```

yielding

```
?    (odd 20)
     39
?    (even 20)
     40
```

and so on. — One or the other of the recursions must be grounded, in other words, but it is arbitrary which. Nor is there any difficulty in generalizing this to any finite number. The real question is how the separate identities could be distinguished, if each has access to the memories of the other, and in fact multiple personality disorders, insofar as they cannot be consigned to urban legend, appear to involve a partitioning of memory and consciousness, which mathematically would correspond to the partitioning of a graph into disconnected components. The idea that alternating personalities might, e.g., see the world differently, and (adopting Russell's simile) switch back and

---

[103] This exchange occurs in the movie, but not the novel. Presumably it should be attributed to Fincher or the screenwriter, Jim Uhls, and not to Mr. Palahniuk.

forth from red- to blue-tinted spectacles, is a trifle more baroque, and a little more difficult to enter into, if not necessarily to understand.

In Shakespeare, the play that is *about* self-consciousness, *Hamlet*,[104] has as the midpoint and summit of its action a play *within* the play, designed by the protagonist to model the action of the play itself. This says everything.

Hegel in the *Phenomenology of Spirit* #80 characterizes consciousness as "explicitly the *Notion* of itself";[105] which I have to admit is the right idea, even though there is very little else in this work that makes any sense. — Perhaps #177: "A self-consciousness exists *for a self-consciousness*" — ? — At any rate he appears to say that consciousness of self presupposes acknowledgement by another consciousness, something like the mirror principle, perhaps something like the principle I intuited on my tricycle.[106]

But this is Hegel, of course, so who the fuck knows.

Umberto Eco points out that Ian Fleming's curiously poetic reflection in *Casino Royale* — "You are about to awake when you dream that you are dreaming" — is stolen from Novalis.[107] — And what is the point of that? that consciousness returns when recursion commences.

See also *Fichte Studies*:[108] "Consciousness is a being outside of being that is within being. … an image of being within being."

---

[104] Bloom refers to Hamlet as "the hero of self-consciousness."

[105] Translated by A.V. Miller. Oxford: Oxford University Press, 1977.

[106] See below.

[107] Specifically it is "Miscellaneous Observation #14" in *Novalis Philosophical Writings*, ed. Marjorie Mahony Stoljar. [Albany: State University of New York Press, 1997.]

[108] Translated by Jane Kneller.

Fichte himself says something about the I positing itself as absolutely self-positing.[109]

---

[109] He doubtless says much more in the same vein, but why court brain damage trying to dredge all of it up.

Wittgenstein remarks to Waismann [22 March 1930]: "Time is the way memories are ordered. Thus time is given in immediate connection with memories. Time is, as it were, the form in which I have memories." — His point here is that there are several different concepts of time, depending on whether the context is that of physics, of personal experience, etc.

Nietzsche [*Twilight of the Idols*]:

> *The error of imaginary causes*. To begin with dreams: *ex post facto*, a cause is slipped under a particular sensation (for example, one following a far-off cannon shot) — often a whole little novel in which the dreamer turns up as the protagonist. The sensation endures meanwhile in a kind of resonance: it waits, as it were, until the causal instinct permits it to step into the foreground now no longer as a chance occurrence, but as "meaning". The cannon shot appears in a *causal* mode, in an apparent reversal of time. What is really later, the motivation, is experienced first— often with a hundred details which pass like lightning—and the shot *follows*. What has happened? The representations which were *produced* by a certain state have been misunderstood as its causes.

"In fact we do the same thing when awake," he continues, but this is less interesting. — The intriguing idea here is that consciousness doesn't simply presuppose temporal ordering, but to some extent *imposes* it. Because in the unconscious dream state the temporal ordering of events does not exist; they are, as it were, just the hypothetical notes scribbled on index cards. It is the act of conscious recollection that puts them into sequence.

Given the obvious analogy to film editing, it seems as though the Kuleshov principle should apply: the raw material of a dream could be rearranged in such a way as to have a completely different interpretation.

(Or if temporal connections are made in the unconscious state, if one card is clipped to another, it's done randomly, as an experiment. Thus the fractured narrative sequences of dreams.)

Hegel was right that consciousness has a history, that it is a work in progress, something in development. I doubt, however, that he realized it has been evolving for a couple of billion years, and that nothing about it is unique to *Homo sapiens*.

Because though on the one hand you suspect some kind of phase transition to consciousness in the higher mammals (perhaps also invertebrates, see the curious case of the octopus), on the other it's obvious that this is a biological invariant, coded in at the lowest level:

> A bacterium is so small that its sensors alone can give it no indication of the direction that a good or bad chemical is coming from. To overcome this problem, the bacterium uses time to help it deal with space. The cell is not interested in how much of a chemical is present at any given moment, but rather in whether that concentration is increasing or decreasing. After all, if the cell swam in a straight line simply because the concentration of a desirable chemical was high, it might travel away from chemical nirvana, not toward it, depending on the direction it's pointing. The bacterium solves this problem in an ingenious manner: as it senses its world, one mechanism registers what conditions are like right now, and another records how things were a few moments ago. The bacterium will swim in a straight line as long as the chemicals it senses seem *better* now than those it sensed a moment ago.[110]

This must be nearly the simplest possible implementation of the sense of time: two states, present and past, and execution of a single command ("change direction") conditional on the computation of a gradient. But it shows that the fundamental operation, the comparison

---

[110] Peter Godfrey-Smith, *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York: Farrar, Strauss, and Giroux, 2016.

of snapshots, is possible even in bacteria. And everything follows from this.

The Interpretation of Dreams (1973)

There's a famous story about a horse named Clever Hans, who was supposed to be able to do arithmetic: some Kantian would ask him what was seven plus five, and he would paw the ground twelve times. This baffled and amazed everyone who witnessed it until someone noticed that, if he couldn't see his trainer, Hans didn't know where to stop: he read the correct answer from subliminal cues. He didn't really know how to add, he just knew what his trainer wanted to hear.[111] — People are (usually) more clever than horses, and can interpret subtler signals.

When I first read Wittgenstein I spent about a month going through the *Philosophical Investigations*, taking copious notes and arguing with the author line by line. I don't think I ever finished the book, at least not on that occasion; the project devolved finally into making notes upon my notes upon my notes. When I had written thirty or forty thousand words, I figured I had done enough.

The experience taught me I could turn the reading of anything into a melodramatic struggle between myself and the author over the real meaning of the text, however, and for a while I read everything that way, because I had developed a narcissistic fascination with watching myself think. — Or at least pretending to think. — "Wrestling is not a sport, it is a spectacle," says Barthes.[112] No shit.

Thus when I decided to make a study of Freud — not that I thought for a moment that psychology (as opposed to, say, anthropology) was

---

[111] The phenomenon is general, in that, e.g., sniffer dogs also respond to cues from their supervisors: Lisa Lit, Julie B. Schweitzer, Anita M. Oberbauer, "Handler beliefs affect scent detection dog outcomes." *Animal Cognition*, 2011 May; 14(3); 387-394.

[112] Roland Barthes, "The World of Wrestling." *Mythologies*. New York: Hill and Wang, 1972.

a real science — let alone psychoanalysis, which was more like a form of literary criticism, an attempt to read the history of the Ego as if it were a work of literature — though there seemed to be a strange sort of depth in it — maybe this was what Kant saw in Swedenborg, it had that element of spirits speaking to us from other dimensions, accessing the unconscious mind in the same way mediums accessed the spirit world in séances — I got a copy of *The Interpretation of Dreams* and began reading it the same way, annotating nearly every line and writing comments continuously as I went. Progress was not rapid, but it was cheap entertainment, and I pressed forward, the intrepid European explorer hacking his way through the hermeneutic jungle with machete and rifle in search of a lost city of gold (hopefully ruled over by Ursula Andress). After a few days I had read several chapters and trashed the best part of a notebook in this heroic quest to fathom the nature of the psyche.

Then I had a dream about fucking my mother. I tossed the notebook into the fireplace, took my paperback Freud to the used book store to trade it in for a couple of space operas, and decided to let someone else refute psychoanalysis.[113]

And went back to reading fast and sloppy.

---

[113] I gather Grünbaum also figured out why the patients of Freudian analysts — subject to far more efficient imprinting than mere readers — had Freudian dreams, and, in his own inimitable fashion, expanded the argument into an enormous book. But of course I never read it. It could only have provoked a relapse.

Compare Harold Bloom:

> Though a kind of Freudian for a few years……a proposed study of him called *Transference and Authority* was the one book I have never been able to finish. And I had to abandon an annual graduate course on Freud, because as the term neared its end, my verbal slips, the parapraxes of Freud's *Psychopathology of Everyday Life*, augmented, until the final class became involuntarily hilarious, since I suffered a kind of misspeaking in tongues.[114]

---

[114] *Genius*, p. 181. New York: Warner, 2002.

Da(i)sein

Parenthetically: though you take it for granted that an artificial intelligence would have a mind and a soul,[115] it seems obvious it would not have a psyche; at least nothing a human would recognize as such. In that respect it would indeed be alien and uncanny.

(Ironically this doesn't mean a robot couldn't be an effective psychotherapist, as Weizenbaum[116] demonstrated long since with ELIZA, a program based on a few simple tricks which required neither mind nor soul. — What this says about psychotherapy I'm not sure. But it can't be good.)

---

[115] In the sense of Aristotle, *De Anima*. We'll get to that.

[116] See Joseph Weizenbaum, *Computer Power and Human Reason* [1976]. — What really amazed him, he said, was that when he had his secretary try talking to ELIZA, even though she had witnessed him writing it and knew it was just a sort of magician's trick, after a couple of minutes interacting with the program she asked him to leave the room to respect her privacy.

Mimicry

— is actually problematic. — As well it ought to be, this is almost the problem of induction.[117]

What did the neoclassicists mean when they said Art put the mirror up to Nature? (What gets reflected? How? — The question of what constitutes *representation*.)

The protagonist of William Gaddis's *The Recognitions* doesn't copy artist's paintings, but artist's styles. How can that be possible?

Bloom quotes a discussion[118] of how Boswell "impregnated himself with the Johnsonian ether"; a paradigmatic example of learning a style in the sense of Gaddis. — For the distinction here between remembering precisely what Johnson had said and reconstructing it from scribbled notes by exerting an uncanny ability to mimic Johnson essentially vanishes.

We learn words and phrases by rote but this isn't all of learning language; we can absorb grammars and generate new ones, internalize theories, fulfill much more elaborately articulated expectations — can interpret much subtler signals, and can respond in more sophisticated

---

[117] In the sense of what it is that allows you to continue a sequence, almost exactly.

[118] By Frederick Pottle: "Does Boswell, then, report Johnson's conversation verbatim? In particular sentences and in some brief passages of an epigrammatic cast, yes. In general, no. The crucial words, the words that impart the peculiar Johnsonian quality, are indeed *ipsissima verba*. Impregnated with the Johnsonian ether, Boswell was able confidently to recall a considerable body of characteristic diction. Words entail sense; and when elements of the remembered diction were in balance or antithesis, recollection of words and sense would almost automatically give 'authentic' sentence structure. But in the main Boswell counted on … an understanding, grown intuitive, of Johnson's habits of composition … to construct epitomizing sentences in which the *ipsissima verba* would be at home." [*Genius,* p. 170.]

fashion. The process is so opaque that Chomsky resurrected the theory of innate ideas to explain it. (His version of the theory of *Meno*.)

We don't simply answer by picking multiple choice.

The brothel scene in *Amadeus*, where Mozart improvises in the style of a series of composers — Salieri last and most ludicrously.

The reductio ad absurdum, Tony Hendra on Douglas Kenney: "The very first time I met Doug, he was talking like William Makepeace Thackeray. ...I mean, he was actually improvising his prose...and it was Thackeray, it wasn't Dickens...and about 30, 45 seconds after that he was demonstrating that he could put his entire fist into his mouth... ."[119]

Not so funny, David Halberstam on Robert McNamara:

> One particular visit seemed to sum it up: McNamara looking for the war to fit his criteria, his definitions. He went to Danang in 1965 to check on the Marine progress there. A Marine colonel in I Corps had a sand table showing the terrain and patiently gave the briefing: friendly situation, enemy situation, main problem. McNamara watched it, not really taking it in, his hands folded, frowning a little, finally interrupting. "Now, let me see," McNamara said, "if I have it right, this is your situation," and then he spouted his own version, all in numbers and statistics. The colonel, who was very bright, read him immediately like a man breaking a code, and without changing stride, went on with the briefing, simply switching his terms, quantifying everything, giving everything in numbers and percentages, percentages up, percentages down, so blatant a performance that it was like a satire. Jack Raymond of the *New York Times* began to laugh and had to leave the tent. Later that day Raymond went up to

---

119 *Drunk, Stoned, Brilliant, Dead.* [Douglas Tirola, 2015.]

McNamara and commented on how tough the situation was up in Danang, but McNamara wasn't interested in the Vietcong, he wanted to talk about that colonel, he liked him, that colonel had caught his eye. "That colonel is one of the finest officers I've ever met," he said.[120]

— Almost like — dare we say it? — the kind of patient one might write a paper about — ? —

---

[120] *The Best and the Brightest*, Chapter Thirteen. Here, alas, we see the origin of the infamous body counts.

Peirce distinguished induction from abduction; a more provocative name for it, though he seems to have meant something more like Popperian hypothesis. — What you wonder about is more like black magic, however, almost a species of Promethean theft: it is as if the idea has been stolen from the mind of God. — Almost like a form of telepathy, you might speak of "channeling" e.g., the Zeitgeist, perhaps, or speaking as the voice of the people.

Can you talk seriously of reading the mind of nature? — Here again though usually you seem to read it from the face, it sometimes seems that some deeper connection has been established.

I did keep track of my dreams for a while in high school, but this had nothing to do with Freud; rather I was intrigued with the speculations of J.W. Dunne, who thought dreams incorporated experiences not only from the past but also the future.[121] (I was not yet enough of a wiseass as to say this meant they employed the advanced as well as the retarded Green's function,[122] but that was the idea he was straining toward.)

So I wrote them all down in a diary for a month or two, to see if I could find evidence of precognition. Apart from a premonition about the outcome of the Rose Bowl,[123] I could find none. And the dreams turned out to be very dull, at least the ones I could remember. — Moreover contrary to accepted wisdom it did not get easier to recall them with practice, and I suspect those who claim that it does have learned instead not to notice they've just started making shit up.

{...}

Nonetheless occasionally I have had good ones. I recall one in particular, when I was seven or eight, in which I flew up and down through history in a time machine that looked like the *Spirit of St. Louis*;[124] it featured stampeding dinosaurs and a supernova

---

[121] J.W. Dunne, *An Experiment With Time*. London: A. & C. Black, Ltd., 1929.

[122] Wheeler, John Archibald, and Richard Feynman, "Interaction with the Absorber as the Mechanism of Radiation." *Reviews of Modern Physics*, Vol. 17 Number 2, April-July 1945.

[123] As has been pointed out by every skeptic since the world began, premonitions recalled after the fact are meaningless. — Conversation with a friend after the Challenger disaster: "I realized that I had had a premonition before the launch that the thing would blow up. And then I realized I had had such premonitions before every single manned launch going back to Alan Shepard." — He: "The exact same sequence of thoughts went through my head."

[124] Lindbergh's memoir was one of my favorite books as a kid. He and Saint-Exupéry epitomized the romance of aviation that NASA later did its best to destroy.

accompanied by a choral ode, and somehow affected me profoundly; at least I've never forgotten it.

{...}

In general, however, the most interesting part of dreaming is waking up and trying to continue the dream into conscious life. — There is a strange character to these meditations. They are like dreams themselves; not conscious analyses, but variations on a theme. — In this cross-conscious state the mind moves sideways at extraordinary speed, like motion on a frictionless surface, a sort of associational superfluidity, with the result that even though the dreams may be pedestrian the interpretations can be remarkable. I have thought of things this way that would never have occurred to me otherwise.

In fact I half suspect that remembered dreams are basically constructed in this state, partial rationalizations of mental states that don't translate into the representations of conscious life.[125] — I might go on at length here about an analogy with the relationship between language and music, but Nietzsche did it all better in *The Birth of Tragedy*.

{...}

On one occasion I was (I guess) engaged in some kind of flashback to my school days, which involved a chase through the steam tunnels in the style of the finale of *The Third Man*. But then my companion and I tried climbing out to escape through some kind of narrow grating: he got away, but I got my head caught in a tight place. — Panicked. —

---

[125] All this would be different if there were some way of monitoring the brain of a sleeping subject and turning that into words and images; as, e.g., in the scene in *Prometheus* [Ridley Scott, 2012] in which the android David/Michael Fassbender watches the dreams of the sleeping Noomi Rapace with a sort of hologram projector. If brain states could be recorded and played back, obviously, in some sense they could be remembered, and thus be accessible to consciousness. — There are recent experiments which seem to indicate this may indeed be possible; if so the terms of the discussion will be transformed completely.

And woke immediately, in full analytical mode: was this really birth trauma? It seemed too cute to be believable. — Then began to drift again, and imagined a cartoon: the moment of birth; doctors, taken aback; no head has yet emerged, but, first, a hat on the end of a stick. — I told this to an artist friend later; he was enormously amused, and at once began to sketch the idea.

Identity

Hume[126] extended skepticism to the existence of the soul or Ego: there is no direct perception of self, he maintained, only of its components, impressions and ideas.

I would say on the one hand yes, this is true, in that self-consciousness in real and limited human life is fragmentary; but on the other hand no, the perception of self is direct and unambiguous, in the same way that you don't perceive the *whole* brick when you look at it, you turn it over and look at different sides, assess its texture, its color, heft it to get a sense of its weight, etc., before (in imitation of Dr. Johnson)[127] you throw it through Hume's window to get his attention.

One can analyze the glimpses down to their individual fragments and there isn't much brick left either; nor window, nor Hume, nor *Treatise*, nor sentences nor words nor sense to be found in them — nor impressions nor ideas either. — It is altogether too easy, with analysis, to reduce the world to undifferentiated dust.

But — granting his hypotheses — ideas are impressions of impressions; still there are ideas of ideas; of ideas; of ideas; of — Hume only carries this to first order, in other words, and there is a hierarchy here of indefinite extent.

(Ulam quotes Banach: "Good mathematicians see analogies between theorems or theories, the very best ones see analogies between analogies.")

---

[126] See his *Treatise*, Book I, Part IV, Section VI, Of Personal Identity.

[127] It is probably superfluous to quote the relevant passage here;. see Boswell's *Life of Johnson* 6 August 1763.

Is it just intension and extension again? — The set versus the predicate that defines it.

(Something about the visible tree, and the root system that is not perceived. — Well, we can multiply metaphors at will….)

Factorization

Adjoint to the problem of identity is a related question which has always bothered me, what might be called the problem of factorization.

That is, we have on the one hand the totalitarian fantasy[128] of the master computer, the monster intellect which like Alpha-60[129] sits at a central location, receives all information, processes it, and issues orders; thus controlling a city, a nation, an empire — whatever. But on the other, the democratic hand, we may imagine a collective intelligence[130] that via some kind of telepathic mechanism shares the thoughts of many individuals, perhaps an entire race — brains as cells in a meta-organism — and multiplies their capacities to godlike proportion.

But there are reasons that the human body does not consist of one big cell, or one undifferentiated colony of cells. The logic of functional organization entails a kind of factorization, a division of labor.

If one tried to build a computer[131] the size of the solar system, it would take several hours at the speed of light to get a signal from a neuron[132] on one side to a neuron on the other; you also know that neurons can't be scaled down indefinitely, that they must have some finite minimal

---

[128] One which for obvious reasons has always been particularly dear to the military mind.

[129] It is, of course, ironic that Godard should have invented this parable of rebellion against totalitarian control, and then embraced Mao, who personified it.

[130] Cf., e.g., Arthur Clarke's *Childhood's End*, or its original, Olaf Stapledon.

[131] We phrase this in the terms of computer architecture because we understand that a little better. Obviously.

[132] There have to be fundamental elements of some kind, why not call them this.

size. So physical limitations suggest that partitioning computations into pieces which can be performed locally must be necessary. Also, the inherent limitations[133] on the efficiency of a single bandwidth-limited central processing unit that performs all computations serially vis-a-vis a parallel architecture that distributes them are also well known, and this argument leads to the same conclusion. So the picture of a distributed network with many centers of computation emerges naturally by more than one line of reasoning.

Godfrey-Smith notes that, in distinction to the design of the chordate nervous system, which channels signals along the spinal cord to a central brain, "… much of a cephalopod's nervous system is not found within the brain at all, but spread throughout the body. In an octopus, the majority of neurons are in the arms themselves—nearly twice as many as in the central brain. The arms have their own sensors and controllers. They have not only the sense of touch, but also the capacity to sense chemicals—to smell, or taste. Each sucker on an octopus's arm may have 10,000 neurons to handle taste and touch. Even an arm that has been surgically removed can perform various basic motions, like reaching and grasping."[134] — Exactly.

So there is a tricky optimization problem here, one whose solution is problem-specific, involving the best balance between the number of cells and their individual capacities. In the original design of his Connection Machine,[135] Danny Hillis imagined the individual elements would be minimal, with very little local memory; in subsequent realizations of his ideas economics dictated wiring together existing computers with rather complex architectures instead, though

---

[133] See John Backus….

[134] Peter Godfrey-Smith, *Other Minds,* Chapter 3. [New York: Farrar, Straus and Giroux, 2016.]

[135] A radically elegant departure in computer architecture, based on the idea of wiring 2^n processors together in parallel as nodes in an n-dimensional Boolean hypercube; described in his thesis: W. Daniel Hillis, *The Connection Machine.* [Cambridge: MIT Press, 1986.]

the trend has been to reduce their size. The design of the neuron, in other words, is still evolving.

So if one really did attempt to build some kind of maximal intelligence, it would look more like a society than an individual mastermind, and it isn't obvious where the tipping-point lies. What you have to guess is that there is some kind of theoretical upper limit to (individual) intelligence, that manifests itself as a threshold of stability, and beyond this if you try to add more processors the whole will begin to behave like a group of individuals communicating — indeed probably arguing — with one another; not like an undivided unity. — That there is a logical necessity that turns Alpha-60 into the City Council, in other words.

And clearly to some extent though there is a unifying identity this has already happened in the brain as we see it, though we don't understand the interaction of the component parts.

Here also I should mention something that I think of as the principle of the cell wall: that parallel computation can only be efficient if possibilities can be evaluated independently, without interference; you must be able to do one thing and work it out to a conclusion. The evolution of life required the separation of genetic material into separate individuals to ensure that experiments would be independent.

— However. — Identity and factorization *are* adjoint, the dialectic of analysis and synthesis is universal (or as universal as anything gets), and though treating an arbitrarily large collection of individual neurons as a single intelligence may not be valid, this doesn't rule out the possibility of hierarchical organization, with collectives at one level forming individuals at the next;[136] this is the idea of the monadology, after all, and this too corresponds to some fundamental principle in nature. — So we haven't really eliminated the possibility of Gaia.

------

[136] It's also possible that "levels" may parse differently depending on point of view. But one conceptual morass at a time.

We're still articulating the reasons that we don't understand how it could work.

Of course given any mathematical model of consciousness the first question is how to generalize it. We can easily imagine more primitive modes of consciousness — I would certainly say that dogs have minds and souls, for instance, but doubt their memories are quite so acute as ours, and suspect therefore that their ability to think *about* what they are thinking is considerably less[137] — there is a continuum, clearly, but there may be some kind of abrupt transition when complexity attains some critical value — but can we imagine higher forms? by which we don't refer to some form of cosmic insight (the drugs trivialized all that, I'm afraid) but some extension of the idea —

— Perhaps to higher dimension? suppose we picture a two-dimensional time, and try the string-theoretical trick of extending world lines to world sheets: one might imagine a manifold of sequences of lived experience, each taken as a state, and a — continuous? — progression through them.

But it isn't obvious this amounts to a real extension: taking an entire world line as a state is simply adopting a more complicated definition of "state", Turing machines with multiple tapes can be linearized, functions of more than one variable can be reduced to functions of a single variable by the trick of "currying", etc.

— A better idea is to relax the requirement of linear memory. (To split the causal chain of consciousness.) Nothing in principle prevents the extension of the reference/containment idea to more general partial orders, and though the notion of a tree of possibilities which branches at every moment is both overfamiliar and rather nebulous, one might, for instance, suppose the possibility of cloning a personality completely, say in a science-fiction scenario in which (I will get to this later) you transmit a copy of yourself to Alpha Centauri, the two

---

[137] They are however obviously conscious, since they can be unconscious, sleep, and dream.

versions of yourself have parallel series of adventures, and then the copy is transmitted back and reintegrated with the original.[138] — It is difficult to believe that the human brain as presently constituted could absorb this shock, but a form of consciousness that could is a very natural idea of the superhuman.

(Perhaps it is worth noting that though there's no logical difficulty in reconciling two separate threads of past experience there wouldn't be any subjective sense of whether an event on one timeline occurred before or after an event on another; no matter that some external clock might be able to decide the proposition objectively. — Threads that are not recombined, on the other hand, are simply separate personalities; rather odd to contemplate but then in some sense every living creature represents a branch on the tree of life, and, etc.)

Strange but true there's even the hint of such speculation in the classical literature: there is an old tradition, referred to in the Second Part of Goethe's *Faust*, that the real Helen went to Egypt and it was a phantom double who went to Troy. When asked about this she says [8880]

> This is a superstition of dark-tangled sense!
> Which of them am I? Even now I do not know.[139]

Though in principle nothing prevents her from having been both.

— The extent to which multiple simultaneous *conscious* threads of execution may be sustained by real existing humans, on the other hand, is not obvious. It can be verified with simple psychological tests that the vast majority of people (in excess of ninety percent) are incapable of true multitasking, for instance; a small minority, on the

---

[138] It would be more natural, of course, to imagine a swap, in which both copies continue to exist after integrating their separate threads of experience.

[139] David Luke translation.

other hand, seem to be much more capable. — It is also said that some gifted individuals can attend to several conversations at once, or read and converse at the same time — Caesar, for instance, is supposed to have had this ability — and if I ask myself whether when Bach played the organ he was doing four intricate things at once, or just one really complicated thing that involved both hands and both feet, I don't know what to answer. — But the *possibility* certainly exists.

<center>{…}</center>

Put another way: the traditional conception of the immortality of the soul is one of an extended life: this world, *and then* the next; the linear continuation of personal identity by the accretion of memories, an uninterrupted thread. But this seems strangely limited. It is more naturalto imagine higher forms of consciousness, e.g., one that began with the original thread, the life-line with its beginning and end, and extended it in another dimension, into a sort of ribbon; this might be a kind of extension into parallel worlds, but there could be other interpretations as well. — But more or less by definition this is beyond human comprehension.

Regarding it as Nietzsche did contemplating Goethe, a life taken as a whole might be regarded as a work of art — though if so one never completed but abandoned — as a kind of moment of apprehension in a larger consciousness, say; then one can ask, by analogy, what might *follow* that, and it would look more like a variation, an imitation, perhaps, or an annotated commentary, or an answer to the question "how else might it have been done?"

It probably isn't an accident that all this occurs to me while listening to Gould play the Eroica Variations.

Memory in childhood

I had noticed by the time that I was seven or eight[140] that my earliest memories were already fossilized; no longer the living originals, but copies, the imprints those had left in stone.[141]

I asked my mother about a few of these, because I had no context for them other than they seemed to come before any others. — One was of very bright sunlight on white sand, with a blue expanse of water beyond it. She said that was Lake Huron, at the tip of the thumb, when I was two years old. — Another was of lunch at a kitchen table, and a radio playing *The Romance of Helen Trent*. She said that was the first apartment they'd had after they got married, and I may have been less than a year old. — I think this was an exaggeration, but I also retained[142] a subliminal impression of a brick building, which must have preceded the first of many suburban houses. — I also remembered my playpen, and the room in which it sat in in the first of those houses, though not what she claimed were my effortless escapes from it. — In any case I have seen home movies of my third birthday, and I remember nothing of it.

---

[140] It had to be about that time, since the analogy would not have occurred to me before I had read George Gamov's *Biography of the Earth*, which explained the formation of fossils.

[141] As for how this metaphor would have occurred to me — well: I had already read George Gamov's *Biography of the Earth*.

[142] The memory emerged in an unexpected fashion, when I found myself, forty-odd years later, by chance navigating an apartment building in Married Student Housing which dated from the same postwar era, and was undoubtedly constructed in the same slapdash-but-curiously-permanent style that characterized, e.g., the Quonset huts which had for decades performed a similar function. — A strange feeling of familiarity came over me; I felt, for reasons I could not identify, that I was at home. — Only later did an explanation for this otherwise uncanny sensation occur to me.

I also remembered standing in our attic bedroom, looking at my sister in her crib, and forming the perfectly subvocalized sentences: "I am three. That is my sister. She is two." — There is a charming simplicity in this. It is almost worthy of a British philosopher.

But the continuous thread of conscious life and memory only began around the time that I was four. My impression, which can only be slightly exaggerated, is that I understood what had happened almost immediately. With this commenced a process of self-examination which continues to this day.

Before that self-awareness had been patchy, and interrupted by blackouts and lapses. — I recall vividly, for instance — or did, it is hard to get the tenses right — one day when I was left alone to take a bath. When my father came back to check he found me playing contentedly with my toys at one end of the tub; but pointed out, as sternly as he could manage, that there was something at the other. I perceived a discoloration in the water, and realized this was the signature of the presence of a large turd, which had somehow materialized on the bottom. Clearly I had to be responsible, but I remembered nothing, and said, truthfully, that I hadn't been aware of what I was doing. He didn't understand me — in fact this was the reason I remembered the incident so clearly, because it puzzled me that he didn't understand — and asked me, in apparent seriousness, whether I thought the Lone Ranger would have done something like that. — I stared at him blankly, because it seemed like such a strange question. — As indeed it still does.

— Well. — What can I tell you. — Marcel Proust I am not.

The unconscious

The inherent necessity of shutting the mechanism down for regularly scheduled maintenance became apparent to me during a (thankfully brief) period when I abused the use of amphetamines, and became habituated to staying up three and four days at a time. What I discovered was that dream interludes began inserting themselves into waking life whether you wanted them or not. — A canonical instance was an occasion when I was (so I thought) trying to follow an argument in Whittaker/Watson, and realized when I turned one page that the cavalry had just arrived in the nick of time on the one preceding. — Wait a minute, said I….

At any rate dreams, though fascinating, are not some kind of unique window into "the unconscious", which is essentially a kind of triviality.

This was less obvious before the invention of digital clocks. Now it happens constantly that I glance at the readout in the upper right hand corner of the screen and notice — what a coincidence — that the time is something like 3:14, or 2:34 (you look up a moment later and it is 2:46), or 3:43, or 12:16, or (this only began after 2001) 9:11.

My laptop has a sexed-up version of a Unix utility which in the MacOS is called the Activity Monitor; it list all the processes resident on the CPU at any moment, and numbers the subprocesses subordinate to each. This fluctuates continuously, but typically there are around 180-190 processes and 900-1000 threads of execution dependent on them.[143]

Clearly something similar holds for the human brain. Some thread is constantly checking the clock for me, and promotes the time to

---

[143] For the new generation of laptops, double the numbers.

attention when there's something significant about it. And whatever it is recognizes arithmetic sequences, and knows Beethoven's birthday, and the cube of 7, and a great deal besides. — Other illustrations may be multiplied at will, but consider, e.g., the common experience[144] of glancing at a page without seeming to see it, having the feeling there was something funny about it, and then realizing on closer inspection that it contains the name of someone you knew in elementary school.

So this answers the question "Where do thoughts come from?" — It is just as with Santa Claus and his toys: an unseen army of little elves labor to produce them.[145]

---

[144] At least, this sort of thing happens to me all the time.

[145] Henri Poincaré had a lot to say about these little elves; see below.

The world's longest "Duh"

The premise of *Blink* [Michael Apted, 1993] is that a woman (Madeleine Stowe) who has been blind since childhood has had her sight restored by an operation; her capacity for processing visual input, however — the software, it is intimated, not the hardware — has atrophied from disuse, and until it is reconstituted she is unable to interpret what she is seeing; images pop into her head at unpredictable intervals after the stimulus that prompted them, often days later. — Of course this is immediately complicated by her being witness to a murder, but — never mind — this really does happen, though I have usually observed it with auditory rather than visual inputs.

In particular it is a frequent occurrence, for someone surrounded by speakers of an unfamiliar language, to hear some word or phrase spoken and be unable to recognize it; until some pattern-recognition engine which has been chugging away in the interim announces its result, often minutes later.

The record, which is indicative either of the unusual strength of my unconscious muscles or of just how stupid I really am, is held by the answer to a question I put to my hosts on the first New Year's Eve I spent in Argentina: we were eating something, I asked what it was, they made some incomprehensible noise[146] in reply, I pawed about aimlessly in the dictionary, I gave up. — "Looks like breaded veal," I thought. — Six months later while browsing in the *carnecería* I saw a sign saying "Milanesa" and knew instantly what I had heard. — And

---

[146] There is no harm in belaboring this elementary point of epistemology: you have no idea what you have heard if you don't know what to listen for; pattern recognition is a matter of matching pre-existing templates — not to concede Chomsky's point, but he was certainly onto something — and when they don't exist, the result is chaos. — The same is true of vision, of course, and even — for reasons inherent to the logic of perception — machine vision; which goes a long way toward explaining the UFO phenomenon, even when it appears to be confirmed by electronic sensors and computer software.

glancing downward, saw a familiar set of cutlets on display in the cooler below it.

Of course there are less obvious cases — waking in the middle of the night, checking the digital clock and noting that it is 4:20 on the morning of 4/20, for instance — is there a clock in my head? maybe, it sometimes seems that way — weirdly inappropriate to apply such mathematical precision to a stoner holiday though; is this the work of the Inner Prankster? — and there is disturbing evidence of a kind of slippery slope: I could laugh it off, for instance, when knowing from previous experience that I would be spending the entire day waiting in office antechambers for someone to hand me the next form to be filled out in the paper chase I carried *Ulysses* with me to my first day of work on a government summer job, read half of it with a curiously obsessive focus, and only realized later that the day in question was, indeed, June 16th — Bloomsday — but it was more perplexing when I was seized one evening by the impulse to read a book about Boethius, and discovered halfway through it that this was his feast day (October 23) — how could I have known that? I couldn't recall ever having read anything about him; Russell mentions him with admiration but never refers to the Church calendar. So where did that come from? How did that fact enter my head? — Jung did have something right. Sometimes it looks as though external reality is also wired into the unconscious mind. That dreams connect us by some subterranean channel to *anima mundi*.

Synchronicity

Jung filled a book[147] with anecdotes illustrating the idea of synchronicity, but one will suffice here to stand for all, because [a] it's a great story, and [b] the stories all sound alike anyway, more or less for the same reason that even though The Hero may have A Thousand Faces, they all look like Luke Skywalker.

It is from Izaak Walton's life of John Donne:

> At this time of Mr. Donne's and his wife's living in Sir Robert's house, the Lord Hay was, by King James, sent upon a glorious embassy to the then French king, Henry the Fourth; and Sir Robert put on a sudden resolution to accompany him to the French court, and to be present at his audience there. And Sir Robert put on a sudden resolution to solicit Mr. Donne to be his companion in that journey. And this desire was suddenly made known to his wife, who was then with child, and otherwise under so dangerous a habit of body, as to her health, that she professed an unwillingness to allow him any absence from her; saying, "Her divining soul boded her some ill in his absence"; and therefore desired him not to leave her. This made Mr. Donne lay aside all thoughts of the journey, and really to resolve against it. But Sir Robert became restless in his persuasions for it, and Mr. Donne was so generous as to think he had sold his liberty, when he received so many charitable kindnesses from him; and told his wife so, who did therefore, with an unwilling-willingness[148], give a faint consent to the journey, which was proposed to be but for

---

[147] Jung, C.G. "Synchronicity: An Acausal Connecting Principle." *The Structure and Dynamics of the Psyche*, pp. 417-531. Transl. R.F.C. Hull. London: Routledge, 1960.

[148] Somebody's been reading Sir Philip Sidney.

two months; for about that time they determined their return. Within a few days after this resolve, the Ambassador, Sir Robert, and Mr. Donne left London; and were the twelfth day got all safe to Paris. Two days after their arrival there, Mr. Donne was left alone in that room in which Sir Robert, and he, and some other friends had dined together. To this place Sir Robert returned within half-an-hour; and as he left, so he found, Mr. Donne alone, but in such an ecstasy, and so altered as to his looks, as amazed Sir Robert to behold him; insomuch that he earnestly desired Mr. Donne to declare what had befallen him in the short time of his absence. To which Mr. Donne was not able to make a present answer, but after a long and perplexed pause, did at last say, "I have seen a dreadful vision since I saw you: I have seen my dear wife pass twice by me through this room, with her hair hanging about her shoulders, and a dead child in her arms; this I have seen since I saw you." To which Sir Robert replied, "Sure, sir, you have slept since I saw you; and this is the result of some melancholy dream, which I desire you to forget, for you are now awake." To which Mr. Donne's reply was, "I cannot be surer that I now live than that I have not slept since I saw you; and am as sure that at her second appearing she stopped and looked me in the face, and vanished." Rest and sleep had not altered Mr. Donne's opinion the next day, for he then affirmed this opinion with a more deliberate, and so confirmed a confidence, that he inclined Sir Robert to a faint belief that the vision was true. — It is truly said that desire and doubt have no rest, and it proved so with Sir Robert; for he immediately sent a servant to Drewry House, with a charge to hasten back, and bring him word whether Mrs. Donne we alive; and, if alive, in what condition she was as to her health. The twelfth day, the messenger returned with this account: That he found and left Mrs. Donne very sad, and sick in her bed; and that, after a long and dangerous labour, she had been delivered of a dead child. And, upon examination, the abortion proved to be the same day, and about the very hour, that Mr. Donne affirmed he saw her pass by

him in his chamber.

> This is a relation that will beget some wonder, and it well may; for most of our world are at present possessed with an opinion that visions and miracles are ceased. And, though it is most certain that two lutes being both strung and tuned to an equal pitch, and then one played upon, the other, that is not touched, being laid upon a table at a fit distance, will — like an echo to a trumpet — warble a faint audible harmony in answer to the same tune; yet many will not believe there is any such thing as a sympathy of souls...

So here we have a vivid portrait of the intimate spiritual connection between Donne and his wife, a striking anecdote — admittedly so perfect that we must assume many layers of embellishment have been interpolated between the story Walton tells and whatever really happened — and the beautiful figure of the attuned lutes — about which it is interesting, considering that this is, after all, an exercise in hagiography, and Walton a fairly pious guy who idolizes the Dean of St. Paul's and mentions young Jack Donne the notorious rake and wit not at all, that he doesn't simply say "God told Donne what had happened to his wife," but suggests a sort of explanatory hypothesis; that he seems to have an intuitive belief that this episode lies within the scope of the philosophy of nature, and that one needn't employ theological cheats like dividing by infinity to set everything equal to zero.

Moreover the idea of sympathetic vibration seems perfect because it not only suggests the kind of resonant excitation that intuitively seems to be involved,[149] but also the host of complications that may beset the

---

[149] Compare also Kepler's theory of astrological influence, that the stars worked their effects upon humanity by a kind of sympathetic resonance with the Music of the Spheres. — The Pythagorean discovery in which mathematical physics originated was that the tone emitted by a string depended on the tension in it; the intuitive belief that the harmony of Nature was intrinsically musical lingered long, long afterward.

connection if the resonant cavities are not perfectly attuned; connotes also the suggestion that the cavities need not be identical in shape and size, but need only agree in a characteristic frequency; and intimates a tuning process may have made Donne and his wife consonant with one another,[150] that people who live together intimately may come to think the same thoughts — incline toward one another; as in the famous image of the compass legs in Donne's own "A Valediction: forbidding mourning".

— Nonetheless: this is mental radio again, and begs the usual question of the constitution of the luminiferous ether. —

{…}

Leibniz in a letter to Basnage de Beauval of January 1696[151] suggests three possible mechanisms to explain how a pair of watches might come to be in agreement. The first, curiously enough, is exactly that of resonant excitation, for which he cites an experiment of Huygens in which two pendulums suspended from the same wooden frame ("the collective unconscious"?) adapted to a mutual agreement in period. — The second is occasional interference by an external observer (something like Newton's assuming God would occasionally have to readjust the solar system to keep the planets in their orbits); the third, the pre-established harmony. — The question, really, is why the third possibility doesn't seem like an *explanation* at all. Since this is certainly the most exact description of the phenomenon in question.

---

[150] Thus the well-known phenomenon of the alignment of the menstrual cycles of women who live together; or for that matter (see Kepler) the alignment of the menstrual cycle generally with the lunar cycle.

[151] Loemker pp. 459-460.

<center>{…}</center>

Presumably the idea of resonance was in the air. (And — yes — what does that mean?) The English Puritan John Flavel, for instance, says about the pre-established harmony

> The soul manifests its dear love and affection to the body, by its sympathy, and compassionate feeling of all its burdens: whatever touches the body, by way of injury, affects the soul also by way of sympathy. The soul and body are as strings of two musical instruments set exactly at one height; if one be touched, the other trembles. They laugh and cry, are sick and well together. This is a wonderful mystery….[152]

Though then just as easily might one soul excite another, or one soul excite another body. (It is odd it went unremarked that, under this argument, one might as easily walk around in someone else's skin as one's own.)

<center>{…}</center>

None of this works, at any rate, and so Jung in a way simply reverts to Leibniz: he tries to explain everything one might think of as paranormal as meaningful coincidence. — This also sounds like a cheat, but the point is subtler: obviously *anything* can be explained away as coincidence — the fact that the ball flies away from the bat when they make contact, the fact that the glass of the window breaks when the baseball strikes it, the fact that the kids' legs respond when their several wills urge them to take flight — it can all be pre-established harmony, if you are so perverse as to insist upon it — but (on the one hand) it is inconsistent to introduce the idea and then not carry it to its logical conclusion, and (on the other) there is the question whether one can grant the whole causal picture with which

---

[152] Where on earth did I find this?

we generally interpret our experience *and then ask*, can there be some additional order? has this left something out? — Can there be an alternative — no, let us be precise: an *additional* — framework of explanation, one based, say, on the internal logic of the psyche, complementary to and existing simultaneously with the physical framework with which we are familiar?

If one believed, for instance, that the relation of cause to effect is (as Hume almost literally put it) itself only coincidence, then why not admit the possibility of acausal connections that are not ordered in time? — And once you have done so it is remarkable how easy it is to explain almost any miracle: the same thought turning up simultaneously in two people's heads — even two people separated in time as well as space, thus visions of the past and future and the transmigration of souls; psychokinetic effects; reversals of entropy, which suffice to explain how the Nazarene (assisted by Maxwell's Demon) could walk on water or rise from the dead (the irreversibility of decay is not absolute, but statistical necessity); conjuring spirits, casting spells, making the rain fall or the crops to fail — luck, pure and simple, can account for almost anything. Since the visible order of the natural world is determined, by and large, by probabilities; most of what seems impossible is only very unlikely.[153] So meaningful coincidence can, should you so desire, explain all paranormal phenomena at once.

It should be made clear that this is not a simple null hypothesis — though it can easily become one, and Jung seemed to be susceptible to that temptation — rather, that we are taking the usual explanatory framework as given, and asking how we can add to it.

That is, if we return to Leibniz, his point of view — this is now long forgotten — was that *every* fact of the world must have some explanation; this was the meaning of the principle of sufficient reason.

---

[153] Admittedly the probabilities involved may be infinitesimal. But still.

But though this principle is still a powerful motivation for the formation of hypotheses, the evolution of mathematical physics, commencing (let me wave my hands) with Bernoulli's explanation of Boyle's gas law,[154] restricted the domain of reason drastically: the facts of the world were divided into the differential equations in which causal laws were expressed, and the initial conditions assumed for their solution; the latter were the domain of chance, the former of necessity.[155] — If you shot an arrow in the air, it fell to earth *exactly* where determined by initial position/angle/flex of bow;[156] but what *those* were could not be the concern of physics. — So the governing assumption in theoretical physics became that the time evolution of a system was rigorously determined, but the state in which that evolution commenced was ultimately the result of the (enormously complicated) initial state of the universe, and thus in the absence of the exercise of will and/or experimental design *as random as possible*. — Leibniz did not live to see the triumph of this point of view, but he would have loathed it; whether or not he would have admitted its pragmatic utility.

{…}

Kepler described not only the shapes and periods of the orbits of the planets, but in the *Mysterium Cosmographicum* had attempted to provide sufficient reason for their number and their relative distances from the Sun by associating them with a system of spheres inscribed and circumscribed about the five Platonic solids. — Newton already had abandoned this ambition, explaining their orbits with his laws of

---

[154] In *Hydrodynamica* [1738].

[155] This is in a way analogous to the way that atomism provided an elegant synthesis of the antithesis of Heraclitus and Parmenides: is everything is in flux, or is everything unchanging? The answer is both: the elements are invariant, but their state is variable. — Atoms are Being, motion is Becoming.

[156] Technically only *in vacuo*, but "I shot an arrow in the void/It fell to earth I know not where" doesn't scan.

motion and gravitation but waving his hands and declaring the arrangement of "the System of the World" to have been the work of God at the Creation, and therefore out of bounds. — Later the point of view was gradually advanced that the solar system we inhabit is only one out of zillions,[157] and the number and arrangement of the rocks that comprise it is the mere work of chance.

{…}

Jung made the connection with theoretical physics, though not through any direct understanding but through the mediation of Wolfgang Pauli, who began as a patient and ended up a kind of co-conspirator.[158] — Jung wanted someone to tell him the causal description of the world was incomplete; who better than one of the pioneers of quantum physics, a disciple of Bohr ready to apply the idea of complementarity to literally anything, even the most general conception of the physical picture of Nature.

{…}

Though oddly enough even Pauli didn't seem to understand the essential point, that even classical physics was drastically underdetermined, that there was an enormous freedom of choice left to exploit after the constraints of causality had been imposed; perhaps

---

[157] Here "zillion" has the technical meaning "a lot more than a trillion but probably less than a trillion trillion, at least in the visible universe."

[158] See in particular Letter [37P], 28 June 1949, in *Atom and Archetype: The Pauli/Jung Letters, 1932-1958*, edited by C.A. Meier, translated by David Roscoe. [Princeton: Princeton University Press, 2001.] — Pauli had gone through a series of personal crises in the late Twenties and early Thirties which led him to seek professional help from Jung. Whether this really improved his mental condition is as always open to question, but he rapidly advanced to the status of a collaborator, was a prolific provider of dreams (over 1300 by one count) which Jung made use of — most of the dreams cited in *Psychology and Alchemy* are Pauli's — and following up on his earlier comments was co-author of the volume *The Interpretation of Nature and the Psyche*, in which the essay on synchronicity was first published, contributing a lengthy essay on Kepler and alchemy.

too completely under the spell of Bohr, he insisted that the quantum-mechanical relaxation of strict determinism was the necessary prerequisite for synchronicity to be possible. — That, and he seems to have fallen in love with the Jungian practice of drawing quaternary diagrams purporting to illustrate the dialectical oppositions of the fundamental conceptual framework, viz.[159]

That is why I would now like to make the following *compromise proposal* for a quaternary schema as a basis for discussion; it avoids the opposing of time and place and perhaps combines the advantages of your schema and the one I drew up in 1948.

$$\text{Energy (conservation)}$$
$$\text{causality} \quad \underline{\quad\quad|\quad\quad} \quad \text{synchronicity}$$
$$\text{space-time continuum}$$

On p. 61, where you talk about the "triadic world picture," perhaps one could replace "by means of space, time, and causality" (8th line from the bottom) by "and the notion of causality." That would also fit in better with the term "three-principles doctrine," since *continuity* (*natura non facit saltus*) can certainly be viewed as a characteristic *principle* for the (classical) scientific age.

— where the vertical axis represents complementarity in the sense of Bohr, and the horizontal axis represents complementarity in the sense of bullshit. — All this is unnecessary and irrelevant.

(A more contemporary interpretation of the diagram would label the horizontal axis "entanglement", and one must suspect Pauli had some intuition of the possibilities quantum mechanics had opened in this regard which led him and Jordan, among others, to take a second look at occult phenomena. But the correlation of photon spins and the correlation of human fate with Destiny are radically different ideas, and it is difficult at first glance to see how they might be connected.)

---

[159] *Op cit.*, 45P.

<center>{…}</center>

It was Pauli who pointed out to Jung that Schopenhauer had addressed the question of meaningful coincidence long before either one of them, in an essay titled "Transcendent Speculation on the Apparent Design in the Fate of the Individual".[160] — This had also been the subject of an address Thomas Mann delivered in Vienna on the occasion of Freud's eightieth birthday in 1936,[161] in which he stated that "the most profound and mysterious point of contact between Freud's natural scientific world and Schopenhauer's philosophic one" was

> The pregnant and mysterious idea … that precisely as in a dream it is our own will that unconsciously appears as inexorable objective destiny, everything in it proceeding out of ourselves and each of us being the secret theatre-manager of our own dreams, so also in reality the great dream that a single essence, the will itself, dreams with us all, our fate, may be the product of our inmost selves, of our wills, and we are actually ourselves bringing about what seems to be happening to us.[162]
> .

Which probably made Freud blow a fuse; what could be more Jungian? — But Mann meant well. He was onto something.

------

[160] Arthur Schopenhauer, *Parerga and Paralipomena, Vol. 1* (pp. 199-224), transl. E.F.J. Payne. [Oxford: Clarendon Press, 1974. (Originally published 1850.)]

[161] "Freud and the Future", included in *Essays of Three Decades*, [New York: Alfred A. Knopf, 1947;] translated by H.T.Lowe-Porter. — This made a profound impression on me when I first read it (circa 1970), and of course I set out immediately to find the essay of Schopenhauer in question. Lacking access to a decent research library, I only succeeded decades later. — Another pothole on the Road to Being, as Heidegger might put it.

[162] Compare Bloom: "Freud, like some of the ancients, believed there were no accidents. Ethos is the *daimon*, your character is your fate, and everything that happens to you starts by being you."

{…}

Schopenhauer warns at the outset that the ideas he presents "might perhaps be termed a mere metaphysical fantasy," but suspects nonetheless that many have thought along the same lines; that there is a general belief in providence, a sense of being guided by fate, which he seats not in knowledge but in the will. He says that though "to attribute to pure evident chance a purpose or intention is an idea of unparalleled audacity" nonetheless "everyone has had at least once in his life a vivid conception of it."

That "everything that happens takes place with *strict necessity*" he regards as truth *a priori*, but — this reads very strangely — not in the sense of Newtonian mechanics, not in the sense of Laplace, rather by what he takes to be the incontrovertible evidence for precognition: "It is confirmed empirically and *a posteriori* by the fact, no longer in doubt, that magnetic somnambulists, persons gifted with second sight, and sometimes even the dreams of ordinary sleep directly and accurately predict future events." And as an illustration he cites the plot of *Oedipus Rex,* which he somehow assumes to be taken from life.

He calls this view *transcendent fatalism*, not demonstrable but something that one believes at least some of the time. — Contradicting Spengler in advance, he insists that "plan and totality are to be found not in world history, as professorial philosophy would have us believe, but in the life of the individual. In fact, nations exist only as *abstracts*; individuals are what is real." He remarks that people looking back on their lives often regard them as having unity of theme and purpose, like that of a literary composition, even though at the time they may have seemed random and meaningless.

In part this is a character-is-destiny argument, an appeal to the determining role of unconscious instinct in steering the course of one's

life — "the turtle in the sand, that is hatched out by the sun, at once goes straight to the water" — and he remarks that the conscious apprehension of the course of events is something like a perception in a distorting mirror, which when corrected will exhibit the underlying order and aesthetic arrangement.

Thus chance is surface appearance, necessity is the underlying reality: "…we think at every moment we are masters of our actions; but if we look back on the course of our lives … it looks as if a strange power has guided our steps." — He appeals to the wisdom of the ancients here, of course, but also to Goethe, e.g. *Egmont* "Man thinks he guides his life and directs himself; and his innermost being is drawn in accordance with his fate." — This however leaves us with two radically different interpretations of any given fact or event, the one physical and contingent, the other moral and metaphysical.

Again he emphasizes that this cannot be *demonstrated* — as Wittgenstein would later say, it is something that shows itself but cannot be said — and appeals again to literary examples; the implicit argument being that the design apparent in the plot of a work of Schiller is, like details of character and description, itself somehow drawn from life; that "a secret and inexplicable power guides all the turns and changes of our lives, indeed often contrary to the intention we had at the time." — That the events of one's life are like the events in a drama, in other words, whose author is presumably "what the ancients called fate … what they understood as the guiding genius of every individual"; in fact here he digresses briefly to discuss the origins of the idea of a guardian angel, and cites a passage in the *Republic* in which Plato imagines the soul before birth choosing its fate and the genius who will be its guardian.

This conception of a divided self is very close, obviously, to the later conception of the unconscious mind, and Schopenhauer clearly states that this guardian is not located without but within: "that occult power that guides even external influences can ultimately have its root only

in our own mysterious inner being; for indeed in the last resort the alpha and omega of all existence lie within us."

He notes a similarity to teleology in the natural world, gives some examples which he doesn't understand very well — but —finally states the crux: that causal chains may be related, that they may have become entangled in the past: "…nothing is *absolutely* accidental… definite causes lying high up in the causal chain have long ago necessarily determined that that something was bound to occur precisely at this moment and, therefore, simultaneously with something else."

<p style="text-align:center">{…}</p>

Schopenhauer forms the picture of a network formed from causal chains, ordered by time, like meridians from south to north, in which though "every event is the particular link in a chain of causes and effects which proceeds in the direction of time…there are numberless such chains side by side… not entirely foreign to one another and without any interconnection; on the contrary, they are intertwined… ."

Thus

> … many causes now operating simultaneously, each of which produces a different effect, have sprung from a common cause higher up and are, therefore, related to one another as great-grandchildren are to their great-grandfather. On the other hand, a particular effect occurring now often requires the coincidence of many different causes which, each as a link in its own chain, have come to us from the past. Accordingly, all those causal chains, that move in the direction of time, now form a large, common, much-interwoven net which with its whole breadth likewise moves forward in the direction of time and constitutes the course of the world. Now if we represent those individual causal chains by meridians that would lie in the direction of time,

then that which is simultaneous, and for this reason does not stand in direct causal connection, can be everywhere indicated by parallel circles. Now although all things situated under the same parallel circle do not directly depend on one another, they nevertheless stand indirectly in some connection, though remote, by virtue of the interlacing of the whole net or of the totality of all causes and effects that roll along in the direction of time. Their present co-existence is therefore necessary; and on this rests the accidental coincidence of all the conditions of an event that is necessary in a higher sense, the happening of that which fate has willed.

and here he appeals directly to dreams, and proposes that life is in some sense itself a dream: "indeed it is this analogy with the dream which enables us to observe … how the mysterious power, governing and controlling the external events that affect us … might yet have its root in the depths of our own unfathomable nature."

He notes that in dreams it can occur that the pursuit of some object of desire (he gives sexual examples) is constantly frustrated — and these frustrations are, of course, of our own invention — accordingly "just as everyone is the secret theatrical manager of his dreams, so too … that fate that controls the actual course of our lives ultimately comes in some way from the *will*."

It follows that that there is a division in the self between the empirical (conscious) and the hidden (unconscious) will; he points out a curious (and certainly heretical) passage from Scotus Erigena attributing the same internal division to the divinity; appeals to the *noumenon/phenomenon* distinction (that infallible catchall) to explain the duality of freedom/necessity, causality/teleology, contingency/moral necessity — here are the ancestors of the doctrine of complementarity — and, like Heraclitus, identifies character and destiny:

> … all the events in a man's life are connected in two

fundamentally different ways; first in the objective causal connection of the course of nature, secondly in a subjective connection that exists only in reference to the individual who experiences them. It is as subjective as his own dreams, yet in him their succession and content are likewise necessarily determined, but in the manner in which the succession of the scenes of a drama is determined by the plan of the poet. Now those two kinds of connection exist simultaneously and yet the same event, *as* a link in two quite different chains, exactly fits them both, in consequence whereof one man's fate is always in keeping with another's, and everyone is the hero of his own drama, but at the same time figures also in that of another. All this, of course, is something that surpasses all our powers of comprehension and can be conceived as possible only by virtue of the most marvelous *harmonia praestabilita.*

Which anticipates exactly the Jungian reduction to the unconscious. As the dream is purely a product of the individual, though it doesn't appear to be, so is the world as a whole.

And so much for the reality principle.

{…}

How Schopenhauer could have understood this question so clearly is a bit of a mystery, since otherwise he shows no evidence of mathematical intuition. But perhaps it can be explained by his musical sense, which — here is another mystery, why doesn't he say this — should have been telling him that the natural model for the relationship is not literary but musical: causality is melody, synchronicity is harmony — or, more precisely, the arrangement of parallel melodic lines in counterpoint.

These are, precisely, his meridians and latitudes.

So if we think of the world as a work of art, the artist we should picture creating it is not the playwright or novelist, but the composer.

And the form of the design, then, would be something like a musical score. With very many parts.

How then do the same thoughts appear in separate minds? — well: here is the theme in the strings, and here it is in variation, over there in the woodwinds. — How can Jung be listening to a patient tell him her dream about a golden scarab when suddenly a scarab beetle taps upon the window and flies into the room?[163] — Same difference.

You picture the author of Nature, then, as a sort of überBach; an artist coordinating a near-infinite number of individual melodic lines in a colossal fugue.

{…}

Mann had adapted Wagner's musical idea of the leitmotif in the composition of *The Magic Mountain*,[164] where it serves as a device which allows moments widely separated within the (enormous) narrative to (as it were) resonate, but more apropos is the musical

---

[163] "A young woman I was treating had, at a critical moment, a dream in which she was given a golden scarab. While she was telling me this dream I sat with my back to the closed window. Suddenly I heard a noise behind me, like a gentle tapping. I turned round and saw a flying insect knocking against the window-pane from outside. I opened the window and caught the creature in the air as it flew in. It was the nearest analogy to a golden scarab that one finds in our latitudes, a scarabaeid beetle, the common rose-chafer (*Cetonia aurata*), which contrary to its usual habits had evidently felt an urge to get into a dark room at this particular moment. I must admit that nothing like it ever happened to me before or since, and that the dream of the patient has remained unique in my experience." [p. 22] — A longer version of the story appears in an additional lecture included as an appendix.

[164] See the brief essay "The Making of *The Magic Mountain*" appended to the English translation of H.T. Lowe-Porter [New York: Alfred A. Knopf, 1965], which first appeared in *The Atlantic*, January 1953.

theory[165] of the protagonist of *Doctor Faustus*, the composer Adrian Leverkühn; who has (or thinks he has) sold his soul to the Devil to gain the creative power of a god — a Demiurge, whose compositions are as intricate as worlds, and employ radically original laws of form.

He is intrigued by the Pythagorean aspect of music, its relations with mathematics and astronomy — "Reason and magic," he says, "may meet and become one in that which one calls wisdom, initiation; in belief in the stars, in numbers …"[166] — incessantly theorizes about the nature of music and its structure — and, former student of theology that he is, the inherent blasphemy of the act of artistic creation: "*Imitatio Dei* — I am surprised it is not forbidden."

The ideal toward which he strives is that of the completely determined composition: one in which every element is related to every other element; one in which every element, like one of Leibniz's monads, reflects the musical universe it inhabits so that the whole may be inferred from any part. — "… I will tell you what I understand by 'strict style'," he says. "I mean the complete integration of all musical dimensions, their neutrality towards each other due to complete organization."

He explains to the narrator that he first approached this ideal in a song which

> "… is entirely derived from a fundamental figure, a series of interchangeable intervals, the five notes B, E, A, E, E-flat, and the horizontal melody and the vertical harmony are determined and controlled by it … . It is like a word, a key word, stamped on everything in the song, which it would like to determine entirely. But it is too short a word and in itself not flexible enough. The

---

[165] These disquisitions were largely subcontracted to Adorno, and stole so many of Schoenberg's ideas that he threatened a lawsuit. See E. Randol Schoenberg, ed., *The Doctor Faustus Dossier*. [Oakland: University of California Press, 2018.]

[166] The translations are those of H.T. Lowe-Porter. [New York: Alfred A. Knopf, 1948.]

tonal space it affords is too limited. One would have to go on from here and make larger words out of the twelve letters, as it were, of the tempered semitone alphabet. Words of twelve letters, certain combinations and interrelations of the twelve semitones, series of notes from which a piece and all the movements of a work must strictly derive. Every note of the whole composition, both melody and harmony, would have to show its relation to this fixed fundamental series. Not one might recur until the other notes have sounded. Not one might appear which did not fulfill its function in the whole structure. There would no longer be a free note. That is what I would call 'strict composition'. … every note, without exception, has significance and function according to its place in the basic series or its derivatives. That would guarantee what I call the indifference to harmony and melody."

"A magic square," I said.

{…}

Jung takes Rhine seriously (at the time, everyone did); argues that time and distance are irrelevant; tries to make an argument for astrology (curiously enough Pauli, no doubt under the influence of Kepler, encouraged him in this), but why align the individual human destiny with the positions of the planets against the fixed stars? or more than once, anyway; the next coincidence might as well be with the arrangement of trash trucks in a parking lot in Philadelphia on the first day of Spring in 1973, or a message spelled out by driftwood on the seashore, or …

{…}

Pauli on Kepler and the *harmonia mundi*:

He was fascinated by the old Pythagorean idea of the music of

the spheres (which, incidentally, also played no small part in contemporary alchemy) and was trying to find in the movement of the planets the same proportions that appear in the harmonious sounds of tones and in the regular polyhedra. For him, a true spiritual descendant of the Pythagoreans, all beauty lies in the correct proportion, for "*Geometria est archetypus pulchritudinis mundi*" (Geometry is the archetype of the beauty of the world).[167]

{...}

One obvious thing about paranormal phenomena that this idea seems to capture is that they should lie beyond the domain of events that can be shaped by the (conscious) will; that the whole idea of occult powers which can be *exerted* to read or control the minds of others, or (say) deliberately affect the spin of the roulette wheel, must be abandoned. — Such things may come to pass, may be destined, may even in a way derive from within the psyche — in some unusual sense of "within" — but they cannot be *made* to happen — they are imposed. You can perceive the reality of your destiny, but not change it. — Only in this fashion could the complementary schemes of explanation be consistent.

{...}

Though it is not causality per se that is being abandoned in this alternative scheme of connection — the traditional idea of action at a distance is quite as strange and wholly unnatural, indeed Newton himself was only able to accept it after his study of alchemy

---

[167] Wolfgang Pauli, "The Influence of Archetypal Ideas on the Scientific Theories of Kepler", *Writings on Physics and Philosophy*, pp. 219-280; ed. Charles P. Enz and Karl von Meyenn; translated by Robert Schlapp. Berlin: Springer-Verlag, 1994.

accustomed him to the concept of occult powers [168] — rather, it is *local* causality, the idea that causal influence can only be exerted in the immediate spatiotemporal neighborhood of a source; that signals *propagate*, rather than being felt instantaneously irrespective of distance or time.[169] If Donne's wife in her anguish had been able to scream loudly enough that he could have heard her on the Continent, she would also have been heard — in sequential order — by everyone else in between; but if the information was transmitted by a vision, not necessarily so — perhaps we can explain away why no one else receives the vision by some idea of being tuned in on some psychic radio dial, but it doesn't seem as though Donne would have had better reception if he had been closer; and indeed the ideas of spatiotemporal transmission and the iterative communication of information by successive intermediary exchanges seem curiously irrelevant. — No, the idea here is purely Leibnizian: that two clocks have been wound up to coincide. Space and time seem to have little or nothing to do with it. These events are correlated in another matrix.

{…}

Suppose, for instance, the kind of meaningful coincidence that would be the basis of a Hollywood RomCom: the couple-to-be are out in the city walking their dogs separately, they approach an intersection from orthogonal directions, arrive simultaneously, and as one rounds the

---

[168] Cf., e.g., B.J.T. Dobbs, *The Foundations of Newton's Alchemy*, Cambridge: Cambridge University Press, 1983. In hindsight it is incredible that Newton's warning that this idea was only provisional was so completely forgotten once it had been adopted, and it subsequently became the model for all theories of interaction. Faraday's intuition of the field should never have appeared as a radical departure.

[169] I should emphasize in this connection that the only serious attempt to eliminate the field concept (thus local causality) from electromagnetism, and replace it with direct interactions between charges at a distance — that of Wheeler and Feynman — required not only that forces should be felt at a distance with a speed-of-light time delay, but that reactions should be felt both forward and backward in time. — I.e., the idea of action at a distance has no regard for temporality, and I might just as easily in shaking a charge in the antenna of my transmitter be sending a signal back to the Jurassic as to the cell-phone tower at the top of the hill.

corner one way and the other, the other, they are simultaneously distracted by a fortuitous fender-bender, they collide, their dogs' leashes become entangled, confusion, embarrassment, chagrin, apologies as they sort themselves out; and thus they Meet Cute. — Does anything in this story contradict causality? Not at all. All that is required, really, is that each should have started at the right moment. That their clocks should have been synchronized. — That their meeting was *meant to be*.

<div align="center">{…}</div>

Do I believe any of this? Do I believe that the fate of the individual is scripted by Necessity — coded, as it were, into the initial conditions of the universe?

Like Schopenhauer I have had this feeling at times, but my reaction to the coincidences that form its basis is more like one of disbelief: they can't be serious, I think, this is too cute. Surely this is just another joke by the gods at my expense.

<div align="center">{…}</div>

Here's a good one: I am sitting at my work table reading with my phone set off to one side, dark and dormant. Without any (apparent) action on my part, it wakes up, displays the time — 4:32 — and promptly goes back to sleep. — Was this the result of my making some unconscious motion that awakened the phone? (In which case, how did I know what time it was?) — A joke programmed into its source code by some comedian? — Or a joke programmed into the source code of the world?

<div align="center">{…}</div>

— all right, I give up — the Uncanny Experience:

I dropped acid one afternoon[170] and, though it was not approved behavior (perhaps I was just trying to avoid more bad interactions), wandered off by myself to my room. Everything seemed unnaturally still. (The cliché was that you could feel the walls breathing, but this time it felt like they were holding their breath. — If that makes sense. — No, it doesn't have to make sense, does it? Anyway that's what it felt like.) I pulled out the *I Ching* and my designated coins, and started flipping hexagrams to tell my fortune. The first few didn't seem to say much, which was rather boring, so I decided to make something interesting happen: I began *selecting* the hexagrams that I wanted to read by *directing* the fall of the coins. Finally I pushed back the pronouncements of the oracle to the one at the beginning of the book; read some suitably dramatic shit about the dragon lying hidden in the deep and the unshakable cool of the Superior Man, and gave up — too frightened, I think, to carry the experiment any further.

So what did this mean?

Well, I was really stoned. But more than that, I seemed to have achieved a state of coherence. I was, as the scientologists aptly[171] put it, Clear. There was a great stillness in my mind. All noise had been removed. I seemed to have attained some kind of perfection. — In this, the Stone Age of music visualization, we would often feed the outputs of the stereo into an oscilloscope; the results on the display (as Fourier analysis would suggest) were usually complicated and somewhat random, but there had once been one magic moment in "Piece of my Heart" where Janis had hit a mathematically perfect note, and to our stoned amazement a perfect ellipse had appeared on the screen. — It felt like that.

Nonetheless I never really understood what happened. It appeared that I was flipping the coins and calling the outcomes *before* they

---

[170] This was probably in the Fall of 1968.

[171] This may be the only thing they got right.

happened. — What was this? — False memory? — Was I just choosing to interpret the results according to my anticipations of what they ought to be? just seeing what I wanted to see and not segregating desire from observation adequately? Hallucinating? — It didn't seem like that, nothing was moving, everything seemed rather flat and colorless, black and white. — I was flipping the coins very slowly, so that the outcome was easy to fix in advance? — The flip wasn't really that slow, but somehow I was physically determining the outcome anyway? (I had seen things like that happen, physical capability strangely amplified, people shooting pool on acid, for instance.) — The drugs had scrambled the time-ordering of my memory? — I simply wasn't sure.

The funny part was that I was wondering all this *while it was happening* and still couldn't decide what was going on. Because even stoned out of my gourd I was skeptical to the core. — Was it that I was living Moore's paradox? that I was making the oracle tell me what I wanted to hear, and I still didn't believe it? — No, more likely this, just the old joke once again: I wouldn't belong to any cosmic destiny that would have me as a member.

{…}

What did Leibniz mean when he said this was the best of all possible worlds? He was one of the inventors of the calculus of variations, in which you solve problems like the curve of fixed length which encloses the maximum area (the circle) and the curve of constraint along which a body under the influence of gravity falls between two fixed points in minimum time (the brachistochrone), and he thought sufficient reason must pick out a world that maximized some quality/quantity as the one that actually exists.

So far as anyone can tell this idea — or at least, this scheme for generating ideas — is exactly right. Variational principles are fundamental in theoretical physics: the evolution of a mechanical

system obeys the principle of least action, the Boltzmann distribution in statistical mechanics is the one that minimizes information/ maximizes entropy under the constraint of constant total energy, transition amplitudes in quantum mechanics are determined by a looking-glass version of the principle, the Feynman path integral; the evolution of the entire cosmos must somehow involve a sum over all possible geometries — all this is testimony to the genius of Leibniz.

But his own version of the idea was different. There is a silly theological interpretation, the one Voltaire ridiculed, which says that all is for the best in this best of all possible worlds (and moves the goalposts in the definition of "good" as needed),[172] but his real idea, the idea of the logician, was something more like mutual consistency: that possible existents were not all mutually compatible, and that the world as it is realized represents a sort of *maximal* compossible set. — Accordingly everything that *could* exist, *did* exist; and thus the principle of sufficient reason and the complementary principle of plenitude.[173]

So how would Leibniz have interpreted Schopenhauer? Once he had understood the factorization of information about the mechanical evolution of the world into initial conditions and laws of motion expressed in terms of differential equations, he would have wanted to select the set of initial conditions that led to the most *interesting* cosmos. — The one, we might guess, that harbored the most meaningful coincidences. The maximally entangled knot of individual destinies. The cosmic fugue with the largest number of independent parts and internal symmetries. — Of course "interesting" and "meaningful" are still problematic, but he would have waved all that away.

---

[172] F.H. Bradley was funnier: this is the best of all possible worlds, and everything in it is a necessary evil.

[173] The principle of the identity of indiscernibles can also be taken as a corollary.

Is this reasoning all that radical? Actually not. In contemporary cosmological speculation theoreticians run up against the embarrassing problem that physical theories, viz. the so-called Standard Model of elementary particle physics, depend upon a fairly large set of free parameters, the values of fundamental physical constants like the strengths of interactions, the masses of the fundamental particles, etc. — a few dozen of them — and that for almost all choices of these apparently arbitrary parameters the physical world we see could not exist. — What does Leibniz tell us here? that there are two alternatives: either sufficient reason can be found for the existing choices — thus e.g. the eccentric literature which attempts to derive the value of the electromagnetic fine structure constant from pure reason — or, if no sufficient reason exists, then (the principle of plenitude) all possibilities must be realized.

And sure enough, the scientific tabloids are full of speculation about the so-called multiverse. — About which let me remark that I actually prefer the version that appears in the comic books: Kang the Conqueror is more entertaining than inflationary cosmology, and the reasoning behind him is probably more rigorous.

One shouldn't exaggerate the degree to which this principle of maximizing (as one might put it) the elegance of the cosmos is identical with Schopenhauer's notion of the Will dreaming his personal destiny: from the standpoint of theoretical physics, human individuals don't really exist, at least not as simple (significant) objects — the idea instead would be to impose some kind of maximal symmetry group, and then (since the vacuum is always the simplest realization of any such symmetry) to ask what happens when the system is tweaked, ever so slightly, away from perfection; the idea of spontaneously broken symmetry is the one really deep philosophical idea to emerge from elementary particle physics.[174]

In any case Leibniz and Schopenhauer would want to do more than turn a few knobs on the front of the console and tune in the cosmos where — say — the zeros of the Riemann zeta function determine the elementary particle spectrum. They'd want the whole state of the universe to follow uniquely from pure reason *a priori*. Or at least they'd want to come close enough to declare victory and throw a party.

{…}

Since there's no use pretending we aren't trying to interpret the world as a work of art — with aesthetic criteria — and imagining ways in which it might attain maximum semantic content — the best guess — here I am almost serious — would be something like a principle of maximum information content; doesn't it seem as though a universe that permits the evolution of life is better than one that does not? for instance. — This would presumably entail properties like the existence of stable matter, etc., that are otherwise not guaranteed.

---

[174] It is analogous to the idea of the swerve in Lucretius.

But really the problem is too difficult to approach directly. We have to try first on simpler models.

$$\{\ldots\}$$

The principle of local causality states that an event can only have a physical effect on other events which are close to it.[175] What one might call the Proustian view of space and time[176] is that events separated by a considerable distance in time and space may in some other sense be essentially adjacent — Donne in France and his wife in her agony, for example.

If purely as a mathematical question we ask whether the same set of events can have two entirely different spatiotemporal interpretations, the answer at the level of point set topology is already trivially yes. If we ask whether this can remain the case at higher levels of structural complexity, the answer is still yes: one familiar example is the classification of norms on the rational numbers consistent with multiplication which allow the definition of a distance satisfying the axioms for a metric space:

$$\|x\| = 0 \iff x = 0$$
$$\|x \cdot y\| = \|x\| \cdot \|y\|$$
$$\|x + y\| \leq \|x\| + \|y\|$$
$$d(x, y) \equiv \|x - y\|$$

implying

$$d(x, y) = 0 \iff x = y$$
$$d(x, y) = d(y . x)$$
$$d(x, z) \leq d(x, y) + d(y, z)$$

---

[175] We actually want to say *infinitesimally* close, which requires a lengthy digression to define adequately. So we won't.

[176] In re which see below (xv), (3/28/01).

Such norms can be shown (up to scaling) to be of two types, the familiar

$$\|x\| = |x|$$

and the p-adic

$$\|x\| = \left\|\frac{a}{b}\right\| = \frac{1}{p^{ord_p(a)}}$$

where p is a prime, and

$$ord_p(a) = max\{n \geq 0 : p^n | a\}$$

Thus 1 and 65537 are close 2-adically, though in the standard metric (or any other p-adic metric) they are rather far apart.

,A more recondite example is provided by the existence of exotic differentiable structures on four-dimensional Euclidean space,[177] but — at any rate — it isn't that difficult to invent possible worlds in which entirely different conceptions of space and time can coexist. Though the question remains, of course, what correspondence they may have with this one.

{…}

The other possibility that always seemed intriguing was the Gödel ambiguity: that the truth of the world might not all be provable from the axioms. — One might think of proof as being like causality, and synchronicity being whatever gets added afterward.

---

[177] See Michael H. Freedman and Frank Quinn, *Topology of 4-Manifolds*. Princeton: Princeton University Press, 1990.

(Though this seems less less a new perspective on synchronicity than a new perspective on logic. — Save that thought…)
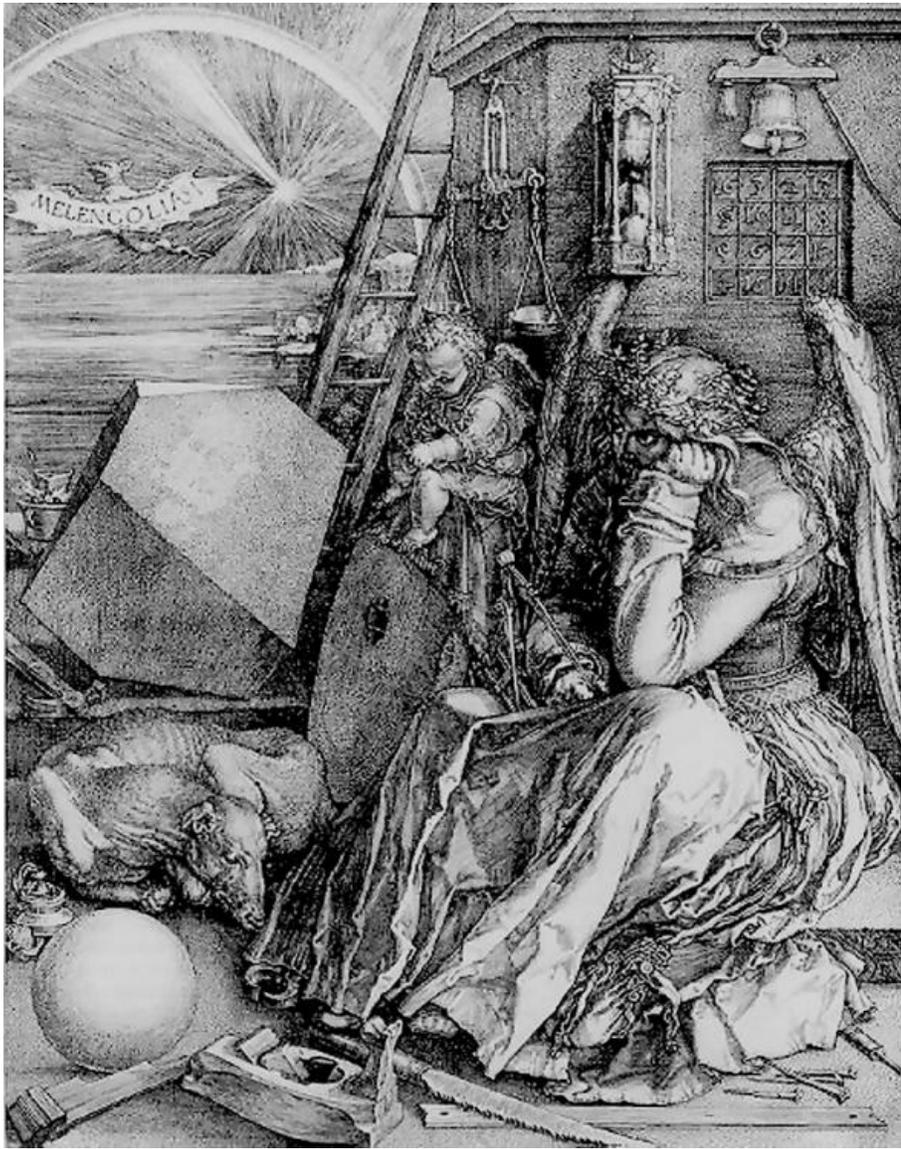
Magic squares (1972)[178]

On the wall above the piano was an arithmetical diagram fastened with drawing-pins, something he had found in a second-hand shop: a so-called magic square, such as appears also in Dürer's *Melancolia*, along with the hour-glass, the circle, the scale, the polyhedron, and other symbols. Here as there, the figure was divided into sixteen Arabic-numbered fields, in such a way that number one was in the right-hand lower corner, sixteen in the upper left; and the magic, or the oddity, simply consisted in the fact that the sum of these numerals, however you added them, straight down, crosswise, or diagonally, always came to thirty four. What the principle was upon which this magic uniformity rested I never made out, but by virtue of the prominent place Adrian had given it over the piano, it always attracted the eye, and I believe I never visited his room without giving a quick glance, slanting up or straight down and testing once more the invariable, incredible result.

— Thomas Mann: *Doctor Faustus*.[179]

There is something entirely characteristic in this obsessive determination to find (or create) a sort of perfection in randomness. A retrodictive version of destiny.

---

[178] All too typically, though the ideas date from the Nixon administration, I only got around to working them out during the pandemic, when I spent a lot of time lying on the couch watching movies and daydreaming about mathematics. The ensuing results, such as they are, were communicated in a series of letters to Mark Lankton and Sam Galetar, who probably only found them entertaining because they were as bored as I was.

[179] Translated by H.T. Lowe-Porter. New York: Alfred A. Knopf, 1948.

A *magic square* of rank n is an n by n array of the numbers[180]

$$0, 1, \ldots, n^2 - 1$$

with the property that the sums along every row, column, and diagonal are identical. If this is the case it is easy to see the distinguished value (sometimes called the *magic number*) must be

$$\frac{n(n^2 - 1)}{2}$$

thus 12 for n=3, 30 for n=4, etc.

Magic squares have been studied since antiquity: the unique example (up to symmetry) of rank 3, known as *lo shu*, was written down by the Chinese mathematician Xu Yu around 190 B.C.:

```
2   9   4
7   5   3
6   1   8
```

(Fortran convention)

and was probably known long before that. Yoke[181] notes that "it was used mainly in astrology and divination ... to interpret not only natural

---

[180] We employ as it were the C (count starting at 0) and not the Fortran (count starting at 1) convention; except when we don't. This can't be any more confusing than usual.

[181] Yoke, Ho Peng: "Magic squares in China"; *Encyclopedia of the History of Science, Technology, and Medicine in Non-Western Cultures*, ed. Helaine Selin (2nd edition). Berlin: Springer-Verlag, 2008; pp. 1251-1252. See also the companion article by Takao Hayaashi, "Magic Squares in Indian Mathematics", pp. 1252-1259.

phenomena, but also philosophy and human behavior." Indian mathematicians later listed among the uses of magic squares of rank 4 neutralizing poison, protecting crops from insects, and exorcizing evil spirits.

The Dürer magic square

```
16   3   2  13
 5  10  11   8
 9   6   7  12
 4  15  14   1
```

which appears in the   etching reproduced above, dates, in fact — clever Albrecht — precisely from the year 1514.

Algorithms for constructing magic squares were devised by Fermat, Euler, and Benjamin Franklin, and can be found in the juvenilia of Ramanujan.

A severe combinatorial explosion seems to place a general classification beyond the realm of possibility: for n = 4, there are 880 squares [Bernard Frénicle de Bessy, 1693]; for n = 5, 275,305,224 [Richard Schroeppel, 1973], for n = 6 doubtless something even more ridiculous — but certain special cases are tractable, and there is an extensive literature on the subject, which as usual I will ignore.

It is obvious that sums of rows remain the same if the last row is moved to the top, and columns similarly if the rightmost is moved to the left or vice-versa.[182] Diagonals are not preserved by these

---

[182] I.e. they are invariant under cyclic permutations. — In fact any permutation will work, but not without scrambling the diagonals. (One may contrive combined permutations of rows and columns that leave the elements on the diagonals invariant, but why bother.)

transformations, however, and thus it is natural to consider instead the *magic toroid*, i.e. a magic square which retains the summation properties when rolled left or right or up or down, and thus also yields the same sum along "broken" diagonals which run off one side and resume on the other.

The simplest example is of order 4

```
 0   7   9  14
11  12   2   5
 6   1  15   8
13  10   4   3
```

which one can imagine either drawn upon a cylinder that rolls up both ways, i.e. a donut, or as an infinite repeating pattern in the plane in which in any row, column, or diagonal any four numbers in a row add up to 30:

```
    .    .    .    .    .    .    .    .    .    .    .    .
      .    .    .    .    .    .    .    .    .    .    .    .
    .    .    .    .    .    .    .    .    .    .    .  .|   .
...  0   7   9  14   0   7   9  14   0   7   9  14 ...
... 11  12   2   5  11  12   2   5  11  12   2   5 ...
...  6   1  15   8   6   1  15   8   6   1  15   8 ...
... 13  10   4   3  13  10   4   3  13  10   4   3 ...
...  0   7   9  14   0   7   9  14   0   7   9  14 ...
... 11  12   2   5  11  12   2   5  11  12   2   5 ...
...  6   1  15   8   6   1  15   8   6   1  15   8 ...
... 13  10   4   3  13  10   4   3  13  10   4   3 ...
    .    .    .    .    .    .    .    .    .    .    .    .
      .    .    .    .    .    .    .    .    .    .    .    .
    .    .    .    .    .    .    .    .    .    .    .    .
```

This matrix has the properties:

— The fundamental array is (up to symmetry) unique.

— It is invariant under the maps

$$(i, j) \rightarrow (i \pm n, j), (i, j \pm n)$$

where here n = 4. Thus we can simply take coordinates modulo n.

— The map

$$(i, j) \rightarrow (i + \frac{n}{2}, j + \frac{n}{2})$$

carries an entry to its *complement*, i.e.

$$v\left(i + \frac{n}{2}, j + \frac{n}{2}\right) = n^2 - 1 - v(i, j)$$

— The sum of the entries in any 2-by-2 subsquare is equal to the magic number.

At this point, of course, it becomes an entertaining exercise to extend this solution by a combination of brute force and intuition (guided largely by considerations of symmetry) to the 8-by-8 [C convention]

```
 0 55 31 40 22 33  9 62
59 12 36 19 45 26 50  5
47 24 48  7 57 14 38 17
20 35 11 60  2 53 29 42
41 30 54  1 63  8 32 23
18 37 13 58  4 51 27 44
 6 49 25 46 16 39 15 56
61 10 34 21 43 28 52  3
```

and (Fortran convention) 16-by-16

1	248	32	233	128	137	161	88	170	95	119	130	23	226	10	255
252	13	229	20	133	116	92	173	83	166	142	123	238	27	243	6
48	217	49	200	145	104	80	185	71	178	154	111	58	207	39	210
213	36	204	61	108	157	181	68	190	75	99	150	195	54	222	43
192	73	97	152	193	56	224	41	215	34	202	63	106	159	183	66
69	180	156	109	60	205	37	212	46	219	51	198	147	102	78	187
81	168	144	121	240	25	241	8	250	15	231	18	135	114	90	175
172	93	117	132	21	228	12	253	3	246	30	235	126	139	163	86
87	162	138	127	234	31	247	2	256	9	225	24	129	120	96	169
174	91	115	134	19	230	14	251	5	244	28	237	124	141	165	84
186	79	103	146	199	50	218	47	209	40	208	57	112	153	177	72
67	182	158	107	62	203	35	214	44	221	53	196	149	100	76	189
42	223	55	194	151	98	74	191	65	184	160	105	64	201	33	216
211	38	206	59	110	155	179	70	188	77	101	148	197	52	220	45
7	242	26	239	122	143	167	82	176	89	113	136	17	232	16	249
254	11	227	22	131	118	94	171	85	164	140	125	236	29	245	4

which have similar properties — the story with the subsquares is a
trifle more complicated, but the diagonal transform is the same.

Of course the whole point of all this was to generalize the idea to higher dimension. For that I found a couple of interesting algorithms — or two versions of the same algorithm — based on what might be called bit-twiddling.

Begin with the 4 by 4 case.

Write numbers in binary.

Start from 0000.

Define immediate neighbors as bit-vectors which agree with one another in *exactly one place*.

Neighbors of 0000 are then 0111, 1101, 1110, 1011, which are arranged as follows:

```
                1101

        1110    0000    0111

                1011
```

(Does the order matter? No, because the properties of the array are invariant under rotations and inversions around axes and diagonals, which allows you to generate all 24 permutations.)

Then the neighbors of 0111 are 0000, 1100, 1010, 1001, and the neighbors of 1011 are 1100, 0000, 0110, 0101, meaning you can fill in the lower right corner uniquely:

```
                1101

        1110    0000    0111

                1011    1100
```

The other corners are determined in similar fashion, giving you:

```
        0011    1101    1010

        1110    0000    0111

        0101    1011    1100
```

Now you exploit the requirement that array coordinates are modulo 4 (it's a torus), so that the entry to the right of 1010 also neighbors 0011 and thus must be 0100, the entry below 0101 also neighbors 0011 and must be 1000, etc. So you can extend the 3 by 3 array to a 4 by 4, which gives you:

```
        0011    1101    1010    0100

        1110    0000    0111    1001

        0101    1011    1100    0010

        1000    0110    0001    1111
```

i.e., the magic toroid

```
 3 13 10  4
14  0  7  9
 5 11 12  2
 8  6  1 15
```

You see from the bit-vectors that shifting over one column and then down one row changes exactly two bits, and repeating this operation changes all of them, i.e. sends a number x to (15 - x), which when you count from 0..15 is the complement in the magic-square sense. That guarantees the toroidal property. —

This construction is unique only when you go 0..3 in each coordinate, but you can do this in arbitrary dimension, and if the ("polar") diagonal mapping is defined as the iteration of

$$(i, j, k, \ldots) \rightarrow (i \pm 1, j \pm 1, k \pm 1, \ldots)$$

then a unit shift along the diagonal changes 3 bits out of 6 in dimension 3, 4 bits out of 8 in dimension 4, etc.[183] — So what this tells you is that you can construct a magic toroid for any dimension d, with 4^d elements.

(There is a missing flattening algorithm to explain how you can turn these into magic toroids in the plane so that the 3-cube becomes an 8 by 8, the 4-cube a 16 by 16, the 5-cube presumably a 32 by 32, etc., but we'll get to that presently.)

After dithering over data representation, we settle on the usual half-man, half-beast combination of elegance and kludge:

_____

[183] In fact this property motivates this definition of the diagonal.

```
31 > *square*
#2A((0 13 3 14) (7 10 4 9) (12 1 15 2) (11 6 8 5))

35 > *cube*
#3A(((0 61 3 62) (55 10 52 9) (12 49 15 50) (59 6 56 5)) ((31 34 28 33) (40 21 43 22) (19
    46 16 45) (36 25 39 26)) ((48 13 51 14) (7 58 4 57) (60 1 63 2) (11 54 8 53)) ((47 18
    44 17) (24 37 27 38) (35 30 32 29) (20 41 23 42)))


35 > *tesseract*
#4A((((0 253 3 254) (247 10 244 9) (12 241 15 242) (251 6 248 5)) ((223 34 220 33) (40 213
    43 214) (211 46 208 45) (36 217 39 218)) ((48 205 51 206) (199 58 196 57) (60 193 63
    194) (203 54 200 53)) ((239 18 236 17) (24 229 27 230) (227 30 224 29) (20 233 23
    234))) (((127 130 124 129) (136 117 139 118) (115 142 112 141) (132 121 135 122)) ((160
    93 163 94) (87 170 84 169) (172 81 175 82) (91 166 88 165)) ((79 178 76 177) (184 69
    187 70) (67 190 64 189) (180 73 183 74)) ((144 109 147 110) (103 154 100 153) (156 97
    159 98) (107 150 104 149))) (((192 61 195 62) (55 202 52 201) (204 49 207 50) (59 198
    56 197)) ((31 226 28 225) (232 21 235 22) (19 238 16 237) (228 25 231 26)) ((240 13 243
    14) (7 250 4 249) (252 1 255 2) (11 246 8 245)) ((47 210 44 209) (216 37 219 38) (35
    222 32 221) (212 41 215 42))) (((191 66 188 65) (72 181 75 182) (179 78 176 77) (68 185
    71 186)) ((96 157 99 158) (151 106 148 105) (108 145 111 146) (155 102 152 101)) ((143
    114 140 113) (120 133 123 134) (131 126 128 125) (116 137 119 138)) ((80 173 83 174)
    (167 90 164 89) (92 161 95 162) (171 86 168 85))))
```

```
35 > *fifth*
#5A(((((0 1021 3 1022) (1015 10 1012 9) (12 1009 15 1010) (1019 6 1016 5)) ((991 34 988 33)
    (40 981 43 982) (979 46 976 45) (36 985 39 986)) ((48 973 51 974) (967 58 964 57) (60
    961 63 962) (971 54 968 53)) ((1007 18 1004 17) (24 997 27 998) (995 30 992 29) (20
    1001 23 1002))) (((895 130 892 129) (136 885 139 886) (883 142 880 141) (132 889 135
    890)) ((160 861 163 862) (855 170 852 169) (172 849 175 850) (859 166 856 165)) ((847
    178 844 177) (184 837 187 838) (835 190 832 189) (180 841 183 842)) ((144 877 147 878)
    (871 154 868 153) (156 865 159 866) (875 150 872 149))) (((192 829 195 830) (823 202
    820 201) (204 817 207 818) (827 198 824 197)) ((799 226 796 225) (232 789 235 790) (787
    238 784 237) (228 793 231 794)) ((240 781 243 782) (775 250 772 249) (252 769 255 770)
    (779 246 776 245)) ((815 210 812 209) (216 805 219 806) (803 222 800 221) (212 809 215
    810))) (((959 66 956 65) (72 949 75 950) (947 78 944 77) (68 953 71 954)) ((96 925 99
    926) (919 106 916 105) (108 913 111 914) (923 102 920 101)) ((911 114 908 113) (120 901
    123 902) (899 126 896 125) (116 905 119 906)) ((80 941 83 942) (935 90 932 89) (92 929
    95 930) (939 86 936 85)))) ((((511 514 508 513) (520 501 523 502) (499 526 496 525)
    (516 505 519 506)) ((544 477 547 478) (471 554 468 553) (556 465 559 466) (475 550 472
    549)) ((463 562 460 561) (568 453 571 454) (451 574 448 573) (564 457 567 458)) ((528
    493 531 494) (487 538 484 537) (540 481 543 482) (491 534 488 533))) (((640 381 643
    382) (375 650 372 649) (652 369 655 370) (379 646 376 645)) ((351 674 348 673) (680 341
    683 342) (339 686 336 685) (676 345 679 346)) ((688 333 691 334) (327 698 324 697) (700
    321 703 322) (331 694 328 693)) ((367 658 364 657) (664 357 667 358) (355 670 352 669)
    (660 361 663 362))) (((319 706 316 705) (712 309 715 310) (307 718 304 717) (708 313
    711 314)) ((736 285 739 286) (279 746 276 745) (748 273 751 274) (283 742 280 741))
    ((271 754 268 753) (760 261 763 262) (259 766 256 765) (756 265 759 266)) ((720 301 723
    302) (295 730 292 729) (732 289 735 290) (299 726 296 725))) (((576 445 579 446) (439
    586 436 585) (588 433 591 434) (443 582 440 581)) ((415 610 412 609) (616 405 619 406)
    (403 622 400 621) (612 409 615 410)) ((624 397 627 398) (391 634 388 633) (636 385 639
    386) (395 630 392 629)) ((431 594 428 593) (600 421 603 422) (419 606 416 605) (596 425
    599 426)))) ((((768 253 771 254) (247 778 244 777) (780 241 783 242) (251 774 248 773))
    ((223 802 220 801) (808 213 811 214) (211 814 208 813) (804 217 807 218)) ((816 205 819
    206) (199 826 196 825) (828 193 831 194) (203 822 200 821)) ((239 786 236 785) (792 229
    795 230) (227 798 224 797) (788 233 791 234))) (((127 898 124 897) (904 117 907 118)
    (115 910 112 909) (900 121 903 122)) ((928 93 931 94) (87 938 84 937) (940 81 943 82)
    (91 934 88 933)) ((79 946 76 945) (952 69 955 70) (67 958 64 957) (948 73 951 74))
    ((912 109 915 110) (103 922 100 921) (924 97 927 98) (107 918 104 917))) (((960 61 963
    62) (55 970 52 969) (972 49 975 50) (59 966 56 965)) ((31 994 28 993) (1000 21 1003 22)
    (19 1006 16 1005) (996 25 999 26)) ((1008 13 1011 14) (7 1018 4 1017) (1020 1 1023 2)
    (11 1014 8 1013)) ((47 978 44 977) (984 37 987 38) (35 990 32 989) (980 41 983 42)))
    (((191 834 188 833) (840 181 843 182) (179 846 176 845) (836 185 839 186)) ((864 157
    867 158) (151 874 148 873) (876 145 879 146) (155 870 152 869)) ((143 882 140 881) (888
    133 891 134) (131 894 128 893) (884 137 887 138)) ((848 173 851 174) (167 858 164 857)
    (860 161 863 162) (171 854 168 853)))) ((((767 258 764 257) (264 757 267 758) (755 270
    752 269) (260 761 263 762)) ((288 733 291 734) (727 298 724 297) (300 721 303 722) (731
    294 728 293)) ((719 306 716 305) (312 709 315 710) (707 318 704 317) (308 713 311 714))
    ((272 749 275 750) (743 282 740 281) (284 737 287 738) (747 278 744 277))) (((384 637
    387 638) (631 394 628 393) (396 625 399 626) (635 390 632 389)) ((607 418 604 417) (424
    597 427 598) (595 430 592 429) (420 601 423 602)) ((432 589 435 590) (583 442 580 441)
    (444 577 447 578) (587 438 584 437)) ((623 402 620 401) (408 613 411 614) (611 414 608
    413) (404 617 407 618))) (((575 450 572 449) (456 565 459 566) (563 462 560 461) (452
    569 455 570)) ((480 541 483 542) (535 490 532 489) (492 529 495 530) (539 486 536 485))
    ((527 498 524 497) (504 517 507 518) (515 510 512 509) (500 521 503 522)) ((464 557 467
    558) (551 474 548 473) (476 545 479 546) (555 470 552 469))) (((320 701 323 702) (695
    330 692 329) (332 689 335 690) (699 326 696 325)) ((671 354 668 353) (360 661 363 662)
    (659 366 656 365) (356 665 359 666)) ((368 653 371 654) (647 378 644 377) (380 641 383
    642) (651 374 648 373)) ((687 338 684 337) (344 677 347 678) (675 350 672 349) (340 681
    343 682)))))
```

I'd do a six-dimensional cube, but this is sufficiently ridiculous.

However by way of illustration the three-cube above can be written as

```
 0 61  3 62
55 10 52  9
12 49 15 50
59  6 56  5

31 34 28 33
40 21 43 22
19 46 16 45
36 25 39 26

48 13 51 14
 7 58  4 57
60  1 63  2
11 54  8 53

47 18 44 17
24 37 27 38
35 30 32 29
20 41 23 42
```

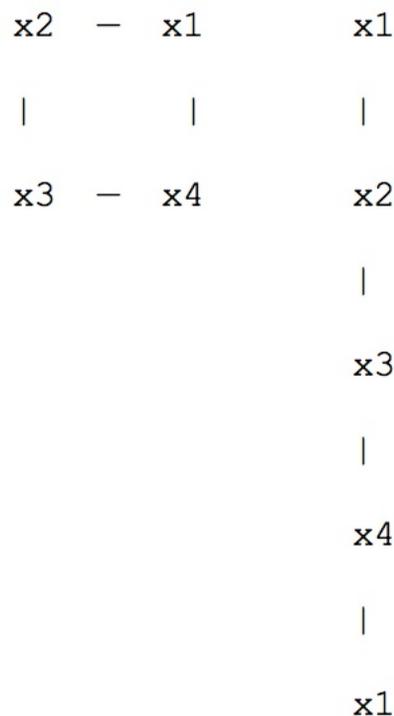which may be pictured   as four squares stacked
on top of one another.   The summation
properties are then easily verified, e.g. $(0 + 61 + 3 + 62) = (0 + 31 + 48 + 47) = (0 + 21 + 63 + 42) = 126$. — A tesseract then consists of four cubes, a 5-cube of four tesseracts, etc.

So why does this work? Really it has almost nothing to do with arithmetic, it's more like a matter of the combinatorics of bit vectors.

The algorithm states that you begin with an initial configuration and then complete the matrix by the rule that adjacent bit vectors are the same in exactly one place.

Also, if you take four steps in one direction you end up where you started.

This says that the square and the loop around the torus are essentially the same:

```
x2  —  x1         x1

 |      |          |

x3  —  x4         x2

                   |

                  x3

                   |

                  x4

                   |

                  x1
```

because either way you're going around a circuit.

This means that, if you look at an individual bit in the vector, every time you take a step it either flips [+] or it doesn't [-], but around a loop it has to flip an even number of times, 2 or 4. It has to flip because it can't not flip twice in a row; if it did, all the other bits in the vector would have to flip twice, and restore a previous value.

So the possibilities are these:

```
        +                                          -
    x2  —  x1                              x2   —   x1

  —  |        |  —                      +  |        |  +

    x3  —  x4                              x3   —   x4

        +                                          -
```

or

```
                    +
                x2  —  x1

              +  |        |  +

                x3  —  x4

                    +
```

Thus, e.g., the original 4 by 4 example

```
        0011      1101      1010      0100

        1110      0000      0111      1001

        0101      1011      1100      0010

        1000      0110      0001      1111
```

viewed bit-column by bit-column looks like this:

[1]	0	1	1	0
	1	0	0	1
	0	1	1	0
	1	0	0	1
[2]	0	1	0	1
	1	0	1	0
	1	0	1	0
	0	1	0	1
[3]	1	0	1	0
	1	0	1	0
	0	1	0	1
	0	1	0	1
[4]	1	1	0	0
	0	0	1	1
	1	1	0	0
	0	0	1	1

Which is a pretty simple pattern, and has a number of properties not previously remarked: parallel rows and columns have the same patterns of bit-flipping, for instance, and knight moves flip one bit in each vector.

Diagonal moves flip two bits. (In the square; in higher dimension exactly half of them, because the cubic diagonal takes three steps, etc.)

Now suppose we stack the bit-vectors one on top of the other — row or column or — ha! something extra! — even around a square. The requirement that there have to be an even number of flips implies that in each column there are two ones and two zeroes. (See above, e.g.)

Now, if you add a list of numbers in binary, if there are always the same number of ones in a column their order is irrelevant, the answer is always going to be the same. (A sort of bitwise commutative law.)

And here there are always going to be two in each column. So the first entry on the right in the sum will be 0, and with carries every entry to the left will be 1, with one additional 1 at the left.

So the total, for bit vectors of length 2n bits, will always be

$$2 \bullet (2^{2n} - 1)$$

For n = 2 this is 30, for n= 3 this is 126, for n = 4 this is 510, and so on.

The requirement for a magic 4-cube of dimension d is that the sum in any row or column will be the sum of all the numbers from 0 to $4^d - 1$, i.e. $4^d \bullet \dfrac{4^d - 1}{2}$, divided by the number of rows/columns/..., which is 4^{d-1}. So the answer in that case is also $2 \bullet (4^d - 1)$, i.e. 30, 126, 510, etc.

— Ah, the diagonal. What happens there? In that case if you make two diagonal steps in a row, (+-2,+-2), you first flip half the bits, and then the other half. This transforms the bit-vector into its complement,

and therefore any diagonal (starting anywhere) contains two numbers m, n, $(4^d - 1 - m, 4^d - 1 - n)$, or to repeat the previous argument, every 1 is paired with a 0, so there are (again) exactly two ones in each bit-vector column. So the sum, again, is $2 \bullet (4^d - 1)$.

— And, again, the sum of any 2 by 2 square is the same as the sum of a row, column, or diagonal.

$$\{...\}$$

An aside on topological variants:

With an odd number of bits you appear to be hosed, but in the plane, at least, one can construct a sort of magic Klein bottle, where you sum four in a row in a 4 by 4 array and all the diagonals work but there's a twist introduced and you use each of the numbers 0 .. 7 twice.

I.e., you have an arrangement

g	h	a	b	....
a	b	g	h	...
c	d	e	f	...
e	f	c	d	...
g	h	a	b	...

which continues in all directions, constructed by pasting the top to the bottom with the same orientation and the third column is identified with the first with opposite orientation. Then you start with 000 and use the same adjacent rule to complete it up and to the right, but since you have only three neighbors the one below it is the complement:

```
000        011

111        100

010        001

101        110
```

(The necessity for something like this becomes obvious if you map adjacency for 3-bit vectors: 000 is connected to 011,101,110 and 111 is connected to 100,010,001, but the two components are disjoint.)

This fills out into a matrix

```
0        3        5        6

7        4        2        1

2        1        7        4

5        6        0        3
```

which has the same properties, i.e. all rows, columns, diagonals, and squares add up to 14, but which only uses half as many numbers and repeats them twice.

After which you naturally start wondering how to pull this off for the Poincaré dodecahedral space, but lacking immediate inspiration we'll drop that line of inquiry and continue —

$$\{\ldots\}$$

The key to extending the algorithm to d-toroids of rank greater than 4 is, perversely, to first solve the problem for rank 2 — i.e., the Boolean n-cube $\{0,1\}^n$, the set of bit-strings of length n.

Assume first that n is even. Then every vertex of the cube has binary coordinates

$$v = (x_1, \ldots, x_n), \ x_i = 0,1$$

Label the vertices as follows: if the number of bits in **v** is even, leave it alone; if the number of bits is odd, take the binary complement. Thus e.g.

$$v = (0,0, 1,1) \rightarrow v' = (0,0, 1,1)$$

but

$$v = (1,0, 1,1) \rightarrow v' = (0,1, 0,0)$$

Then, again, the sum of the labels around any face of the cube is

$$2 \cdot (2^n - 1)$$

because every bit appears twice. (One can call this the "sort-of-magic" number.)

In the simplest nontrivial even-dimensional case, d = 4, the cube before inversion is:
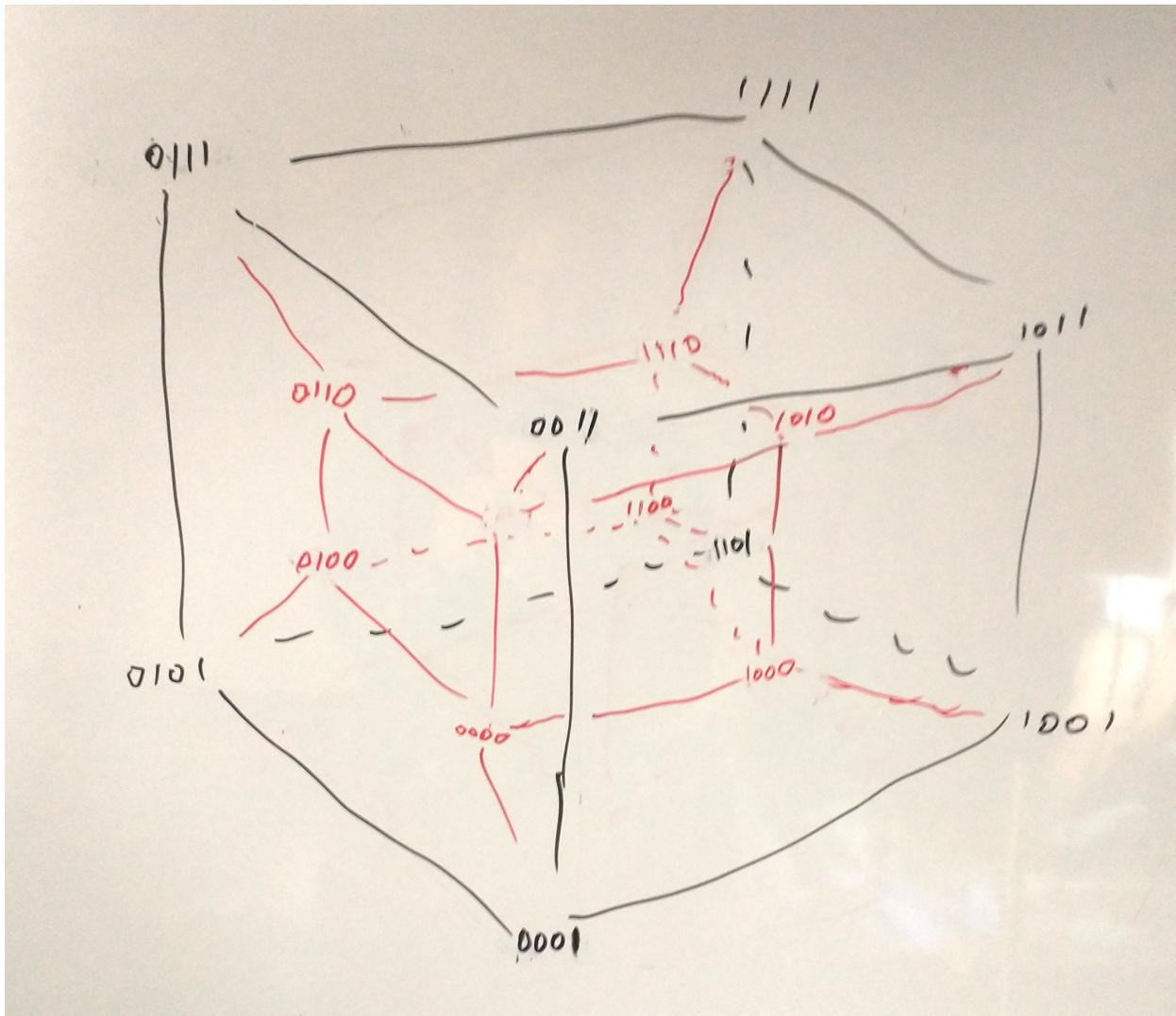
And afterward:

Here every vertex has four adjacent vertices with coordinates differing in every place but one, the sums around squares (faces) are identical (and equal 30), and opposite vertices are Boolean complements. — I.e. this is equivalent to the previous algorithm, and geometrically more perspicuous; at least, once we have described how to flatten it.

Flattening is defined via what might seem a peculiar way of linearizing the bit vectors, i.e. $\{0,1\}^3$ flattens as

1000

1111

0100

0001

0110

0011

1010

110T

1100

1011

0010

0101

0111

0000
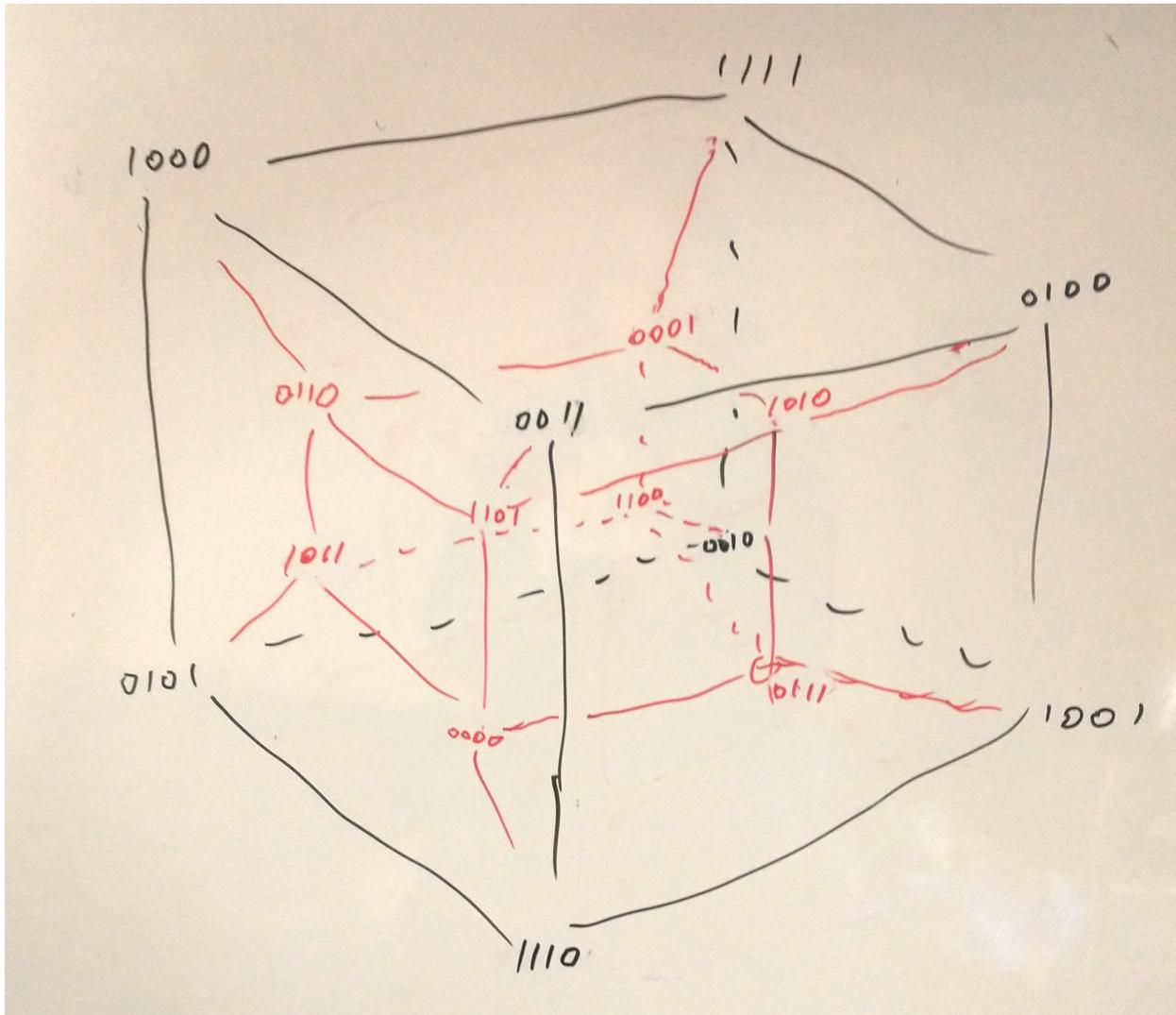
1001

1110

((0 0 0) (1 0 0) (1 1 0) (0 1 0) (1 1 1) (0 1 1) (0 0 1) (1 0 1))

and not, say, as the vectors in lexicographic order. But the definition is designed [a] to be recursive and [b] to ensure that in a vector of 2^k elements the second half are the complements of the first, in the same order.

I.e. we define

$$vectors[0] = nil$$
$$vectors[1] = ((0)(1))$$
$$vectors[n+1] = vectors[n] \times 0 + complement(vectors[n] \times 1)$$

where "+" has the meaning of list concatenation. Then in a product

$$2^{kd} = 2^{k^d}$$

you take d k-bit vectors as coordinates and apply the even/odd rule to fill in the entries in the d-cube.

For $2^4 = 4^2$ this is a table of

$$((0\ 0)\ (1\ 0)\ (1\ 1)\ (0\ 1)) \times ((0\ 0)\ (1\ 0)\ (1\ 1)\ (0\ 1))$$

where the [2,3] entry (C convention!), e.g., is (1 1 0 1) which flips to (0 0 1 0).

Writing this out yields

$$(0\ 13\ 3\ 14)$$
$$(7\ 10\ 4\ 9)$$
$$(12\ 1\ 15\ 2)$$
$$(11\ 6\ 8\ 5)$$

which is toroidal as defined previously, has the shift property on diagonals, and has all the clever little internal symmetries like the sums of all the 2 by 2 squares being the same. — To enumerate them explicitly, there are exactly 24 = 16 plus 8 2 by 2 squares plus rows/columns, and these correspond exactly to the 24 2-dimensional 2 by 2 faces of the Boolean cube in 4 dimensions. There are two complementary pairs in each of 8 diagonals, each counted twice, and there are 8 pairs of polar opposites in the Boolean 4-cube. So every identity is, as it were, accounted for.

In the general case of a cube of rank $r = 2^k$ and dimension d there are a great many internal symmetries, but the pattern isn't quite as simple. The Boolean cube of dimension $n = kd$ has 2^{n-1} pairs of polar opposites, corresponding exactly to the 2^{kd-1} complementary pairs of

entries related by diagonal shifts of $\frac{r}{2} = 2^{k-1}$ in the flattened object; however for $n > 2$, $\{0,1\}^n$ has

$$(n^2 - n)2^{n-3}$$

faces, 6 for $n = 3$, 24 for $n = 4$, 80 for $n = 5$, etc., each providing an equation for 4 entries, meaning that the number of relations far exceeds the requirements of magicality per se.

Explicitly, in a flattened product

$$I \times I \times \ldots \times I \, (d \, copies)$$

where

$$I = \left(x_1, \ldots, x_k\right), \ x_i = 0,1$$

a (generalized) row consists of the 2^k values which result from fixing coordinates in each of the copies save one:

$$v_1 \times \ldots \times I \times v_{j+1} \times \ldots \times v_n$$

and there are $\frac{k(k-1)}{2}$ faces of the Boolean cube that map into I, meaning there are that many ways of proving that the sum of those 2^k values is $2^{k-2} \bullet 2 \bullet (2^k - 1)$. The additional relations provided by faces which map into two different components yield squares whose values add to the sort-of-magic number.

King Kong versus the minor hack (9/8/2020)

Here in the original letter, after a lengthy discussion of the correct definition of "diagonal" in higher dimension,[184] I note

> … alarming new development: after incorporating peanut butter and honey sandwiches into my diet again, I find that notebook pages now have a tendency to stick together…

and append a listing for a magic toroid of dimension 3 and rank 16 with 4096 entries:

---

[184] One question I originally set out to answer was whether it is possible to construct a d-cube in which not simply the polar diagonals, but the face diagonals sum to the magic number. The usual combination of argument and experiment indicates that there is no elegant solution to this problem; which, accordingly, I dismissed.

```
(((0 4094 3 4093 4088 6 4091 5 15 4081 12 4082 4087 9 4084 10)
    (4079 17 4076 18 23 4073 20 4074 4064 30 4067 29 24 4070 27
     4069)
    (48 4046 51 4045 4040 54 4043 53 63 4033 60 4034 4039 57 4036
     58)
    (4063 33 4060 34 39 4057 36 4058 4048 46 4051 45 40 4054 43
     4053)
    (3983 113 3980 114 119 3977 116 3978 3968 126 3971 125 120 3974
     123 3973)
    (96 3998 99 3997 3992 102 3995 101 111 3985 108 3986 3991 105
     3988 106)
    (4031 65 4028 66 71 4025 68 4026 4016 78 4019 77 72 4022 75
     4021)
    (80 4014 83 4013 4008 86 4011 85 95 4001 92 4002 4007 89 4004
     90)
    (240 3854 243 3853 3848 246 3851 245 255 3841 252 3842 3847 249
     3844 250)
    (3871 225 3868 226 231 3865 228 3866 3856 238 3859 237 232 3862
     235 3861)
    (192 3902 195 3901 3896 198 3899 197 207 3889 204 3890 3895 201
     3892 202)
    (3887 209 3884 210 215 3881 212 3882 3872 222 3875 221 216 3878
     219 3877)
    (3967 129 3964 130 135 3961 132 3962 3952 142 3955 141 136 3958
     139 3957)
    (144 3950 147 3949 3944 150 3947 149 159 3937 156 3938 3943 153
     3940 154)
    (3919 177 3916 178 183 3913 180 3914 3904 190 3907 189 184 3910
     187 3909)
    (160 3934 163 3933 3928 166 3931 165 175 3921 172 3922 3927 169
     3924 170))
  ((3839 257 3836 258 263 3833 260 3834 3824 270 3827 269 264 3830
     267 3829)
    (272 3822 275 3821 3816 278 3819 277 287 3809 284 3810 3815 281
     3812 282)
    (3791 305 3788 306 311 3785 308 3786 3776 318 3779 317 312 3782
     315 3781)
    (288 3806 291 3805 3800 294 3803 293 303 3793 300 3794 3799 297
     3796 298)
    (368 3726 371 3725 3720 374 3723 373 383 3713 380 3714 3719 377
     3716 378)
    (3743 353 3740 354 359 3737 356 3738 3728 366 3731 365 360 3734
     363 3733)
    (320 3774 323 3773 3768 326 3771 325 335 3761 332 3762 3767 329
     3764 330)
    (3759 337 3756 338 343 3753 340 3754 3744 350 3747 349 344 3750
     347 3749)
    (3599 497 3596 498 503 3593 500 3594 3584 510 3587 509 504 3590
     507 3589)
    (480 3614 483 3613 3608 486 3611 485 495 3601 492 3602 3607 489
     3604 490)
    (3647 449 3644 450 455 3641 452 3642 3632 462 3635 461 456 3638
     459 3637)
    (464 3630 467 3629 3624 470 3627 469 479 3617 476 3618 3623 473
     3620 474)
    (384 3710 387 3709 3704 390 3707 389 399 3697 396 3698 3703 393
     3700 394)
    (3695 401 3692 402 407 3689 404 3690 3680 414 3683 413 408 3686
```

```
 411 3685)
(432 3662 435 3661 3656 438 3659 437 447 3649 444 3650 3655 441
 3652 442)
(3679 417 3676 418 423 3673 420 3674 3664 430 3667 429 424 3670
 427 3669))
((768 3326 771 3325 3320 774 3323 773 783 3313 780 3314 3319 777
 3316 778)
(3311 785 3308 786 791 3305 788 3306 3296 798 3299 797 792 3302
 795 3301)
(816 3278 819 3277 3272 822 3275 821 831 3265 828 3266 3271 825
 3268 826)
(3295 801 3292 802 807 3289 804 3290 3280 814 3283 813 808 3286
 811 3285)
(3215 881 3212 882 887 3209 884 3210 3200 894 3203 893 888 3206
 891 3205)
(864 3230 867 3229 3224 870 3227 869 879 3217 876 3218 3223 873
 3220 874)
(3263 833 3260 834 839 3257 836 3258 3248 846 3251 845 840 3254
 843 3253)
(848 3246 851 3245 3240 854 3243 853 863 3233 860 3234 3239 857
 3236 858)
(1008 3086 1011 3085 3080 1014 3083 1013 1023 3073 1020 3074
 3079 1017 3076 1018)
(3103 993 3100 994 999 3097 996 3098 3088 1006 3091 1005 1000
 3094 1003 3093)
(960 3134 963 3133 3128 966 3131 965 975 3121 972 3122 3127 969
 3124 970)
(3119 977 3116 978 983 3113 980 3114 3104 990 3107 989 984 3110
 987 3109)
(3199 897 3196 898 903 3193 900 3194 3184 910 3187 909 904 3190
 907 3189)
(912 3182 915 3181 3176 918 3179 917 927 3169 924 3170 3175 921
 3172 922)
(3151 945 3148 946 951 3145 948 3146 3136 958 3139 957 952 3142
 955 3141)
(928 3166 931 3165 3160 934 3163 933 943 3153 940 3154 3159 937
 3156 938))
((3583 513 3580 514 519 3577 516 3578 3568 526 3571 525 520 3574
 523 3573)
(528 3566 531 3565 3560 534 3563 533 543 3553 540 3554 3559 537
 3556 538)
(3535 561 3532 562 567 3529 564 3530 3520 574 3523 573 568 3526
 571 3525)
(544 3550 547 3549 3544 550 3547 549 559 3537 556 3538 3543 553
 3540 554)
(624 3470 627 3469 3464 630 3467 629 639 3457 636 3458 3463 633
 3460 634)
(3487 609 3484 610 615 3481 612 3482 3472 622 3475 621 616 3478
 619 3477)
(576 3518 579 3517 3512 582 3515 581 591 3505 588 3506 3511 585
 3508 586)
(3503 593 3500 594 599 3497 596 3498 3488 606 3491 605 600 3494
 603 3493)
(3343 753 3340 754 759 3337 756 3338 3328 766 3331 765 760 3334
 763 3333)
(736 3358 739 3357 3352 742 3355 741 751 3345 748 3346 3351 745
 3348 746)
(3391 705 3388 706 711 3385 708 3386 3376 718 3379 717 712 3382
```

```
  715 3381)
 (720 3374 723 3373 3368 726 3371 725 735 3361 732 3362 3367 729
  3364 730)
 (640 3454 643 3453 3448 646 3451 645 655 3441 652 3442 3447 649
  3444 650)
 (3439 657 3436 658 663 3433 660 3434 3424 670 3427 669 664 3430
  667 3429)
 (688 3406 691 3405 3400 694 3403 693 703 3393 700 3394 3399 697
  3396 698)
 (3423 673 3420 674 679 3417 676 3418 3408 686 3411 685 680 3414
  683 3413))
((2303 1793 2300 1794 1799 2297 1796 2298 2288 1806 2291 1805
  1800 2294 1803 2293)
 (1808 2286 1811 2285 2280 1814 2283 1813 1823 2273 1820 2274
  2279 1817 2276 1818)
 (2255 1841 2252 1842 1847 2249 1844 2250 2240 1854 2243 1853
  1848 2246 1851 2245)
 (1824 2270 1827 2269 2264 1830 2267 1829 1839 2257 1836 2258
  2263 1833 2260 1834)
 (1904 2190 1907 2189 2184 1910 2187 1909 1919 2177 1916 2178
  2183 1913 2180 1914)
 (2207 1889 2204 1890 1895 2201 1892 2202 2192 1902 2195 1901
  1896 2198 1899 2197)
 (1856 2238 1859 2237 2232 1862 2235 1861 1871 2225 1868 2226
  2231 1865 2228 1866)
 (2223 1873 2220 1874 1879 2217 1876 2218 2208 1886 2211 1885
  1880 2214 1883 2213)
 (2063 2033 2060 2034 2039 2057 2036 2058 2048 2046 2051 2045
  2040 2054 2043 2053)
 (2016 2078 2019 2077 2072 2022 2075 2021 2031 2065 2028 2066
  2071 2025 2068 2026)
 (2111 1985 2108 1986 1991 2105 1988 2106 2096 1998 2099 1997
  1992 2102 1995 2101)
 (2000 2094 2003 2093 2088 2006 2091 2005 2015 2081 2012 2082
  2087 2009 2084 2010)
 (1920 2174 1923 2173 2168 1926 2171 1925 1935 2161 1932 2162
  2167 1929 2164 1930)
 (2159 1937 2156 1938 1943 2153 1940 2154 2144 1950 2147 1949
  1944 2150 1947 2149)
 (1968 2126 1971 2125 2120 1974 2123 1973 1983 2113 1980 2114
  2119 1977 2116 1978)
 (2143 1953 2140 1954 1959 2137 1956 2138 2128 1966 2131 1965
  1960 2134 1963 2133))
((1536 2558 1539 2557 2552 1542 2555 1541 1551 2545 1548 2546
  2551 1545 2548 1546)
 (2543 1553 2540 1554 1559 2537 1556 2538 2528 1566 2531 1565
  1560 2534 1563 2533)
 (1584 2510 1587 2509 2504 1590 2507 1589 1599 2497 1596 2498
  2503 1593 2500 1594)
 (2527 1569 2524 1570 1575 2521 1572 2522 2512 1582 2515 1581
  1576 2518 1579 2517)
 (2447 1649 2444 1650 1655 2441 1652 2442 2432 1662 2435 1661
  1656 2438 1659 2437)
 (1632 2462 1635 2461 2456 1638 2459 1637 1647 2449 1644 2450
  2455 1641 2452 1642)
 (2495 1601 2492 1602 1607 2489 1604 2490 2480 1614 2483 1613
  1608 2486 1611 2485)
 (1616 2478 1619 2477 2472 1622 2475 1621 1631 2465 1628 2466
```

```
   2471 1625 2468 1626)
  (1776 2318 1779 2317 2312 1782 2315 1781 1791 2305 1788 2306
   2311 1785 2308 1786)
  (2335 1761 2332 1762 1767 2329 1764 2330 2320 1774 2323 1773
   1768 2326 1771 2325)
  (1728 2366 1731 2365 2360 1734 2363 1733 1743 2353 1740 2354
   2359 1737 2356 1738)
  (2351 1745 2348 1746 1751 2345 1748 2346 2336 1758 2339 1757
   1752 2342 1755 2341)
  (2431 1665 2428 1666 1671 2425 1668 2426 2416 1678 2419 1677
   1672 2422 1675 2421)
  (1680 2414 1683 2413 2408 1686 2411 1685 1695 2401 1692 2402
   2407 1689 2404 1690)
  (2383 1713 2380 1714 1719 2377 1716 2378 2368 1726 2371 1725
   1720 2374 1723 2373)
  (1696 2398 1699 2397 2392 1702 2395 1701 1711 2385 1708 2386
   2391 1705 2388 1706))
((3071 1025 3068 1026 1031 3065 1028 3066 3056 1038 3059 1037
   1032 3062 1035 3061)
  (1040 3054 1043 3053 3048 1046 3051 1045 1055 3041 1052 3042
   3047 1049 3044 1050)
  (3023 1073 3020 1074 1079 3017 1076 3018 3008 1086 3011 1085
   1080 3014 1083 3013)
  (1056 3038 1059 3037 3032 1062 3035 1061 1071 3025 1068 3026
   3031 1065 3028 1066)
  (1136 2958 1139 2957 2952 1142 2955 1141 1151 2945 1148 2946
   2951 1145 2948 1146)
  (2975 1121 2972 1122 1127 2969 1124 2970 2960 1134 2963 1133
   1128 2966 1131 2965)
  (1088 3006 1091 3005 3000 1094 3003 1093 1103 2993 1100 2994
   2999 1097 2996 1098)
  (2991 1105 2988 1106 1111 2985 1108 2986 2976 1118 2979 1117
   1112 2982 1115 2981)
  (2831 1265 2828 1266 1271 2825 1268 2826 2816 1278 2819 1277
   1272 2822 1275 2821)
  (1248 2846 1251 2845 2840 1254 2843 1253 1263 2833 1260 2834
   2839 1257 2836 1258)
  (2879 1217 2876 1218 1223 2873 1220 2874 2864 1230 2867 1229
   1224 2870 1227 2869)
  (1232 2862 1235 2861 2856 1238 2859 1237 1247 2849 1244 2850
   2855 1241 2852 1242)
  (1152 2942 1155 2941 2936 1158 2939 1157 1167 2929 1164 2930
   2935 1161 2932 1162)
  (2927 1169 2924 1170 1175 2921 1172 2922 2912 1182 2915 1181
   1176 2918 1179 2917)
  (1200 2894 1203 2893 2888 1206 2891 1205 1215 2881 1212 2882
   2887 1209 2884 1210)
  (2911 1185 2908 1186 1191 2905 1188 2906 2896 1198 2899 1197
   1192 2902 1195 2901))
((1280 2814 1283 2813 2808 1286 2811 1285 1295 2801 1292 2802
   2807 1289 2804 1290)
  (2799 1297 2796 1298 1303 2793 1300 2794 2784 1310 2787 1309
   1304 2790 1307 2789)
  (1328 2766 1331 2765 2760 1334 2763 1333 1343 2753 1340 2754
   2759 1337 2756 1338)
  (2783 1313 2780 1314 1319 2777 1316 2778 2768 1326 2771 1325
   1320 2774 1323 2773)
  (2703 1393 2700 1394 1399 2697 1396 2698 2688 1406 2691 1405
```

```
 1400 2694 1403 2693)
(1376 2718 1379 2717 2712 1382 2715 1381 1391 2705 1388 2706
 2711 1385 2708 1386)
(2751 1345 2748 1346 1351 2745 1348 2746 2736 1358 2739 1357
 1352 2742 1355 2741)
(1360 2734 1363 2733 2728 1366 2731 1365 1375 2721 1372 2722
 2727 1369 2724 1370)
(1520 2574 1523 2573 2568 1526 2571 1525 1535 2561 1532 2562
 2567 1529 2564 1530)
(2591 1505 2588 1506 1511 2585 1508 2586 2576 1518 2579 1517
 1512 2582 1515 2581)
(1472 2622 1475 2621 2616 1478 2619 1477 1487 2609 1484 2610
 2615 1481 2612 1482)
(2607 1489 2604 1490 1495 2601 1492 2602 2592 1502 2595 1501
 1496 2598 1499 2597)
(2687 1409 2684 1410 1415 2681 1412 2682 2672 1422 2675 1421
 1416 2678 1419 2677)
(1424 2670 1427 2669 2664 1430 2667 1429 1439 2657 1436 2658
 2663 1433 2660 1434)
(2639 1457 2636 1458 1463 2633 1460 2634 2624 1470 2627 1469
 1464 2630 1467 2629)
(1440 2654 1443 2653 2648 1446 2651 1445 1455 2641 1452 2642
 2647 1449 2644 1450))
((3840 254 3843 253 248 3846 251 3845 3855 241 3852 242 247 3849
 244 3850)
(239 3857 236 3858 3863 233 3860 234 224 3870 227 3869 3864 230
 3867 229)
(3888 206 3891 205 200 3894 203 3893 3903 193 3900 194 199 3897
 196 3898)
(223 3873 220 3874 3879 217 3876 218 208 3886 211 3885 3880 214
 3883 213)
(143 3953 140 3954 3959 137 3956 138 128 3966 131 3965 3960 134
 3963 133)
(3936 158 3939 157 152 3942 155 3941 3951 145 3948 146 151 3945
 148 3946)
(191 3905 188 3906 3911 185 3908 186 176 3918 179 3917 3912 182
 3915 181)
(3920 174 3923 173 168 3926 171 3925 3935 161 3932 162 167 3929
 164 3930)
(4080 14 4083 13 8 4086 11 4085 4095 1 4092 2 7 4089 4 4090)
(31 4065 28 4066 4071 25 4068 26 16 4078 19 4077 4072 22 4075
 21)
(4032 62 4035 61 56 4038 59 4037 4047 49 4044 50 55 4041 52
 4042)
(47 4049 44 4050 4055 41 4052 42 32 4062 35 4061 4056 38 4059
 37)
(127 3969 124 3970 3975 121 3972 122 112 3982 115 3981 3976 118
 3979 117)
(3984 110 3987 109 104 3990 107 3989 3999 97 3996 98 103 3993
 100 3994)
(79 4017 76 4018 4023 73 4020 74 64 4030 67 4029 4024 70 4027
 69)
(4000 94 4003 93 88 4006 91 4005 4015 81 4012 82 87 4009 84
 4010))
((511 3585 508 3586 3591 505 3588 506 496 3598 499 3597 3592 502
 3595 501)
(3600 494 3603 493 488 3606 491 3605 3615 481 3612 482 487 3609
 484 3610)
```

(463 3633 460 3634 3639 457 3636 458 448 3646 451 3645 3640 454
3643 453)
(3616 478 3619 477 472 3622 475 3621 3631 465 3628 466 471 3625
468 3626)
(3696 398 3699 397 392 3702 395 3701 3711 385 3708 386 391 3705
388 3706)
(415 3681 412 3682 3687 409 3684 410 400 3694 403 3693 3688 406
3691 405)
(3648 446 3651 445 440 3654 443 3653 3663 433 3660 434 439 3657
436 3658)
(431 3665 428 3666 3671 425 3668 426 416 3678 419 3677 3672 422
3675 421)
(271 3825 268 3826 3831 265 3828 266 256 3838 259 3837 3832 262
3835 261)
(3808 286 3811 285 280 3814 283 3813 3823 273 3820 274 279 3817
276 3818)
(319 3777 316 3778 3783 313 3780 314 304 3790 307 3789 3784 310
3787 309)
(3792 302 3795 301 296 3798 299 3797 3807 289 3804 290 295 3801
292 3802)
(3712 382 3715 381 376 3718 379 3717 3727 369 3724 370 375 3721
372 3722)
(367 3729 364 3730 3735 361 3732 362 352 3742 355 3741 3736 358
3739 357)
(3760 334 3763 333 328 3766 331 3765 3775 321 3772 322 327 3769
324 3770)
(351 3745 348 3746 3751 345 3748 346 336 3758 339 3757 3752 342
3755 341))
((3072 1022 3075 1021 1016 3078 1019 3077 3087 1009 3084 1010
1015 3081 1012 3082)
(1007 3089 1004 3090 3095 1001 3092 1002 992 3102 995 3101 3096
998 3099 997)
(3120 974 3123 973 968 3126 971 3125 3135 961 3132 962 967 3129
964 3130)
(991 3105 988 3106 3111 985 3108 986 976 3118 979 3117 3112 982
3115 981)
(911 3185 908 3186 3191 905 3188 906 896 3198 899 3197 3192 902
3195 901)
(3168 926 3171 925 920 3174 923 3173 3183 913 3180 914 919 3177
916 3178)
(959 3137 956 3138 3143 953 3140 954 944 3150 947 3149 3144 950
3147 949)
(3152 942 3155 941 936 3158 939 3157 3167 929 3164 930 935 3161
932 3162)
(3312 782 3315 781 776 3318 779 3317 3327 769 3324 770 775 3321
772 3322)
(799 3297 796 3298 3303 793 3300 794 784 3310 787 3309 3304 790
3307 789)
(3264 830 3267 829 824 3270 827 3269 3279 817 3276 818 823 3273
820 3274)
(815 3281 812 3282 3287 809 3284 810 800 3294 803 3293 3288 806
3291 805)
(895 3201 892 3202 3207 889 3204 890 880 3214 883 3213 3208 886
3211 885)
(3216 878 3219 877 872 3222 875 3221 3231 865 3228 866 871 3225
868 3226)
(847 3249 844 3250 3255 841 3252 842 832 3262 835 3261 3256 838
3259 837)

```
  (3232 862 3235 861 856 3238 859 3237 3247 849 3244 850 855 3241
   852 3242))
((767 3329 764 3330 3335 761 3332 762 752 3342 755 3341 3336 758
  3339 757)
 (3344 750 3347 749 744 3350 747 3349 3359 737 3356 738 743 3353
  740 3354)
 (719 3377 716 3378 3383 713 3380 714 704 3390 707 3389 3384 710
  3387 709)
 (3360 734 3363 733 728 3366 731 3365 3375 721 3372 722 727 3369
  724 3370)
 (3440 654 3443 653 648 3446 651 3445 3455 641 3452 642 647 3449
  644 3450)
 (671 3425 668 3426 3431 665 3428 666 656 3438 659 3437 3432 662
  3435 661)
 (3392 702 3395 701 696 3398 699 3397 3407 689 3404 690 695 3401
  692 3402)
 (687 3409 684 3410 3415 681 3412 682 672 3422 675 3421 3416 678
  3419 677)
 (527 3569 524 3570 3575 521 3572 522 512 3582 515 3581 3576 518
  3579 517)
 (3552 542 3555 541 536 3558 539 3557 3567 529 3564 530 535 3561
  532 3562)
 (575 3521 572 3522 3527 569 3524 570 560 3534 563 3533 3528 566
  3531 565)
 (3536 558 3539 557 552 3542 555 3541 3551 545 3548 546 551 3545
  548 3546)
 (3456 638 3459 637 632 3462 635 3461 3471 625 3468 626 631 3465
  628 3466)
 (623 3473 620 3474 3479 617 3476 618 608 3486 611 3485 3480 614
  3483 613)
 (3504 590 3507 589 584 3510 587 3509 3519 577 3516 578 583 3513
  580 3514)
 (607 3489 604 3490 3495 601 3492 602 592 3502 595 3501 3496 598
  3499 597))
((2047 2049 2044 2050 2055 2041 2052 2042 2032 2062 2035 2061
  2056 2038 2059 2037)
 (2064 2030 2067 2029 2024 2070 2027 2069 2079 2017 2076 2018
  2023 2073 2020 2074)
 (1999 2097 1996 2098 2103 1993 2100 1994 1984 2110 1987 2109
  2104 1990 2107 1989)
 (2080 2014 2083 2013 2008 2086 2011 2085 2095 2001 2092 2002
  2007 2089 2004 2090)
 (2160 1934 2163 1933 1928 2166 1931 2165 2175 1921 2172 1922
  1927 2169 1924 2170)
 (1951 2145 1948 2146 2151 1945 2148 1946 1936 2158 1939 2157
  2152 1942 2155 1941)
 (2112 1982 2115 1981 1976 2118 1979 2117 2127 1969 2124 1970
  1975 2121 1972 2122)
 (1967 2129 1964 2130 2135 1961 2132 1962 1952 2142 1955 2141
  2136 1958 2139 1957)
 (1807 2289 1804 2290 2295 1801 2292 1802 1792 2302 1795 2301
  2296 1798 2299 1797)
 (2272 1822 2275 1821 1816 2278 1819 2277 2287 1809 2284 1810
  1815 2281 1812 2282)
 (1855 2241 1852 2242 2247 1849 2244 1850 1840 2254 1843 2253
  2248 1846 2251 1845)
 (2256 1838 2259 1837 1832 2262 1835 2261 2271 1825 2268 1826
  1831 2265 1828 2266)
```

```
   (2176 1918 2179 1917 1912 2182 1915 2181 2191 1905 2188 1906
    1911 2185 1908 2186)
   (1903 2193 1900 2194 2199 1897 2196 1898 1888 2206 1891 2205
    2200 1894 2203 1893)
   (2224 1870 2227 1869 1864 2230 1867 2229 2239 1857 2236 1858
    1863 2233 1860 2234)
   (1887 2209 1884 2210 2215 1881 2212 1882 1872 2222 1875 2221
    2216 1878 2219 1877))
 ((2304 1790 2307 1789 1784 2310 1787 2309 2319 1777 2316 1778
    1783 2313 1780 2314)
   (1775 2321 1772 2322 2327 1769 2324 1770 1760 2334 1763 2333
    2328 1766 2331 1765)
   (2352 1742 2355 1741 1736 2358 1739 2357 2367 1729 2364 1730
    1735 2361 1732 2362)
   (1759 2337 1756 2338 2343 1753 2340 1754 1744 2350 1747 2349
    2344 1750 2347 1749)
   (1679 2417 1676 2418 2423 1673 2420 1674 1664 2430 1667 2429
    2424 1670 2427 1669)
   (2400 1694 2403 1693 1688 2406 1691 2405 2415 1681 2412 1682
    1687 2409 1684 2410)
   (1727 2369 1724 2370 2375 1721 2372 1722 1712 2382 1715 2381
    2376 1718 2379 1717)
   (2384 1710 2387 1709 1704 2390 1707 2389 2399 1697 2396 1698
    1703 2393 1700 2394)
   (2544 1550 2547 1549 1544 2550 1547 2549 2559 1537 2556 1538
    1543 2553 1540 2554)
   (1567 2529 1564 2530 2535 1561 2532 1562 1552 2542 1555 2541
    2536 1558 2539 1557)
   (2496 1598 2499 1597 1592 2502 1595 2501 2511 1585 2508 1586
    1591 2505 1588 2506)
   (1583 2513 1580 2514 2519 1577 2516 1578 1568 2526 1571 2525
    2520 1574 2523 1573)
   (1663 2433 1660 2434 2439 1657 2436 1658 1648 2446 1651 2445
    2440 1654 2443 1653)
   (2448 1646 2451 1645 1640 2454 1643 2453 2463 1633 2460 1634
    1639 2457 1636 2458)
   (1615 2481 1612 2482 2487 1609 2484 1610 1600 2494 1603 2493
    2488 1606 2491 1605)
   (2464 1630 2467 1629 1624 2470 1627 2469 2479 1617 2476 1618
    1623 2473 1620 2474))
 ((1279 2817 1276 2818 2823 1273 2820 1274 1264 2830 1267 2829
    2824 1270 2827 1269)
   (2832 1262 2835 1261 1256 2838 1259 2837 2847 1249 2844 1250
    1255 2841 1252 2842)
   (1231 2865 1228 2866 2871 1225 2868 1226 1216 2878 1219 2877
    2872 1222 2875 1221)
   (2848 1246 2851 1245 1240 2854 1243 2853 2863 1233 2860 1234
    1239 2857 1236 2858)
   (2928 1166 2931 1165 1160 2934 1163 2933 2943 1153 2940 1154
    1159 2937 1156 2938)
   (1183 2913 1180 2914 2919 1177 2916 1178 1168 2926 1171 2925
    2920 1174 2923 1173)
   (2880 1214 2883 1213 1208 2886 1211 2885 2895 1201 2892 1202
    1207 2889 1204 2890)
   (1199 2897 1196 2898 2903 1193 2900 1194 1184 2910 1187 2909
    2904 1190 2907 1189)
   (1039 3057 1036 3058 3063 1033 3060 1034 1024 3070 1027 3069
    3064 1030 3067 1029)
```

```
(3040 1054 3043 1053 1048 3046 1051 3045 3055 1041 3052 1042
 1047 3049 1044 3050)
(1087 3009 1084 3010 3015 1081 3012 1082 1072 3022 1075 3021
 3016 1078 3019 1077)
(3024 1070 3027 1069 1064 3030 1067 3029 3039 1057 3036 1058
 1063 3033 1060 3034)
(2944 1150 2947 1149 1144 2950 1147 2949 2959 1137 2956 1138
 1143 2953 1140 2954)
(1135 2961 1132 2962 2967 1129 2964 1130 1120 2974 1123 2973
 2968 1126 2971 1125)
(2992 1102 2995 1101 1096 2998 1099 2997 3007 1089 3004 1090
 1095 3001 1092 3002)
(1119 2977 1116 2978 2983 1113 2980 1114 1104 2990 1107 2989
 2984 1110 2987 1109))
((2560 1534 2563 1533 1528 2566 1531 2565 2575 1521 2572 1522
 1527 2569 1524 2570)
(1519 2577 1516 2578 2583 1513 2580 1514 1504 2590 1507 2589
 2584 1510 2587 1509)
(2608 1486 2611 1485 1480 2614 1483 2613 2623 1473 2620 1474
 1479 2617 1476 2618)
(1503 2593 1500 2594 2599 1497 2596 1498 1488 2606 1491 2605
 2600 1494 2603 1493)
(1423 2673 1420 2674 2679 1417 2676 1418 1408 2686 1411 2685
 2680 1414 2683 1413)
(2656 1438 2659 1437 1432 2662 1435 2661 2671 1425 2668 1426
 1431 2665 1428 2666)
(1471 2625 1468 2626 2631 1465 2628 1466 1456 2638 1459 2637
 2632 1462 2635 1461)
(2640 1454 2643 1453 1448 2646 1451 2645 2655 1441 2652 1442
 1447 2649 1444 2650)
(2800 1294 2803 1293 1288 2806 1291 2805 2815 1281 2812 1282
 1287 2809 1284 2810)
(1311 2785 1308 2786 2791 1305 2788 1306 1296 2798 1299 2797
 2792 1302 2795 1301)
(2752 1342 2755 1341 1336 2758 1339 2757 2767 1329 2764 1330
 1335 2761 1332 2762)
(1327 2769 1324 2770 2775 1321 2772 1322 1312 2782 1315 2781
 2776 1318 2779 1317)
(1407 2689 1404 2690 2695 1401 2692 1402 1392 2702 1395 2701
 2696 1398 2699 1397)
(2704 1390 2707 1389 1384 2710 1387 2709 2719 1377 2716 1378
 1383 2713 1380 2714)
(1359 2737 1356 2738 2743 1353 2740 1354 1344 2750 1347 2749
 2744 1350 2747 1349)
(2720 1374 2723 1373 1368 2726 1371 2725 2735 1361 2732 1362
 1367 2729 1364 2730)))
```

In odd dimension the application of the algorithm isn't straightforward. In even dimension even numbers of bits map to even numbers of bits, odd numbers of bits map to odd numbers of bits, and the mapping is one to one. In odd dimension even goes to even but odd goes to even as well, so half the values disappear:
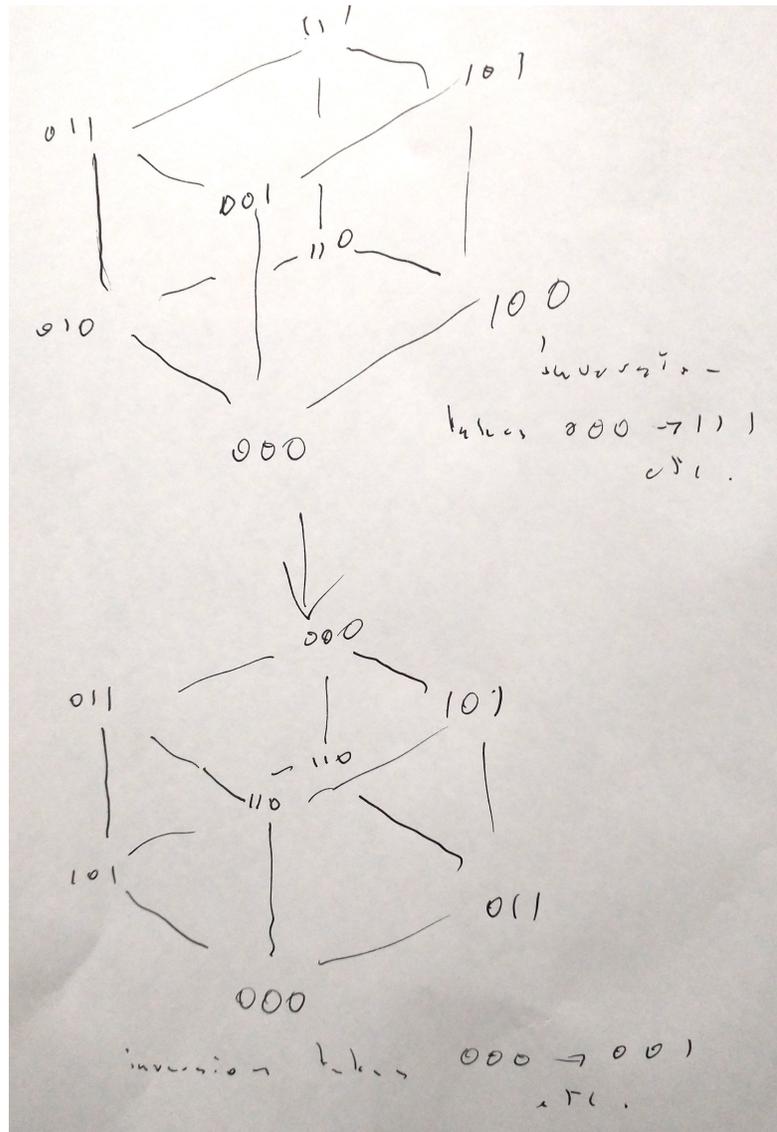
But sums around faces are invariant under bit complementation, so there is a looking-glass cube with the other half of the values, and you

can construct an alternating infinite array as indicated here in dimension 3:
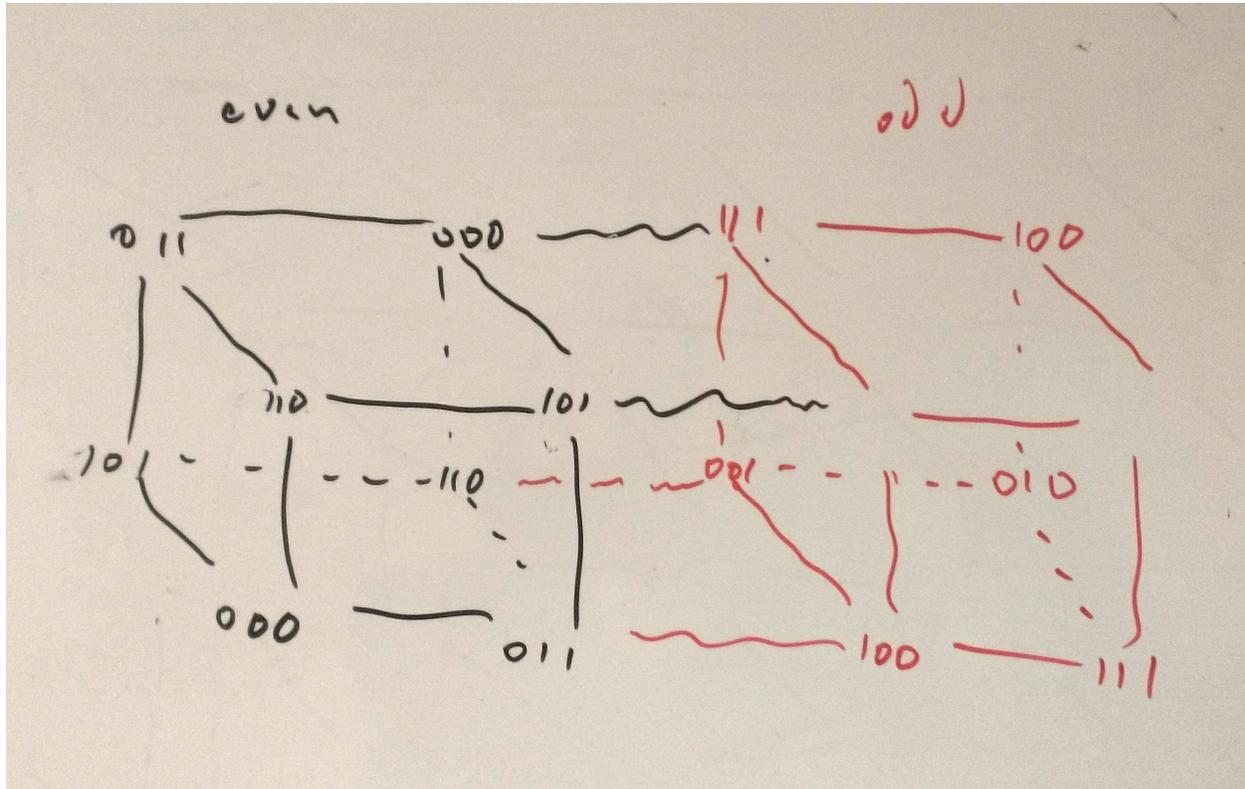
repeated in every direction. Then within the even and odd cubes the adjacency rule holds, the sum around every square is, again, the sort-of-magic number (here 14), and

$$v(i + 4, j, k) = v(i, j + 4, k) = v(i, j, k + 4) = v(i, j, k)$$

$$v(4 - i, j, k) = etc. = complement(v(i, j, k))$$

ensuring that the sum of any four consecutive values is, again, the sort-of-magic number.

The magic carpet

It is known[185] that the Penrose tiling is a projection of a section of a five-dimensional cubic lattice onto the plane. In analogy with the cubic lattice in three dimensions, the vertices can be labeled with the integers 0, …, 31 in such a way that the sum around any square is 62, and accordingly the tiling can be so labeled. In part it looks like this:

---

[185] N. G. de Bruijn, "Algebraic theory of Penrose's non-periodic tilings of the plane [I, II]", *Indagationes Mathematicae* **84**, 38–52, 53–66 (1981).
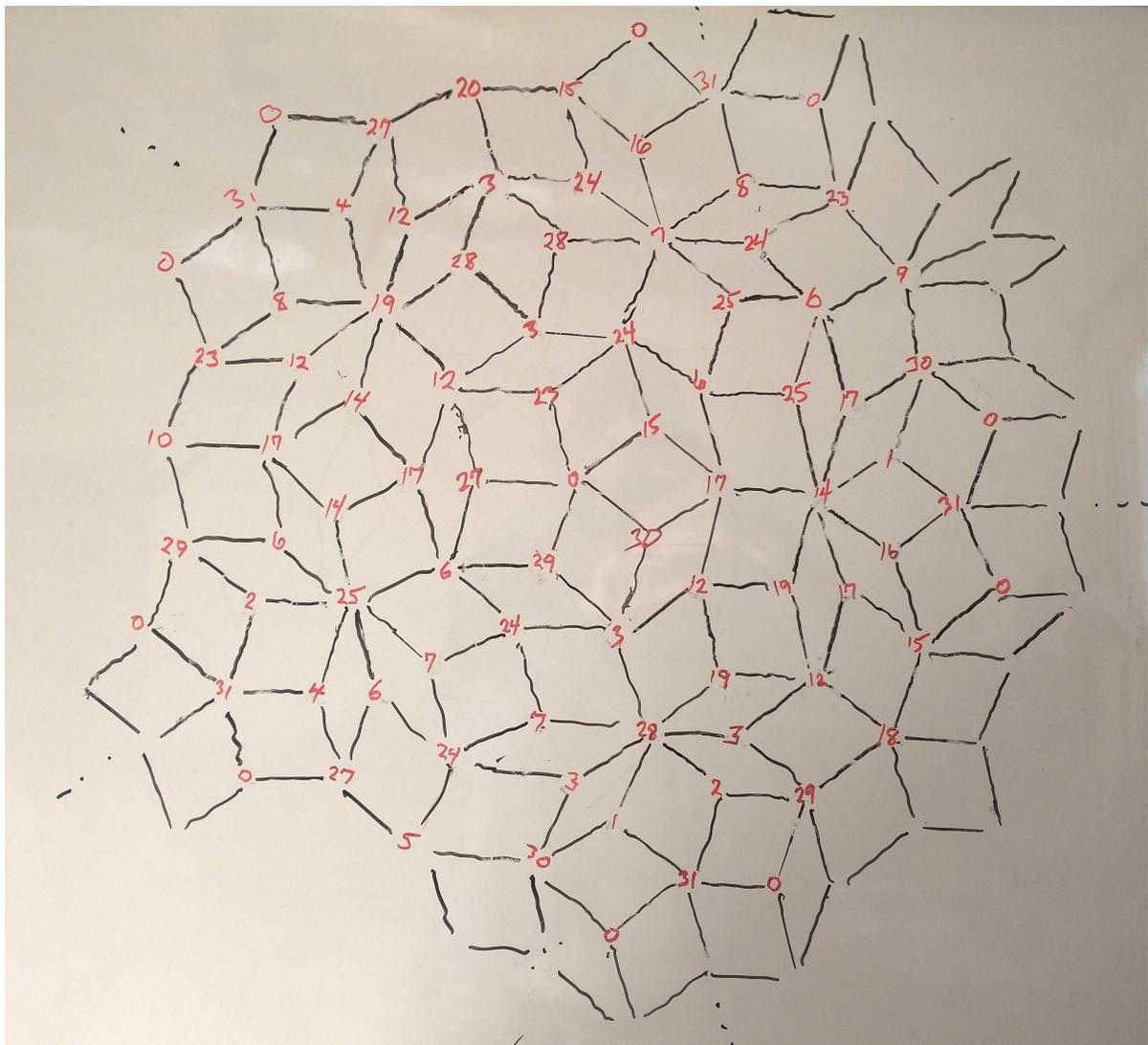
This has the following properties:

— There is a central axis around which the graph has a fivefold symmetry, i.e. rotating it any multiple of 72 degrees carries it into itself.

— Similarly there is a reflection symmetry about the horizontal axis, the top and bottom are mirror images.

(These properties hold only for this particular representation of the tiling, which corresponds to a projection onto a plane passing through the origin in five-space.)

— The center is the vertex labelled with 0.

— The horizontal axis passes through the vertices labelled 27, 0, 17, etc.

— The labels of adjacent vertices, regarded as bit-vectors, are related either by the requirement that they agree in exactly one place, as e.g. from the origin

0 = (0 0 0 0 0) <—> (0 1 1 1 1) = 15, (1 0 1 1 1) = 23, (1 1 0 1 1) = 27, (1 1 1 0 1) = 29, (1 1 1 1 0) = 30

or (when crossing a boundary between adjacent five-cubes) as complements, e.g.

17 = (1 0 0 0 1) <—> 14 = (0 1 1 1 0)
24 = (1 1 0 0 0) <—> 7 = (0 0 1 1 1)

(etc.)

— When the graph is rotated counterclockwise 72 degrees, the bit representation of the label does a cyclic shift right, e.g. around the center decahedron

17 = (1 0 0 0 1) —> 24 = (1 1 0 0 0) —> 12 = (0 1 1 0 0)
—> 6 = (0 0 1 1 0) —> 3 = (0 0 0 1 1) —> 17

— Reflection about the horizontal axis carries a label bitwise into its inversion, thus

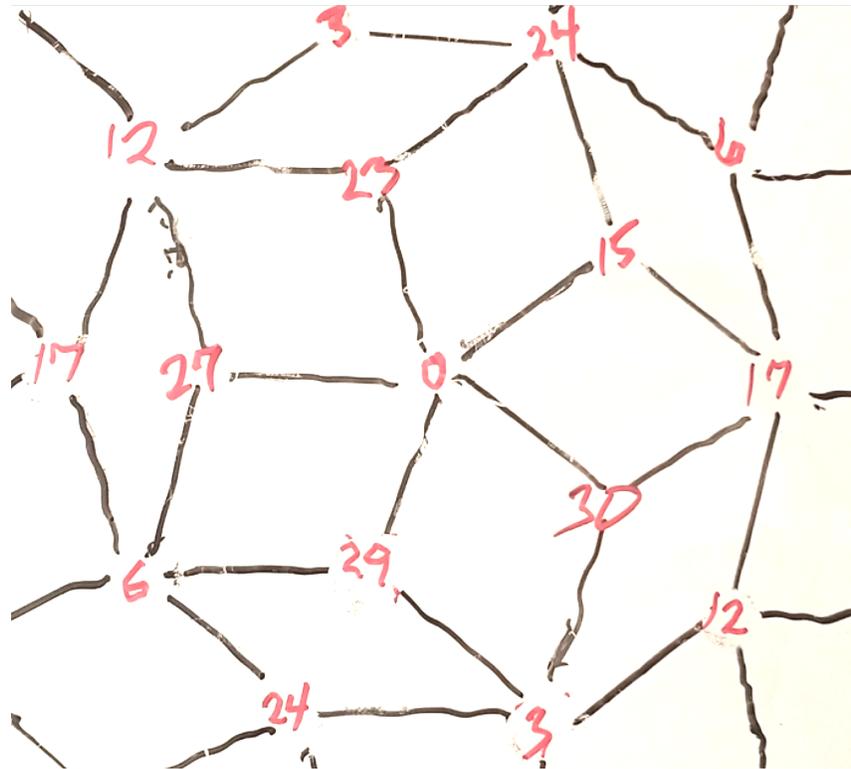[a] The numbers on the axis are bitwise palindromic, e.g. left to right

10 = (0 1 0 1 0), 17 = (1 0 0 0 1), 27 =(1 1 0 1 1), 0 = (0 0 0 0 0), 14 = (0 1 1 1 0), 31 = (1 1 1 1 1), and

[b] proceeding right from the center

15 = (0 1 1 1 1) <—> 30 = (1 1 1 1 0)
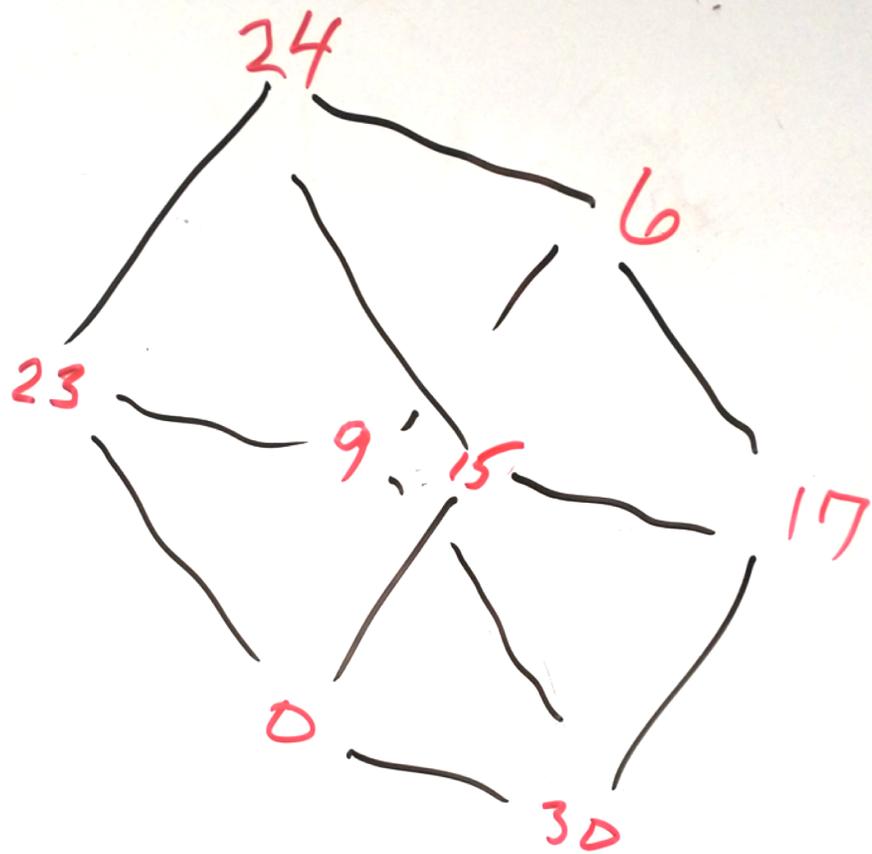6 = (0 0 1 1 0) <—> 12 = (0 1 1 0 0)
25 = (1 1 0 0 1) <—> 19 = (1 0 0 1 1)

etc.

— Moreover the hidden values of the projected 3-cubes can be inferred as follows: looking, e.g., at the central decahedron

the sum around any square must be 62, and thus the missing values are 9, 20, 10, 5, and 18. (These obey the cyclic shift rule.)

The results can be easily extended to tilings based on projections of higher-dimensional cubic lattices onto the plane (or lower-dimensional surfaces generally), but leave that for the next pandemic.