Reflections on Algorithmicity and Consciousness

Why should anyone believe that conscious awareness is solely the result of the combinatorial complexity of the brain - the idea that the brain is some kind of computer performing calculations? What is the connection between consciousness and calculation: could it be transformation of quantity into quality? The claim that the one is reducible to the other seems unconvincing—as unlike as chalk and cheese! In his book *The Emperor’s New Mind*, Roger Penrose at least attempts to compare like with (something like) like: the enigma of consciousness with the progress of physics.

In any case, supposing that one were to become convinced on “scientific” grounds that there is no option but to regard consciousness as, ultimately, a matter of calculation, how would one go on to formulate a cogent account of the evident and irreducible subjectivity of consciousness: the fact that a conscious being has access only to its own—and to no other—consciousness?

I have long been convinced that whether any "objective" account of consciousness will fail to do justice to its uniquely subjective character. In the words of Hermann Weyl, such an account

*may be objectively adequate, but it is shattered by the desperate cry of Judas: Why did I have to be Judas! The impossibility of an objective formulation to this question strikes home, and no answer in the form of an objective insight can be given. Knowledge cannot bring the light that is I into coincidence with the murky, erring human being that is cast out into an individual fate.*

Both Penrose and the AI people appear to share the belief that such an account is possible—and presumably necessary! Indeed the latter, at least in Penrose’s portrayal, claim already to be in possession of it. Penrose himself is skeptical of this claim, but still cherishes the hope that a satisfactory “objective” account of consciousness will
somehow emerge from future developments in physics. In any event, the fact that both parties believe such an account to be formulable in principle sharply distinguishes them from philosophers who take consciousness seriously, e.g. Descartes, Husserl who regarded consciousness as an ineluctable “given”.

I believe that the inescapably “external” or “extensional” character of any “objective” account of consciousness renders it incapable of addressing the irreducibly subjective or “intensional” phenomenon of actually being conscious - just as mathematical theories of time necessarily fail to capture the phenomenon of being in time. Scientific theories are not intended to capture the essence of the subjective - to that extent scientific knowledge is limited.

Whether thought and objective reality are, in the last analysis, algorithmic in character is a question of some subtlety. The Gödel incompleteness theorem establishes the “nonalgorithmic” character of the truth of a particular arithmetical statement $A$ at the level of a specified formalism. On the other hand, the truth of $A$ can be established within a strengthened formalism, thereby reinstating “algorithmicity” of the truth of $A$. But now, inevitably, new “non-algorithmically verifiable” statements appear at the level of the strengthened formalism, and the whole business starts all over again. The point would seem to be that, while “algorithmicity” is manifested locally at higher and higher levels of description and is likely to be an essential constituent of such descriptions, nevertheless diagonal arguments show that algorithmicity cannot be imposed globally, that is, uniformly and simultaneously, on all levels of description.

The apparent ubiquity of algorithmicity, or, more generally, symbolizability, should not mislead us into believing that the world is itself an algorithm or a symbol.
Reflections on the oneiric self, time, and all that.

Cast a cold eye
On life, on death
Horseman, pass by!

___ Yeats

To awaken is to collapse the wave function of dreaming

----- Bemnhall

Reflecting on the differences between the two selves each of us possesses—let me term them the conscious self—the self of which one is aware when one is awake—and the oneiric self, the self that emerges, if murkily, when one is asleep, and whose only experiences are dreams—it is quite striking that, while the oneiric self sees and hears, it lacks certain senses, for example (in my case, at least, the olfactory and the tactile, with the exception of “externally induced” sensations such as that arising from bladder pressure). But this very fact confers certain “advantages” on the oneiric self. For instance, the oneiric self’s lack of tactile sensibility reduces what the conscious self would feel as actual pain to a kind of diffuse anxiety. In contrast with the conscious self, the oneiric self is characterized by its lack of volition, its passivity, its incapability of understanding tenses, its inability to control the direction of its temporal tangent vector. From the point of view of the conscious self, it resembles the past (conscious) self, over which the conscious self has no control. On the other hand the essential ephemerality of the oneiric self enables it to avoid the boredom, the meaningless repetition, that burdens the conscious self: while my oneiric self, at any rate, is often subject to anxiety, it never experiences boredom—or at least, my conscious self, in the effort of recollection, has filtered that out! The conscious self, when it awakes, recalls its oneiric self as a kind of
fleeting version of its past conscious self, as a kind of automaton, a version of itself deprived of the volition and self-consciousness that the conscious self ineluctably attributes to itself. So the conscious self has something denied to its oneiric equivalent: however bored, however anxious, however tormented, the conscious self always has the potential escape into sleep, that is, into an escape from full consciousness, yet at the same time accompanied by the hope, the conviction even, that consciousness will soon be fully restored —thereby, in recollection, liberating the oneiric self from its fleeting existence - an existence entirely dependent on the reawakening, the re-emergence, of the conscious self into objective time. (This is, in turn, dependent on the continued functioning of that conscious self’s material brain, which disrupts the continuity. In any case, as Shakespeare has observed: “life’s time’s fool, and time must have a stop”). Nevertheless, the oneiric self has another advantage over the conscious self, namely, the plasticity of the dream-world within which it dwells, the world of pure willless imagination where natural law is suspended, and also through the idea suggested by stratified dreams that there is an inscrutable relationship between mind and reality—or at least what the mind takes to be reality—which is simply not accessible within the waking state.

While the oneiric self may experience what the conscious self would regard as quite startling deviations from natural law, distortions of the familiar — all those defiances of gravity, etc., that make (the recollection of) dreaming so delightful, and on occasion, in the form of nightmares, so harrowing— the oneiric self seemingly cannot escape the spatiotemporal categories to which its conscious counterpart is subject. For example, it still experiences linear time, even if with discontinuities (the “jump-cuts” of dreams), it does not travel into the past, it is never in two places at once, etc. In fact the (or at least my) oneiric self is quite conventional, a mere distortion of my conscious self, except for the fact that only it is the subject of stratified dreams—leading to the idea, already mentioned, that there may be a hidden arrangement to the world which is not immediately accessible to the senses, in which, perhaps, the same spatiotemporal (basic
categorical) laws prevail but within which another dimension beyond the temporal-linear also figures. Pure SF of course—but I can only blame my oneiric self for these fantasies!

The content of the sensations or experiences of the oneiric self resides in the recollections of the conscious self upon awakening from a dream. So if in this sentence we replace “oneiric” by “past”, “conscious” by “present” and “on awakening from a dream” by “reflecting on the past” we obtain the assertion: the content of the sensations or experiences of the past self resides in the recollections of the present self reflecting on the past. This shows that the oneiric self bears a certain resemblance to the past self. But of course there is a crucial difference: from the standpoint of the present self, the past self is merely the present self as it was in the past, with no essential change in the meaning of “self”. To put it in terms of a proportion,

\[
\text{Present self} : \text{Past self} = \text{Present} : \text{Past.}
\]

The cancellability of the “self” term here indicates that it is an invariant, a substrate, the terms “present” and “past” then being no more than temporal modifiers. Now entertain the corresponding proportion

\[
\text{Conscious self} : \text{Oneiric self} = \text{Conscious} : \text{Oneiric}
\]

What can be made of this? Is the sense of the term “self” here the same as in the above proportion? Clearly there is a difference between the two: the present and the past can be interpreted objectively, outside the mental states of the subject, so that the first proportion expresses a correlation between the subjective and the objective. The second proportion, on the other hand, expresses a correlation between the subjective and the metaphysical. The cancellation of the term “self” here leads to genuinely new categories:
the dreamt and the actual—transcending the linearity of objective time and suggesting possibilities even, so to speak, “undreamt”.

Can the past self be viewed by the present self as an “idiot” with the same irritation that the present self can be so regarded by itself? No, because time has honed down those irritations and pains experienced by the (recollected) past self to the point at which the present self can regard them with equanimity.

**The Tangent Space of Dreams**

*From the remorseless twist, the grip, of the actual*

* A dream, or a nightmare, is the ego’s escape.
* The ego set free to follow its own bent.
* Dreams are tangent vectors of a sort.
* Only not straight
* But still curved, still twisted.

The retreat into sleep, and thence to dreams, from Morpheus to Hypnos, should overcome the reverses encountered in actual life (“knitting up the ravelled sleeve of care”), even if the oneiric self reproduces, in an immediate yet curiously reduced (but not colourless) way, all the mental pain experienced by its correlated conscious self. After all, the oneiric self resembles the conscious self to the extent that both experience anxiety, fear, pleasure, and also understand language, etc. What the oneiric self lacks is not consciousness, but *self-consciousness,* —better, the awareness of its own existence—the grip of the idea that it is *actual,* it genuinely *exists.* The oneiric self is a kind of abstraction, a mere sketch of the actual self. Lacking a body, it has only the bare notion of such, and so is, in particular, rendered immune to pain. The price paid for this mercy is the loss of *will,* the directed volition, the *conatus* of the conscious self. Accordingly it
might be said that the oneiric self dwells in a kind of tangent space to the manifold of the conscious self. Indeed, on awakening, the conscious self, in recollecting a dream, is often struck by the idea that its oneiric self is a kind of “first approximation” to its own nature - the first couple of terms of its Taylor expansion, so to speak - perhaps lacking only that self-consciousness - embodied in the higher-order terms of the Taylor series-which somehow provide the ultimate confirmation of one’s waking existence.

In its efforts to control the future, the conscious self strives to bend the outer world to its will, and so necessarily lives in a curved, even twisted space. The oneiric self, by contrast, rushes off at a tangent. Just as the tangent bundle is a linearized presentation of a curved manifold, the oneiric self is a simplified, graphic counterpart to the conscious self. But, unlike its mathematical correlate, the oneiric self retains some of the "curvature" of the conscious self.

Sceptical philosophers have denied the existence of the self. Hume famously reported that when he examined his own thought processes he turned up a mere flux of impressions, with nothing like a “self” binding these impressions together. I do not doubt that this is a faithful, objective report of his own subjective investigations. My own self-analysis has led to different conclusions. When, “to the sessions of sweet silent thought, I summon up remembrance of things past”, it seems clear to me that it is a grasped entity called I – my self - that is doing the summoning. What the self is summoning through memory is its previous mental states – both in subjective and objective time. That being the case, from a mathematical point of view, the “self” might be thought of as a map of some kind, denote it Self. The domain of Self should be time, taken in the subjective sense. The codomain of Self will presumably be the mental states of the subject, as these vary in subjective time. The recognition that one actually has a self then amounts to acknowledging the objective existence of the correlation between the two subjectivities of time and mental states correlated by the map Self. Should the codomain of Self include just conscious mental states or should it also embrace
unconscious, or less than fully conscious mental states, for instance oneiric states? I am not sure.

**Reflections on Consciousness and Subjectivity**

*Consciousness is a form of temporal embroidery which only the embroiderer can see. It is a display visible only to the displayer.*

- The great enigma: how can the brain, as an objectively existing entity, engender the subjectivity of the “I”? Put this question aside for the moment - as Husserl would say, apply the *epoché*. This amounts to starting with what is immediately given. And what is immediately given to me? Just my own awareness and its content, sensations etc. And that very immediacy further requires that I say that my access to my awareness is given in the *present*, the “*now*”. While (my) awareness is *necessarily in the* present, it can, through a kind of stretching of the (specious) present, be directed at various objects such as sensations, appearances etc. These objects vary with the time that they are presented to awareness. Awareness can also, by the application of a certain mental discipline, be directed at *itself*. From this emerges a new object of awareness, one that in its essence does not vary with the moment at which it is grasped, namely, the awareness of *subjective temporal flow* (STF), Husserl’s *time-consciousness*, the experience of a flux or flowing or changing present in its purest sense. The flow is itself an embodiment of change, but its intrinsic form is unchanging. But, while pure in its fluidity, awareness of the temporal flow reveals it to possess an additional quality, namely, *one-dimensionality*, and hence, like a river, a *direction* in which it flows. This directed flow is described as *from past to future through the present, the “now”*. The content of the *present* is what I am *immediately aware of*, immediately “presented”. That content has an intimacy which is denied to the content of the *past* which is what I summon up through *memory*- through those “sessions of
sweet silent thought” - which are presented indirectly to my awareness, as representations of contents that, so to speak, were part of the present, that is, contents of which I have been immediately aware. Note that these representations of past content are themselves actually “present” insofar as I am immediately aware of them: it is their content which is not experienced directly. Thus in summoning up the past awareness divides its content into two components: an immediately gasped representation of the contents of the past as remembered, together with an awareness that what is immediately grasped is such a representation. Finally, the content of the future is created by imagination, employed in such a way as to present to awareness, in an indirect way, configurations which could become part of the content of the present, that is, configurations of which I could become immediately aware.

When awareness is directed at itself, many different qualities, or objects, may emerge. For example, if I see a green leaf, I am aware of its greenness, and if I direct my awareness to my awareness of the greenness, in doing so I become aware of a particular quality of my awareness, namely that it is directed towards a green leaf. Similarly for any other object of my attention. Now there is one object, or quality which always emerges when awareness is directed at itself with sufficient focus, namely subjective temporal flow. This leads to the insight that my awareness is a manifestation of subjective temporal flow associated with me, a flow of consciousness whose specific form uniquely determines my subjectivity, indeed my very identity. It is a short step from that insight to the recognition that the awareness of any conscious being is a version of subjective temporal flow shaped in the manner specific to that being. To employ an analogy, due essentially to Heraclitus, rivers have in common the fact that each is a manifestation of flow, but the form of that flow varies with the river and determines its identity.
The practice of meditation is connected with this insight. It enjoins us to still the flux of thoughts and impressions coursing through our minds, through our individual awareness, whose varied forms distinguish us. Meditation is intended to replace that varied flux by stasis, which through its nature admits no distinctions, so revealing a unity, a single awareness, a single consciousness.

- **Direct accessibility to consciousness.** I have no direct access in the present to the content of my own past awareness, only indirect access through memory. My constant access in the present to my own awareness leads me to affirm with certainty Descartes’ *cogito ergo sum*. We can be certain of nothing except as it is given in the present. But whatever certainty we achieve in the present is instantly snatched away by temporal flux, with the sole exception of the certainty that a flux is still in place. The *cogito* is confined to the present.

Nevertheless, through a shift of tenses I can also affirm, admittedly with a reduced degree of certainty, *cogitabam, ergo eram*: I thought, therefore I existed. This is put objectively. Subjectively, it should be phrased, *I remember that I thought, therefore I existed at the time of thinking.* (Not up to the Latin here!) This seems to me, *contra* Hume, to furnish a grammatical basis for the belief in the idea of the self as subsisting through time.

If I am materialist and accept that my consciousness is in some mysterious way a product of (or at least correlated with) neural processes in my brain I also have to acknowledge that I have no direct access to the physical activity going on there. When I close my eyes I may see flashes in my retina but I cannot see the underlying neural “wiring” that presumably produces these flashes. Nevertheless, unlike direct access to past mental states, the brain *could* be wired up in such a way as to give consciousness direct access to its own neural activity. This is the essential point of Philip K. Dick’s story *The Electric Ant*, in which an
android is directly presented with evidence that its own subjective awareness is the result of nothing more than the scanning of a tape unspooling in its chest - note, not its head!

I can say that I have direct access to the content of my present awareness, and I had direct access to my past awareness, even if I do not have such access now. This is essentially predicated on the certainty to me of my existence in the present. But, even in the present, I do not have, and indeed have never had, direct access at all to the content of the awareness of other minds, in particular that of your mind. I am certain of my own existence in my present, but I am also certain that I have no direct access to your certainty of your existence in your present. The fact that I now have no direct access to my past awareness is strikingly analogous to the fact that I never have direct access to your awareness. Above I attempted to infer the existence of my past awareness (not that I ever doubted it!) by subjecting Descartes’ cogito to a shift of tenses. To affirm the existence to me of your awareness (not that I ever doubted that either!) in the same spirit I could resort to some further grammatical conceits. First, invert Descartes’ maxim to obtain sum, ergo cogito, and then shift persons to affirm es, ergo cogitas: you exist, therefore you are aware. Voila! Ah, if only philosophy were as clear-cut as grammar.

Let us return to the great enigma with which we started: how can the brain, as an objectively existing entity which can (in principle) be inspected by anybody (including oneself) engender the subjectivity of the “I”, which is accessible only to me? I am sufficiently materialist to accept that thinking, awareness, subjectivity do not take place without the presence of a functioning brain, just as I accept that my existence is tied to a certain region in objective spacetime. I accept that, before I was born, and my brain was formed, my subjectivity did not exist. I accept that, when my brain dies, my consciousness will die with it. Credo in doctrino materialismo. I can even accept naturalist explanations of the emergence of consciousness, and the claim that conscious awareness is an attribute of brain activity that is “local” in the sense that it is directly accessible
only to the individual brain in which that activity is actually taking place. But the recognition of the local nature of conscious awareness in this objective sense does not, and cannot, do justice to the experience of this local phenomenon by the subject associated with the individual brain. As philosophers have long recognized, the existence of subjectivity is an objective (or at least transsubjective) fact, but the essence of subjectivity is not conceivably reducible to any objective description of the material conditions necessary for its emergence.

**Reflections on Time Travel**

One can conceive the possibility of travelling into the past, avoiding the generation of logical contradictions such as the Grandfather Paradox and the like. In such scenarios, time is radically “split” into the “subjective” time of the timetraveller and the “objective” time of the rest of the universe. In fact this oversimplifies the matter. Think of worldlines (spacetime trajectories) of objects as being parametrized by objective time. Then the worldline of any conscious entity is already “split” into two components: the objective time associated with the objective temporal parameter, and the subjective temporal flow of the entity’s consciousness. In undertaking a journey into the past, the timetraveller’s world-line, formerly directed into the future along with the worldlines of the rest of the objects in the universe, becomes redirected into their past. Yet, supposing the timetraveller’s trip into the past not to be instantaneous, during its trip time will elapsed for it in the usual way and will still be split into subjective and objective components. While the subjective component is simply a continuation of the subjective temporal flow of the timetraveller’s consciousness before embarking on its journey into the past, the nature of the objective temporal component is less clear. In travelling into the past, the timetraveller’s world-line has “deviated” from the bundle of worldlines of the rest of the objects in the universe which are still proceeding into their (supposed common) future. It follows that the objective temporal component of the timetraveller’s worldline while undergoing its journey must have a local nature: it is a feature of the
timetraveller’s worldline which is not shared by the worldlines of the rest of the objects in the universe travelling into the common future. But, despite its locality, its objectivity is indisputable since, for example, during the trip the timetraveller’s body ages and its chronometer advances.

An old man may in principle step into a time machine and revisit the past, even *his* past, but when he steps out of the machine he is no younger *himself* than when he entered it. In that respect there is no essential difference between the “objective” temporal journey and from the summoning up of the past through memory, or the contemplation of photographs of one’s younger self. Gazing at a photograph of myself taken in my twenties, at a party, say, I might wish to be again the young man in the photograph. Now by this I surely do not mean that my present older self should employ a time machine to travel into the past and crash the party. (There, of course, I would have the curious experience of encountering my younger self, with the attendant problems of self-duplication.) In fact the wish to be young again could reasonably be taken to have two meanings. One is easily understood. It is the purely physical desire that one’s body (and, perhaps, one’s emotional perspective) should be restored to the state it was when one was younger. A possible way of achieving this is for my present older consciousness to be projected onto, or injected into, my past self in such a way as to share the mental space occupied by my past self. If both consciousnesses – past and future - were aware of the presence of the other, the principle of the unity of consciousness would be violated, possibly leading to joint mental breakdown. Another possibility is that the older self gains access to the contents of the consciousness of the younger self while the latter remains unaware of the presence of the older consciousness. (This possibility is explored in Olaf Stapledon’s novel *First and Last Men*.)
The wish to be young again could also be taken in the literal sense that my present older self should somehow come to coincide in the present with my younger self. This could be achieved by having both objective and subjective time run backwards in such a way as to induce my present self to metamorphose into my younger self. In classical mechanics this could, in principle, be done by reversing the direction of motion of all the particles in the universe. The result can be compared with a film run backwards. A good illustration here is Jean Cocteau’s film *Orphée*. Cocteau shows his characters passing into the world of death by stepping through mirrors, and when he wants a character to spring back to life, he simply runs the film backwards. A memorable line from the film, bitter in its truth, is: *Mirrors are the doors through which death comes and goes. Look at yourself in a mirror all your life and you’ll see death do its work.* Suppose that Cocteau had actually filmed someone staring into a mirror for his whole life and then reversed the film. One would see the person initially as old, and then, through time-reversal, slowly undergo a transformation into his younger self. This an objective, observable effect. If sound recording had been used, the person’s utterances would be reversed, and so become unintelligible as actual speech, but they would still be objective as sounds. Objectively, then, time-reversal seems clear enough. But subjectively, the effect of time-reversal is most unclear. What would it mean for the person’s consciousness to “run backwards”? Of course, this question does not arise in the case at hand because the film only provides a record of the appearance of the person’s body and sound of his speech as it is varies in objective time. No recording has been made of the flow of thoughts in the person’s consciousness as it varies in (that person’s) subjective time. But even if that subjective flow could be “recorded” in some way, it is not at all clear what temporal reversal would mean from a subjective point of view. The closest one can come to this seems to be the effort one sometimes makes in memory to return to a previous point in the flow of one’s thoughts, for example when one attempts to trace the source of a thought by summoning up previous mental associations, a process that Freud and Proust had both mastered in their different ways. But of course this is not done by “reversing” the flow of consciousness. It is not the subjective temporal equivalent of
tying a knot and then untying it. In attempting to return through memory to a previous point in one’s flow of consciousness still proceeds “forward”. What memory presents is a discrete chain of previous associations, from which the continuous flow accompanying their actual emergence in one’s mind has been discarded. Nevertheless, the consciousness in which the memory operates in the “now” is always continuous, even if its object of attention, in this case a sequence of previous associations, is discrete.

The fact is that if one’s wish to be young again requires that objective and subjective time run backwards in such a way as to induce one’s present older self to metamorphose into one’s younger self, then that would also seemingly require the obliteration of one’s older self (and, perhaps, everything contemporaneous with it.) But this would amount to truncating the universes’ past at the time the photograph of one’s younger self was taken; in other words, chopping off all the worldlines of objects (including, of course own body along with its consciousness) at that time. This, of course can be achieved subjectively in memory, as Proust records in A la recherche du temps perdu. One can recall one’s past self, and imagine returning to the state of that past self, at the same tine summoning up the universe in which that past self was immersed, thus obliterating the future. Objectively, this “truncation” of world-lines can be illustrated by the cutting back of a vine. At any given moment, the vine in its entirety displays all of its stages of growth up to that moment. When it is cut back, each of its tendrils (and so the whole vine, if the cutting is performed uniformly) is “returned” to a stage of growth corresponding to an earlier time.