

Mind: Before & Beyond Computation

Michael S. Jones

A Sceptical View of the Cognitive Sciences

As a scientist (speaking hypothetically), I am unable to think entirely outside of my own thought processes, or, for that matter, inside the thought processes of another. Therefore, considering that whatever analyses we may make of human thinking they will inevitably be conditioned by an identical thought process, and hence be constrained by the subjectivism of the very processes under investigation, without the safeguard of an independent control for comparison, it is ambitious to expect to be able to arrive at a fully objective, or value-free, understanding of those processes. To what extent therefore might it be implicitly *dangerous* (in that it might entail unforeseen and irreversible consequences) for us to attempt to model those processes of thought, to fashion them as a 'technology', with a view to exporting them to inanimate objects or machines, with the potential to assume certain critical decision-making roles? Considering that such efforts to model human intellectual processes have been the driving force behind much of the technological innovation of the last half-century, it seems pertinent to enquire: Why were these problematic objections not raised a good deal earlier, or are we so blinded by technological optimism that we must remain inured to all its negative and disruptive consequences?

Due to the inherent difficulties in approaching the subject from a purely empirical perspective, I do not subscribe to the hard-empiricist position of the cognitive sciences, which view all aspects of thought and of language in terms of computational systems, and which limits the scope of the enquiry to explanation in terms of functional problem-solving mechanisms. There is more to mind than functional problem-solving. That seems a far too reductive approach to an enquiry which can only realistically proceed on the basis of intuition (since empiricism alone cannot provide all the answers). The dominant tendency of 20th Century debates on the philosophy of mind was that of the physicalist identification of 'mind' with 'brain', as a further symptom of this wholesale reductionism in the approach to the study of human intellectual processes. It is as if the concept of mind (a modern descendent of the metaphysical concept of soul), with its promise of independence for all who possessed it, were something of an anachronism for post-

19th Century science, and one which science sought to 'bring into the laboratory', failing to anticipate that such an attempt to 'lock-down' the object of attention would be in effect to deprive us of an important context for discussing issues having little to do with functional computation, or with individuals' discrete physical organs and biological processes.

This insistence to render all aspects of psychical and linguistic operations to the level of functionally describable physical or biological systems is predicated on an assumption that an adequate understanding of these operations is possible on the basis of current or future knowledge of neurophysical and neurochemical criteria in the brain, and deductions thereof, with the addition of the use of advanced scanning techniques. In other words, there is little of consequence to learn about mental activity, other than what may eventually be revealed at the internal empirical level. In the first place, this overestimates the scope of current and future methods and tools of observation in representing with adequacy the biological systems under investigation, and seems to be unwise to the probability that whatever the current state of knowledge about the brain, science may well forever be committed to a greater or lesser degree of hypothesis and speculation over the subject. Secondly, why would one insist upon such a stultifying and restrictive set of analytic criteria, over-estimating the efficacy of empirical knowledge, if one were not already predisposed to constructing, as a standalone technological artefact, a synthetic model of intellectual operations, as a subset of those actual operations, and one which could be made appropriable to a process of mechanisation – that is, through the application of approximate cybernetic models?

Perhaps then it is not quite the case that the wholesale theoretical reduction of mental operations to the level of the physical and the biological results in a more objective, empirically-evidenced, and value-free understanding of them. It rather helps to define, in strictly mechanistic terms, what we might need to extract as the computational 'essence' of cerebral processes, in order to provide the 'blueprints' for a set of radical instrumental and technological ambitions. Science does not develop in a vacuum, and the emergence of the information and cognitive sciences during the mid-20th Century gained its primary impetus in the context of two devastating world wars, and hence from the need to develop new forms of sophisticated weapons technology, and to enhance the computational power of military code-breaking systems.

In this context it is interesting to note that Noam Chomsky's 1950s research into computational linguistics – which lay much of the theoretical groundwork for the project of Artificial Intelligence in its approach to natural language processing –

was undertaken with the financial support of the US Army Signal Corps; the US Air Force Office of Scientific Research; and the US Navy Office of Naval Research.¹

Alongside these clear military incentives promoting research and development into the computational aspects of intelligence and linguistic processing, developments in semi-conductors, solid-state electronics, and integrated circuits led to the incursion into the mass-market during the 1970s of new and revolutionary brain-saving devices: the ubiquitous pocket calculator, and various future-oriented digital timepieces, courtesy of light-emitting diodes and liquid-crystal displays. Indeed, if the gadget-buying public was not encouraged to reflect upon the generative impulse behind all this exciting new technology, it might ideally position itself as the principle target and beneficiary of this incipient technological revolution.

Professor of biology Steven Rose:

“Science cannot happen without major public or private expenditure but its goals are set at least as much by the market and the military as by the disinterested pursuit of knowledge. This is why neuroscientists have a responsibility to make their subject and its potentials as transparent as possible, and why the voices of concerned citizens should be heard not ‘downstream’ when the technologies are already fully formed, but ‘upstream’ while the science is still in progress. We have to find ways of ensuring that such voices are listened through the cacophony of slogans about ‘better brains’ – and the power of the military and the market.”²

In the early 1990s, at the time of the first widespread influx of mobile telephones into the market, there was an enormous amount of personal resistance to the adoption of this new communications strategy. I recall that overtly using a mobile phone in public was rather frowned upon, as if it were a sign of excessively brash and showy behaviour. It was also extremely difficult to get business contacts to accept a mobile number as a contact detail without in addition offering a ‘respectable’ landline number. Mobile phones were surrounded by an aura of bad taste, associated with the image of the itinerant pushy businessman, or the hipster cocaine dealer. Over a number of years however, that resistance was gradually worn down through the relentless marketing takeover of the telecoms companies. If the telcos had not had the advantage of unrestricted advertising, and had been obliged to put it to a public vote in the early stages (‘upstream’ in Steven Rose’s terminology), for instance with the question: *Do you accept the more or less obligatory*

1. See: *Three Models for the Description of Language* (Chomsky, N., 1956); and: *On Certain Formal Properties of Grammars* (Chomsky, N., 1959).

2. *We are moving ever closer to the era of mind control*, *The Observer*, 5 February 2006 (Rose, S., 2006).

round-the-clock use of a mobile phone in your life?, the proposal would, without any doubt, have been pre-emptively rejected.

More contemporaneously, the advent of 'smartphones' into the market did not face the same kind of hurdle. The telcos easily capitalised on their earlier marketing coup, the population having become naturalised to the need to carry around small pocketable communication devices. However, a similar kind of resistance does now seem to affect the reception of such technological advances as *Google Glass* into the marketplace. Wearing the *Glass*, it is no longer possible to maintain the pretence of undivided attention to the person directly in my midst, and it represents a decisively new kind of intervention of technology into the social sphere. Perhaps eventually this resistance too will be successfully overcome by advertising, and we will all be walking around with digital prostheses routinely strapped to our eyeballs. Or is there a threshold beyond which technological incursions on our bodies, rather than merely into our pockets, become morally or aesthetically intolerable?

Or perhaps we have just ceased to be intrigued by the innovations that technocracy, in its endless need to service growth in the economy, continues to throw at us – we are no longer wooed by the prospect of gadgets possessed of *artificial intelligence* because experience shows that, for the most part, they are not quite fully responsive to the nuances of our day-to-day requirements, and the inevitable further trade-off against sociability with a product like the *Glass* is unjustifiable. Or is it the case that we have simply become frustrated because the 'sentient being' we expect may be lurking in the machine is unable to understand a joke?

Some Assumptions of Computational Linguistics

Whatever it is that forms the kernel of our resistance, for many theorists in the cognitive sciences, the failure of current incarnations of machine intelligence to reach any kind of parity with human intelligence (for instance in the tendency of products like *Siri* to make glaring inferential errors in response to the most mundane queries; or its failure to apply context intuitively in order to resolve ambiguities which follow from the polysemous nature of certain words) is due principally to limitations in current hardware capacity, and such shortcomings will be overcome following projected exponential improvements in hardware design and capacity. So, as our minds seem to be uniquely interwoven in our personal and emotional experience, is all that is preventing us from forming satisfying interpersonal relationships with our digital devices simply the problem that computers are just not yet able to do computation fast enough? That seems to be the

implication of recent narrative excursions into the domain of artificial intelligence as exemplified by the movie *Her*, where the protagonist, at some imagined not-too-distant future time, enters into just such a relationship with the 'OS' of his personal computer (nonetheless, the *voice* he falls in love with is the disembodied voice of Scarlett Johansson – reading from a script, written by another *actual human* – rather than that of an inanimate machine responding to its own self-instructions).

The expectation that such spirited congress of humans with machines might become realisable at any time in the near future is predicated on an assumption that both brain and mind (including language and emotion) may be fully describable within the terms of the current state of scientific knowledge – that is, according to the 'known laws of physics' (which underpin all the other sciences). The brain is understood as a biological organ whose cognitive functions are rooted in computational processes. Computation implies a linear sequence of logical operations on data values, with predictive, or algorithmic, properties. Hence, it is envisaged that the entirety of the brain's cognitive functions might be reproduced in the form of commercial electronics. In this projected model the role of 'mind' tends to be represented as the equivalent of a collection of programmable software running on the 'hardware' of the brain. Hence, on the proviso that everything relating to the brain's intellectual operations can be reduced to the level of computation, and that mind can be understood as a collection of algorithms, contemporary shortcomings in the practical implementation of this model of intelligence can be interpreted as a deficiency *in quantity* of some mechanism.

The validation of this last principle will depend upon whether it is possible to determine a computational basis for language, for while computers may seem capable of conducting most routine computational tasks with consummate speed and accuracy, they are beset by recondite problems in interpretation and usage in their attempts at natural language processing.

Chomsky's 1950s research, which I mentioned earlier, can be viewed as an attempt at a quantitative analysis of natural language, specifically that of English, in terms of its grammatical 'phrase structure'. It was an attempt at 'predictive enumeration', that is, to analyse the logical relationships between a finite set of observed sentences, and a projected infinite set of possible 'legal' sentences, in such a way that the natural language could be modelled in ways conformable to an automated computational process, commonly represented in the form of a *Turing machine*.³ A Turing machine is a hypothetical machine model which cognitive and

3. Turing machines are hypothetical devices employed by information scientists as a means of determining whether a particular function will be computable by a potential digital computer. The 'machine' typically consists of the idea of an infinite length of tape (cont.)

computer scientists employ to decide upon the computability of functions. It stands as the technical model for all computer algorithms, as a means of representing functions in a form suitable for processing by potential digital computers. Such computable functions are defined as *recursive* functions. Recursive functions are those in which the definition of the function includes an instance of the function nested within itself; in other words, they are defined *self-referentially*.⁴

For Chomsky, the phrase structure of natural language permits analysis in terms of a recursive function – as sentences may be constructed in the form of, for example: *Alice thinks that size is everything*, the smaller component grammatical sentence *size is everything* is a case of a discrete function nested recursively within a larger (self-same) sentence function. Chomsky et al⁵ identified an analogy between recursive sentence construction and the set of the natural numbers, which they refer to as “discrete infinity”. The set of the natural numbers is subject to a recursive definition: *0 is a natural number* defines the nested instance of the function as a

divided into squares, on each of which may be printed a symbol, together with a scanning/printing device which may stop at any square on the tape to read or write its content. A square may contain only one symbol, and only a single square may be read at any one time. The functional properties of the ‘machine’ consist in: a predefined list of discrete *states* and the ability to change from one state to another; the *read/write* functions; and the *move* functions (one square only and in either left or right direction). In addition a Turing machine depends upon a *table of rules*, which defines the sequence of operations involved in moving from a *start-state* to a *halt-state*. In most examples of Turing machines (i.e., those appropriate to the function of digital computing) the set of symbols which may be recorded on the tape is restricted to {0, 1} – which is a *unary*, rather than a binary, notation – the ‘0’s having the property of ‘blanks’ or spacing elements between blocks of ‘1’s – the latter signifying ‘meaningful’ segments of the tape according to the length of the blocks. Turing had referred to his initial proposal for the model as the “universal computing machine” in his 1936 paper: *On Computable Numbers, with an Application to the Entscheidungsproblem* (Turing, 1936).

4. As a recursive function is defined essentially by reference to itself, through a nested instance of the function, this means that computable functions do not in principle invoke any universally available functional definition – each algorithm is therefore, functionally speaking, unique. Turing’s designation of the Turing machine model as the “universal computing machine” is therefore open to some misinterpretation, for the term “universal” relates only to the adaptability of the model to act as a theoretical host for any number of diverse routines, by successive digital encoding. Importantly, the set of rules that define a particular machine's operations upon the data in its memory tape is always necessarily unique, and hence possesses no universal functional applicability. For a detailed description of the Turing machine model and its application to examples of simple functions, see: *Turing Machines* (Barker-Plummer, D., 2013).

5. *The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?* (Chomsky, N.; Fitch, W. T.; Hauser, M. D., 2002).

discrete whole number, and the remainder of the series is defined as the succession of each natural number by another whole number by adding '1'. The resulting set of discrete whole numbers is an infinite one. Analogously:

"Sentences are built up of discrete units: there are 6-word sentences and 7-word sentences, but no 6.5-word sentences. There is no longest sentence (any candidate sentence can be trumped by, for example, embedding it in "Mary thinks that . . ."), and there is no nonarbitrary upper bound to sentence length. In these respects, language is directly analogous to the natural numbers..."⁶

The quotation is from an article published in 2002, and the use of *words* as the unit division does not really convey the substance of Chomsky's 1950s research, which was generally addressed at "phrase structure grammar", with phrases, or subsidiary sentences, forming the unitary divisions. For a sentence to be infinitely extendable, the minimal unit must be a phrase (*Mary thinks that..*), as successively adding single words will not result in successive grammatical sentences. As a means of analysing recursive sentence construction (i.e., sentences within sentences) we cannot describe words as 'units' because an individual word does not have the grammatical integrity of a sentence.

The choice of words as the minimal units in the above quotation, while somewhat misleading, seems to have been with the aim of simplifying the demonstration, because single words exhibit greater *apparent* integrity as units than do the several words that constitute a phrase. Generally speaking, the attribution of 'unity' to any object implies that the object is 'integral with itself', capable of functioning independently of its specific location, with no unresolved external dependencies. In terms of recursively defined series, if we refer to elements that are nested within the larger series as 'units', we end up with units *inside* other units, which implies a contradiction.⁷ In terms of linguistic constructions specifically, the attribution of unity to internally nested phrases implies that the 'unit' has at least quasi-independence from determinations of external syntax and of context, a consequence which does not really have any ecological validity with respect to the

6. Ibid., p. 1571.

7. I feel that the attribution of 'unity' to the items under analysis is a kind of intellectual luxury, and a form of idealism, which permits an understanding of items through artificial abstraction, in isolation from their structural contexts, in which however in actuality they are always necessarily embedded. I have made a comparable critique applied to the series of the natural numbers; i.e., with respect to the definition of integers as 'integral wholes' (rather than, as I feel they ought to be considered, 'relative indices of numerical value') and the consequences of this critique for expectations of proportionality in quantitative systems – see: *The Limits of Rationality*; and: *Integers and Proportion* (both at URL: <http://somr.info/xcetera>).

communicative content of natural language utterances. Hence the emphasis upon isolable functional units within language suggests that the units themselves exert a principle causal or intentional influence upon meaning, and encourages the tendency for both context and global syntax to appear as concatenated *effects* of language, rather than, respectively, as its structural conditions and motivation.

To make this point more explicitly, consider the following examples with a view to analysing their meanings in relation to context. Take the sample sentence already referred to above: *Alice thinks that size is everything*. We can compare this sentence with another one – for instance: *When estimating the total yield from an oil field, size is everything*. Both examples contain the identical subsidiary grammatical sentence: *size is everything*. The first example might have appeared, for instance, in a commentary on the tale *Alice in Wonderland*; the second one perhaps in an article discussing geological surveillance techniques. In terms of the role of the subsidiary sentence within each of these larger narratives, there is little that is shared between the two complete sentences, except for a degree of hyperbole (whatever the relevance of *size* in either case, it is unlikely to be literally ‘everything’). In the second example the meanings of each of the nouns *size* and *everything* can be inferred locally from the preceding phrase. In the first example, however, the meanings of the nouns are quite ambiguous. Are we to infer that *size* relates to Alice’s own bodily proportions (an inference in conformance with what we already know about the story), or is the writer implying that Alice has made some kind of philosophical abstraction from her own experiences about the nature of ‘things’ in general? We can only know what is intended by these words with reference to the larger narrative of the commentary and possibly to the original story itself. If while reading the commentary we came across the sentence *Alice thinks that size is everything*, we would most likely have already been prepared for the meaning; which is to say that the meaning does not reside integrally within the sentence, but in a larger non-linear narrative space. In Chomsky’s terms, both instances of *size is everything* are functionally equivalent (though not necessarily identical), because the semantic potential of the subsidiary sentence is understood to be a factor of its discrete grammatical integrity. In this view meaning is seen to derive from the presence of meaning-*full* units within a linear arrangement of formal grammatical units, rather than from contextual references within a network of idiomatic associations, such as the non-linear associations suggested above for the example *Alice thinks that size is everything*. In the case of natural language utterances such as this, a functionalist analysis of grammatical phrase structure, in terms of its amenability to a computational structure, is insensitive to context, and therefore will not provide a framework for the accurate parsing of meanings. The discrete integrity of the phrase is not a key to its meaning, and furthermore, it will not

enable a computational device to distinguish instances of literalism from the instances of hyperbole exhibited in either of the two examples given above.

There is a clear tendency within the cognitive sciences of describing linguistic subdivisions as discrete functional units, in spite of the fact that notions of functionality in this sense have no real relevance to the construction of meaning in natural language. I suggest that this tendency is dictated by the requirement for the analysis to conform to the structure of the established and preferred model of the Turing machine, and therefore to render the analysis amenable to a computational structure, rather than for any ontological correspondence such a model might have to organic natural language processes.

Turing Machines and Logical Inconsistency

Turing machines operate on the basis of discrete data values, represented by variable strings of '1's separated by non-data-bearing '0's, arranged in discrete cells upon a linear one-dimensional tape. Each action of the machine (read/write/move-left/move-right/stop), and each of the states it may occupy, is also finite and discrete. One cannot make language computable unless it also conforms to this structure. But the assumption that something functionally corresponding to this array of mechanical logical procedures must lie at the root of cerebral linguistic processes is entirely a deductive inference, without any empirical evidence in support of it. After all, how could one possibly arrange an experimental scenario, involving molecular examination of brains engaged in language production, which could provide any such empirical evidence?

Unlike the semantic elements of natural language, data values in the Turing machine are static and functionally invariable – logical consistency demands that, according to the selective state of the machine, the action it will take (or the value it returns) upon reading, for example, a '0' following a string of five '1's in succession, is fixed and invariable wherever the sequence may appear on a particular machine's memory tape. On the basis that Turing machine computability relates to recursive functions, this is analogous to saying that, in the set of natural numbers, the value of the integer '4' (commonly represented as a string of five '1's in unary notation) is *proportionally* consistent (i.e., by definition, *logically* consistent) in its relation to all other natural numbers. This much appears to be uncontroversial. However, for a particular Turing machine, its instruction to act in a certain way following a string of five '1's is dependent on the specific machine's table of rules (its program), which does not reside on the memory tape, but somewhere external

to it.⁸ That is to say that *logical consistency is not an integral feature of the data itself*, as it is dependent upon a system of rules necessarily located remotely from the data. Those rules however must be unique for each individual Turing machine (algorithm), so that the data residing on a machine's tape acquires its logical consistency only by virtue of an explicit or implicit reference to that particular machine's table of rules.

Natural numbers are commonly represented in decimal notation, but they may also be represented in any other number base: binary, octal, duodecimal (base-12), etc. If one wished to design a Turing machine with the task of outputting the sequence of natural numbers between '0' and '10' in decimal, its table of rules would need to make explicit the rule that the maximum writeable digit is '9' by specifying exactly nine iterations of its incremental function, before that digit must 'roll over' and revert to a '0', and a new digit be spawned to the left with the value '1'. In general, when working in decimal, there is no need to state the rules explicitly – they are just assumed, as decimal is the conventional default system of numerical notation. However, by making those rules operationally explicit, in conformance with the requirements of the Turing machine model, it assists in clarifying for us that the logic of those rules must be unique. Hence, it follows by

8. Although in a *universal Turing machine* (cf. the "universal computing machine", as specified in Turing's 1936 paper (op cit., Note3 above), which serves as a theoretical model for what we have come to know as the digital computer), it is possible to encode the instructions for subsidiary individual Turing machines ('programs') hosted in discrete sections at the beginning of the master machine's memory tape (see: Barker-Plummer, D., op cit., Section 4), the parent machine must still require its own master table of rules, which tells it how to operate upon the encoded child machines. Clearly, the master table of rules cannot itself be located on the universal machine's tape, or the machine would not be able to read its own rules. In my estimation, this appears as a serious oversight in the design of the Turing machine hypothesis; for it is simply taken for granted that the machine 'just knows' the operational instructions in its table of rules, without any provision for how the machine actually accesses those instructions (bearing in mind that the rules do not have universal applicability and must be unique to each machine, or class of machine – the term "universal" in the machine's designation relating to its capacity for encoding subsidiary machines, *not* the universality of its particular logical mechanism). This assumption that Turing machines are somehow 'divinely' instructed suggests an analogy with the way in which we customarily take for granted the rules of the decimal system as the universally appropriate rules for representing the natural numbers. The choice of decimal is in fact quite an arbitrary one; and in general there is a failure to appreciate that the proportionality attributed to the natural numbers is a unique property of that system – one deriving exclusively out of the rules that define the restrictive array of digits available to decimal notation, and which are therefore inconsistent with those defining alternative numerical radices. This issue is discussed further in subsequent paragraphs.

definition that the proportionality of numerical values expressed in decimal must also be considered as a unique property that accrues to those values by exclusive virtue of the fact that they are expressed in decimal notation, and hence that proportionality cannot be considered to be freely transferable across diverse numerical radices.⁹

The choice of decimal as the default system of notation is made on the basis of an external arbitrary rule (analogous to the external location of the tables of rules for specific Turing machines) – we could be using any other numerical radix as a default; and indeed we do in fact use a combination of sexagesimal (base-60), duodecimal, and octal when representing divisions in time. Some native Meso-American cultures (e.g., the *Pamean* in Mexico) employ octal rather than decimal for everyday counting purposes. The notes of a musical scale form an octal series, as do the separable colours of white light. What is important to emphasise at this point is that the rules which define these various notations are *incompatible* – which is to say that they are logically and proportionally inconsistent with each other. What has not yet been acknowledged, not only by previous mathematicians, but also by information scientists devising Turing machines, is that the proportional consistency of values in a decimal series is a *unique product* of the external rule governing the system of available writable digits in decimal notation; and that therefore it cannot be assumed that the relations of proportionality pertaining

9. I have shown elsewhere (re: *Radical Affinity & Variant Proportion in Natural Numbers*, at URL: <http://somr.info/xcetera/ra.php>) that with respect to the decimal exponential series: 10^z , for $z=(0, [...], 10)$, if one represents that same series across a range of alternative numerical radices (I have used those from *binary* to *nonary* – base-9) and then calculates the logarithmic differences between successive integers in each series (i.e., using the derived radical logarithms \log_b), the differences are found to be proportionally inconsistent in each case with those in the decimal series (\log_{10} – where the logarithmic difference between successive exponentials is equal to 1). The logarithmic function is intended to express *common ratios of proportion* and radical logarithms (e.g., \log_8 in the case of octal) are conventionally derived from ‘common’ logarithms (\log_{10}) according to the formula: $\log_8 x = \log_{10} x / \log_{10} 8$. The graph of the results for the decimal series is clearly a horizontal straight line at $y = 1$, and if the ratios of proportion were indeed ‘common’ for the same values when expressed across diverse radices, we should expect to see horizontal straight lines in the graphs for each radical series. The resulting graph in the case of each radical series however displays a series of variegated peaks and troughs, displaying proportional inconsistency. These results confirm beyond doubt that the rules of proportion pertaining between values within decimal notation are inconsistent with those between *numerically equal* values when expressed in alternative radices – a point which appears to have escaped the attention of mathematicians since the invention of logarithms 400 years ago (for our purposes these findings are of particular significance in the case of both *binary* and *octal* notations).

between values when expressed in decimal will be seamlessly transferable to their numerically equal values when expressed as, say, octal, or as binary, or as hexadecimal values. That prevailing assumption is therefore a mathematical error-in-principle.

Logical consistency in the Turing machine corresponds to proportional consistency in the set of natural numbers, which according to the analysis above cannot be considered as freely transferable across diverse numerical radices. Therefore, analogously we can say that the logical import of data values in a Turing machine arises uniquely out of the relationship between the specific machine's table of rules and its memory tape, and cannot be exported to another Turing machine operating upon a different set of rules (however the corresponding data values are translated and represented in the new machine) without consequently incurring a failure in logical consistency.¹⁰

The upshot of this for natural language processing is that there can be no *logically consistent* universal computational algorithm suitable for encoding even English language into machine readable form, for the reason that language is never functionally transparent (likewise, the data on a Turing machine's tape is not transparent when viewed in isolation from the machine's unique table of rules). Analogous to this is the fact that the logic of any specific natural language utterance will be determined by non-universal discourse-specific rules according to the cultural, academic, or professional affiliations of its users.¹¹

10. In terms of the practical application of information technologies, the vulnerability to such a failure would relate to all data procedures involving the passing of *more than one* item of associated data between one digital application and another; for example where a web server passes data to-and-from a database server. For the most part, where cases are considered in isolation, the effects of such a failure in logical consistency would be unnoticeable (bearing in mind that the issue relates to variations in the comparative relations *between* data elements across systems, rather than to phenomenal changes in data elements themselves). However, in general terms, the conglomerate effect would be to undermine the representative value of data processed in this way, in particular where the processing may involve comparative assessments upon quantitative data.

11. With particular reference to language of scientific communities, see Thomas S. Kuhn's discussion of the breakdowns in communication within scientific communities over competing scientific theories, at moments of paradigm change within the natural sciences. Kuhn makes the point that where translation and conversion to new scientific models are required, the process is impeded by the fact that the disputants have no recourse to a neutral language by which competing theories may resolve their differences – the terminology through which a scientific community sustains a prevailing paradigm arises implicitly out of its commitment to certain exemplars, or special cases, which are exactly the items thrown into question during moments of revolutionary change within (cont.)

“Universal Computation” as a Grandiose Conceit

“... I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.”¹²

These prophetic words are those of Alan Turing, written in 1950. What is important to note is that in the same article Turing had already abandoned any serious enquiry into the nature and speciality of human thinking, deciding that to pose the question: *Can machines think?* to a contemporary public consensus would have been “meaningless”, and tantamount to an invitation to ridicule. Turing’s bold prediction therefore anticipates a radical change, less perhaps in the capabilities of machines themselves, but more crucially in the use and definition of fundamental words and concepts.¹³

The meaning of the word ‘thinking’ (or its linguistic antecedents) had probably not undergone much substantial change at all for centuries, or perhaps even for millennia; so for Turing to anticipate that the idea of a ‘thinking machine’ could undergo such a seismic alteration (from the ridiculous to the sublime) within a relatively brief span of fifty years, was to place an inordinate degree of faith in the power of technological advancement – a confidence that has since reached the status of a virtual hegemony amongst a majority of cognitive, computer, and neuroscientists around the world.

After the passage of seventy years are we now any closer to the realisation of Turing’s self-fulfilling prophecy?

the natural sciences: *Exemplars, Incommensurability, and Revolutions*, Section 5 of the *Postscript* to his *The Structure of Scientific Revolutions* (Khun, 1996, pp.198-204):

“The commitments that govern normal science specify not only what sorts of entities the universe does contain, but also, by implication, those that it does not. It follows [...] that a discovery like that of oxygen or X-rays does not simply add one more item to the population of the scientist’s world. Ultimately it has that effect, but not until the professional community has re-evaluated traditional experimental procedures, altered its conception of entities with which it has long been familiar, and, in the process, shifted the network of theory through which it deals with the world.” (ibid., p.7).

12. *Computing Machinery and Intelligence* (Turing, 1950, p.442).

13. Turing had instead recast the problem in terms of what he called *The Imitation Game*, whereby he proposed a thought experiment in which a human addresses a set of questions to a remote computer and consequently unwittingly mistakes the responses to those questions believing them to have been given by a human not a machine. The scenario since became known as the *Turing test*. For further discussion and criticism on Turing’s proposition, see my article: *Is Artificial Intelligence a Fallacy?* (at URL: <http://somr.info/xcetera/aif.php>).

It is true that the general notion of ‘intelligence’ (as a corollary of ‘thinking’) has changed somewhat, so that we now often find the word frequently being employed to indicate the mere possession of, or access to, valuable recorded information; with less emphasis upon the traditionally human processes of reflective or creative understanding – a shift from a dynamic, intellect-based, and performative definition of intelligence, to a static, digitised, information-based one. Even so, the school of Artificial Intelligence has recently decided that it is appropriate to invoke a two-tier qualitative distinction in its program, between ‘AI’ – which has become familiar to us to the point of banality in the various commercial implementations of ‘smart’ technology (largely an engineering project, sustained by virtue of the evolved, static notion of intelligence) – and ‘AGI’ (‘artificial general intelligence’), which aims at the creation of machines possessed of creative and reflective aspects of conscious self-awareness comparable to those of the human mind.

Even amongst experts in the field however, the question of how such AGIs will become technologically feasible remains rather vague and ill-defined. Remarking upon the fact that AI has made no progress whatever (towards AGI that is) during six decades of its existence, Professor David Deutsch, a physicist at Oxford University, wrote:

“Despite this long record of failure, AGI must be possible. That is because of a deep property of the laws of physics, namely the universality of computation. It entails that everything that the laws of physics require physical objects to do can, in principle, be emulated in arbitrarily fine detail by some program on a general-purpose computer, provided it is given enough time and memory.”¹⁴

I am not aware that “the universality of computation” had been identified as a “deep property of the laws of physics” prior to Turing’s formalisation of the principle of computability in his 1936 paper (see Note 3 on p.5 above). Since Turing, computation has come to be understood as a property of information processing systems, in so far as those systems are designed to process solutions to problems identified in the management and assessment of information. It involves the establishment of certain well-defined functional procedures, and therefore depends upon the recognition of certain values or conditions which, when subjected to such a predefined and finite procedure, will result in a new set of values or conditions, at which point, having arrived at a solution, the computation terminates.

It seems that Prof. Deutsch’s statement can be interpreted in two ways. Either he is intending to say that computation is a universal feature of human (i.e., physicists’) *understanding* of the laws of physics in action, as they may be observed

14. *Philosophy will be the key that unlocks artificial intelligence*, *The Guardian*, 3 October 2012 (Deutsch, D., 2012).

in a series of finite scenarios; or (which is a much bolder statement) he is suggesting that computation is actually a *bona fide* physical principle in itself, part of the very structure of the universe, as it were independently of any human process of observation and understanding of physical events. This ambiguity is implicit in Prof. Deutsch's statement, and I believe that he is quite unaware that the ambiguity exists.

To be clear, when Prof. Deutsch refers to "the laws of physics" he means the *known* laws of physics which, as we are all well aware, are a creation of human scientific endeavour, and are subject to historical revisions. For instance, the quantum mechanical properties of light and the value of the speed of light as a fundamental measurement constant were not a part of the known laws of physics until after Einstein's theory of Special Relativity. It is not out of the question that future developments in physics will lead to revisions or reversions in the established laws, in the same way that quantum mechanics leads to a theory about light which includes principles incompatible with either the wave theory of light that preceded it, or Newtonian corpuscular theory before that.¹⁵ In this sense, Prof. Deutsch's phrase: "everything that the laws of physics require physical objects to do" tends to put the cart before the horse, as it is the historically non-continuous 'known laws of physics' that have followed from observations made upon physical events, and which in any given period are constrained within the limits of contemporary understanding.

I have no objection in principle to the first interpretation mentioned above. Computational algorithms may indeed have become an indispensable tool for physicists in their understanding and interpretation of natural and physical phenomena, based upon their set of finite observations of those phenomena. However, in order to undertake such computational analysis it is first necessary to establish a finite procedure, which will describe the transition from one physical state of affairs to another. There is no *continuous* computation (i.e., which does not result in a closed infinite loop). The identification of discrete physical states of affairs (of ends and beginnings) is purely a factor of human understanding, and the need for a function to finitely resolve the transition between a beginning and an end state does not occur in nature. To ascribe computation as a "deep property" of natural physical laws, independently of the human need to understand those laws, is therefore teleological, as it imputes to nature notions of purpose and required

15. Kuhn, op cit., pp.12-13. See also pp.14-15 of: Heisenberg, W., *Planck's Discovery and the Philosophical Problems of Atomic Physics* (1958); published in: *On Modern Physics*, 1962.

ends which can only be justified by an appeal to some sort of divine will.

It is incorrect therefore to claim that AGI must be possible by an appeal to the metaphysical notion of the “universality of computation”, which ignores the fact that there are no universal rules of computation. Such a statement does little more than perpetuate arbitrarily the wish-fulfilling and self-fulfilling prophesy initiated by Turing.

The ambiguity in Prof. Deutsch’s statement tends to blur the distinction between the inherently limited and contingent need for human understanding of the laws of physics in action on the one hand, and the absolute, transcendent, and universal explanatory power of the laws so perceived on the other; the effect of which is to grant to physics an authority over knowledge which it could never and should never justly deserve. To adopt such a position is all very well for a physicist, as it gives to physicists as a community the grandiose privilege of explaining the universe to us mere mortals, through exclusively reductive functionalist principles, in obeisance to which authority all other systems of knowledge must ultimately pale into insignificance.

By his ambiguous incorporation of computation into the laws of physics, and by implication into the structure of the universe, Prof. Deutsch is actually implying something along the lines of: *‘Computation is everything’*, or at least: *‘Everything can be understood through computation’*. But the computational procedures employed in the understanding of exemplary physical scenarios, let’s say at the quantum mechanical level, have no universal applicability – they are meaningful in terms of quantum mechanics alone, as they include references to entities as ‘units’ that have logical import only in the field of quantum mechanics. They cannot be applied with logical consistency to the computations of Newtonian mechanics for instance, which still retain explanatory value at the macro scale. As the rules of computation therefore do not apply consistently even across all the subdomains of physics, computation considered as a ‘universal property’, i.e., considered independently from its required unique set of governing rules and entity definitions, becomes a rather empty term, as all that the residual term implies is that there are identifiable problems that require identifiable solutions – which indeed we might recognise as a universal characteristic of all kinds of human situations. To describe such a universal characteristic in terms of a “deep property of the laws of physics” seems however an exercise in hyperbole.

Enlightenment Reason as a Cipher for Metaphysics

In the previous section, I have identified the instance of a principle (the supposed “universality of computation”) being incorporated into the laws of physics, in order to provide authoritative validation for a speculative hypothesis about the capabilities of future machine technology, which derives solely from metaphysics (i.e., not deriving from any empirical method of proof), by a Professor of Physics who clearly has an abiding preference in favour of such a validation. Moreover, Prof. Deutsch does not even acknowledge, and seems to all intents and purposes to be quite unaware of, the difference between a metaphysical principle and an empirical observation.

This failure to recognise an implicit dependence upon metaphysical principles within scientific deliberation is not an isolated instance however. Both mathematics and physics, along with the other natural sciences to the extent that they rely upon fundamental mathematical and physical principles, routinely employ metaphysical principles in their patterns of explanation, in order to establish the ground rules for their respective practices, while usually insisting on the empirical validity of experimental findings, but in a way that systematically avoids (because it cannot be conducted empirically) a scientific approach to an understanding of metaphysics as a system of thought.

In the doctrines of scientific method, one will frequently come across appeals made to scientific *Reason*, as a governing principle in the formation of scientific judgements. Reason usually implies the employment of the principles of logic, rationality, and proportion, in assessments made upon experimental data, as a palliative to the influence of subjective bias, or even of superstition, into scientific deliberation. During the 17th and 18th Centuries, at the time of great and systemic changes within European Enlightenment Science (or ‘natural philosophy’ as it was then known) Reason became allied principally to the evidence provided from *sense data*: in particular the visual sense. Under the influence chiefly of Francis Bacon’s inductive methodology¹⁶, which established radically new approaches to the collection and interpretation of empirical data, a systematic attempt was made to rid scientific method of centuries-old habits of metaphysical (or ‘syllogistic’) reasoning, by the attempt to eliminate judgements based on intuition – in effect an attempt to preserve the ‘objectivity’ of raw sense data from interpretations by the mind. Intuition, in so far as it tended to generate metaphysical conclusions, became identified as a source of error in the practical applications of science. For the Greeks

16. *Novum Organum, Or True Directions Concerning the Interpretation of Nature* (Bacon, F., 1620).

however, who were the progenitors of the concept of Reason inherited by Enlightenment scientists, intuition had been an essential component in the application of Reason, without which there could be no certain knowledge of Nature.

To the extent that all the sciences rely upon universal rules established within mathematics and physics, it is unlikely that such an attempt to eradicate intuition from scientific judgements could ever have been carried to completion. It is indisputable that mathematics, at least, relies upon core principles which cannot be derived empirically – *a priori* logical concepts such as number, function, relation, infinity, equality, etc.; which concepts therefore must be admitted into the canons of science as the pure forms of logic. The logically pure concepts of mathematics must be exercised through the intuition – the concept of number (in general) is formed without reference to any specific (empirical) instance of quantity and is applied intuitively. The important consideration is how such concepts arise in the mind, since their stability is unaffected by experience, or the information gathered from sense data. It appears at times that experience must even be mitigated to conform to intuitions which arise out of the core principles of mathematics and physics. Therefore, if intuition might conceivably play an important role in the regulation of sensory experience, how *reasonable* was it for Bacon and his adherents to repudiate intuition as the chief source of error in pre-Enlightenment science?

The English empiricist philosophers John Locke and David Hume had arrived at the conclusion that human understanding prior to any sensory experience was impossible¹⁷, and this encouraged the idea that the contents of the mind could therefore be considered wholly as the accumulated results of experience. Intuition, although acknowledged by Locke only in the later parts of his treatise on human understanding as that faculty upon which “depends all the certainty and evidence of all our knowledge”¹⁸, then appears as a *learned* capacity, one that is derived *a posteriori* to experience. Hence intuition, at one stage removed from direct

17. See Locke’s *An Essay Concerning Human Understanding* (Locke, J., 1690); and Hume’s *A Treatise of Human Nature* (Hume, D., 1739).

18. Locke, J., *ibid.*, Book IV, Chapter 2, *Of the Degrees of our Knowledge*, p.521. The word “intuition” does not appear in Locke’s treatise until Book IV (the final book – “intuitive knowledge” appears for the first time in Book III, Ch. 8, p.462). Locke had by that point already undertaken a thorough discussion of the concepts of sensation, perception, reflection, ideas, complex ideas, association, cause & effect, the modes of thinking, etc.; which suggests that he had intentionally avoided the subject of intuition until the later sections, in spite of the fact that in Book IV (p.521) he then declares (extemporaneously) that “bare intuition; without the intervention of any other idea” is the source of the clearest form of human knowledge!

experience, can also appear as a potential source of error, since the more reliable route to greater accuracy and utility in practical knowledge would seem to lie in the restless expansion of data acquired through direct 'unfettered' observations of nature.

Empiricists were keen to dispel the theory that human understanding develops on the basis of certain *innate ideas* – a philosophy that derives from Plato, and which was popular amongst European rationalists. For the empiricists, the belief in innate ideas was a source of mysticism, and tended to reinforce metaphysical principles dogmatically, resulting in inertia and stagnation in scientific thinking. Hence Locke begins his *Essay Concerning Human Understanding* (1690) with the premise that there are no innate principles or ideas in the mind prior to its reception of the data from sensory experience. The mind of a new-born child was essentially a *tabula rasa* (Locke used the analogy of a sheet of blank paper) upon which would be written the "simple ideas of sensation", and around which the principles of the understanding would subsequently be constructed. The initial condition of the mind is therefore perceived to be one of pure receptivity.

In dissatisfaction with this argument, Kant had proposed (nearly a century after Locke's *Essay*..) that on the basis of empirical experience alone one could not explain the formation of the categories of reason which distinguish between necessary and contingent truths, without which it would be impossible to arrive at the idea of a 'universal law' – such laws must have transcendental potential, and be capable of being applied *a priori* to experience.¹⁹ Similarly, the principle of *causality* entails the idea of an effect being "posited by and through the cause and resulting from it", according to the principle of necessity, which could not be arrived at by empirical induction, as this would only show an effect as "merely annexed to cause", i.e., contingently. Regardless of the frequency with which one might witness the same relationship to occur between comparable events, one could not acquire merely by numerical addition the dignity of a necessity required to transcribe the relationship as a universal law – it requires a 'leap of faith', rather than simply an increment to experience.²⁰

For Kant, all concepts of pure reason, exemplified by the pure *a priori* logical concepts of mathematics must, by definition, have the capability to transcend experience; otherwise we would be continually faced with the prospect of

19. *The Critique of Pure Reason* (Kant, I., 1787), Chapter II, *Of the Deduction of the Pure Concepts of the Understanding* – SS 10. *Transition to the Transcendental Deduction of the Categories*.

20. *Ibid.*, Chapter II, SS 9. *Of the Principles of a Transcendental Deduction in general*.

experience undermining reason, and there would be no grounds for certainty. There must therefore be a primary mode of pre-cognition (intuition), which is not determined by experience (through sensory perception), but which nevertheless continually seeks to prove itself (to *represent* itself) in relation to experience:

“The “I think” must accompany all my representations, for otherwise something would be represented in me which could not be thought; in other words, the representation would either be impossible, or at least be, in relation to me, nothing. That representation which can be given previously to all thought is called intuition. All the diversity or manifold content of intuition, has, therefore, a necessary relation to the “I think,” in the subject in which this diversity is found. But this representation, “I think,” is an act of spontaneity; that is to say, it cannot be regarded as belonging to mere sensibility.”²¹

Kant is suggesting that intuition is not, as it appears in the final parts of Locke’s *Essay..*, a post hoc refinement of a matured (or perhaps misguided) understanding, but rather a faculty which operates *at the roots* of the understanding, *a priori* to all sensory experience. The key principles underpinning the subject’s capacity for representing experience to itself are the primary intuitions of *space* and *time*; which are not to be conceived, as is perhaps customary amongst physicists, as concepts which may be deduced purely empirically (since there are no sensible material properties belonging to either space or time *in themselves*), but rather as the “pure forms of sensible intuition”, which form the ‘seat of consciousness’ (a synthesis of internal and external apperceptive states with respect: a) to time, as the internal experience of succession; and: b) to space, as the external condition for the perception of objects), and unaccompanied by which no experience could ever assume form as a coherent representation for the subject.²²

21. *Ibid.*, Chapter II, SS 12. *Of the Originally Synthetical Unity of Apperception.*

22. Prior to the implementations of thought and of logic in association with the cognition of objects, there are, in Kant’s view, two modes of primary sensible intuition through which the subject apprehends empirical reality. These are space and time, which, having no palpable physical properties or appearance in themselves, are to be understood not as items known empirically but as *a priori* intuitive representations of the external and internal senses (respectively), which provide the necessary conditions for the reception of empirical phenomena, externally, in space, and the subject’s internal relations to those phenomena, in time:

“By means of the external sense (a property of the mind), we represent to ourselves objects as without us, and these all in space. Herein alone are their shape, dimensions, and relations to each other determined or determinable. The internal sense, by means of which the mind contemplates itself or its internal state, gives, indeed, no intuition of the soul as an object; yet there is nevertheless a determinate form, under which alone (cont.)

Kant maintains a necessary twofold distinction between the cognition of objects through means of sensory perception, and non-sensible ideas of 'things in themselves', where the latter are apprehended purely intellectually. As the objects of sensory experience are apprehended by us necessarily through the 'manifold of the intuitions of space and time', he argues that we cannot acquire any speculative knowledge of these objects as things in themselves, but only as *phenomena* conditioned through the intuitions of time and space, and consequently as modes of mental representation. It is through these intuitions that we understand the principle of causality in nature, and all objects existing as material phenomena are subject to determination by external causes. Conversely, the mind cannot be apprehended through sensory perception, but only immanently, from within.

Therefore, if the mind is to be understood as possessing the capacity for free will, it can only be so conceived as a thing in itself, that is, intellectually. If we do not maintain a categorical distinction between sensible objects as material phenomena and ideas of things in themselves, and attempt to view the mind as a phenomenon like any other, this must be to subject mind to external causal determinations – to make it a mere *effect* of experience – and which will negate its capacity for freedom.²³

In Locke's analysis of human cognition, sensory perception, together with reflections upon ideas derived from sensory experience, are the formative principles of all understanding. The "simple ideas" of objects (or of their attributes) derived from sense impressions are distinct "positive" or "absolute" ideas of "things in themselves";²⁴ and are contained in the mind in abstraction from the causal relationships in which, as empirical objects, they are necessarily embedded. To understand the relations between objects, or between objects and properties, such as the relations of causality, is to "superinduce" something extraneous onto

the contemplation of our internal state is possible, so that all which relates to the inward determinations of the mind is represented in relations of time. Of time we cannot have any external intuition, any more than we can have an internal intuition of space." (ibid., Introduction – *Transcendental Doctrine of Elements*. First Part. *Transcendental Aesthetic* – Section 1. *Of Space* – SS2. *Metaphysical Exposition of this Conception*).

23. Ibid., Preface to the Second Edition.

24. "Whatsoever doth or can exist, or be considered as one thing is positive: and so not only simple ideas and substances, but modes also, are positive beings: though the parts of which they consist are very often relative one to another: but the whole together considered as one thing, and producing in us the complex idea of one thing, which idea is in our minds, as one picture, though an aggregate of divers parts, and under one name, *it is a positive or absolute thing, or idea.*" (Locke, op cit., Book II, Ch. 25, *Of Relation*, p.304 – my emphasis).

“the real existence of things”,²⁵ which otherwise have a kind of free-floating independence in the mind, as positive ideas of things in themselves. This is in contradistinction to Kant’s view, in which it appears impossible to form any positive cognition of things in themselves, but only as phenomena conditioned through modes of representation, and hence also the determinations of causality. For Kant, with regard to the objects of sensory experience, the unconditioned ‘thing in itself’ cannot be thought without contradiction.²⁶ Thus, the granting of transcendental positive truth value to the simple ideas derived from sensory experience in Locke’s analysis is illegitimate, as it fails to appreciate that the conditions for the reception of such impressions are *a priori* faculties of the understanding, in particular the intuitions of space and time; such that it is fair to say that the understanding operates, in some degree, as the author of its own experience. To the extent that any understanding of empirical relations is grounded upon the metaphysical synthesis of the intuitions of space and time, analyses of human understanding which exclude metaphysical considerations must remain indifferent to their own non-empirical foundations.

The history of empiricism from the 17th Century onwards is the history of this indifference. The simple ideas of objects and of their attributes, acquired through sense perception, serve as the basic units for an instrumental interpretation of the world according to a new system of logic. While certainty in natural knowledge had previously depended upon a contractual acknowledgement of the limits of knowledge ultimately derived through intuition, this began to appear as an impediment to advancement of Science in its practical mastery over nature. Empiricism was to invigorate scientific method by the implementation of a new system of logic whereby certainty is granted instead through the direct correspondence of distinct ideas with their empirical referents. By breaking down the structure of human understanding to its discrete positive components, that is, to those elements that could be assured to derive purely from ‘unmediated’ sense perception, concerns over subjectivism or over the non-empirical antecedents of pure reason, were effectively abrogated. While the formative principles of the new system of logic may have continued to be scrutinised within philosophy, and within the arts (in particular by Romanticism), the post-Enlightenment schism of science from philosophy meant that, in instrumental terms, the empirical sciences

25. “This further may be considered concerning relation, *that though it be not contained in the real existence of things*, but something extraneous and superinduced, yet the ideas which relative words stand for are often clearer and more distinct than of those substances to which they do belong.” (ibid., p.305 – my emphasis).

26. Kant, op cit., Preface to the Second Edition.

became institutionally immune to the need to reflect upon, or even to comprehend, their own metaphysical foundations.

Conclusion: There is No Algorithm

Although empiricism has its origins in Aristotle and Stoic philosophy, Locke is generally acknowledged as the founder of British empiricism in its modern form. It would be difficult to overestimate Locke's influence, not only upon the Sciences, but also upon social, political, and economic theory since the end of the 17th Century. If one wished to identify a single most important contributor to modern secular humanist and libertarian Capitalist thinking, and to contemporary definitions of wealth, property rights, and individual liberty, that person would be Locke. He is also cited as a significant influence upon the *American Declaration of Independence* and on the formation of the *United States Constitution*. He is described by Thomas Jefferson,²⁷ together with Bacon and Newton, as:

“[One of] the three greatest men that have ever lived, without any exception, and as having laid the foundation of those superstructures which have been raised in the Physical & Moral sciences [...]”

There is a more or less unbroken line of influence following from Locke's philosophical positivism, through Hume and Berkeley in the 18th Century, to Mill's utilitarianism in the 19th, to the analytical philosophy and logical positivism of Wittgenstein, Russell, and the Vienna Circle in the 20th. However, it is the attempts at radical *formalisation* of thought and of language by logical positivists in which Locke's influence is most poignant, and which furnished the key epistemological premises informing 20th Century developments in the information and cognitive sciences. For example, in Turing's early speculations upon the criteria for designing machines with the capacity to imitate human intelligence, he wrote:

“Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain. Presumably the child brain is something like a notebook as one buys it from the stationer's. Rather little mechanism, and lots of blank sheets. (Mechanism and writing are from our point of view almost synonymous.) Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed.”²⁸

27. *The Letters of Thomas Jefferson, Letter to Richard Price Paris*, (Jefferson, T., 1789).

28. *Computing Machinery and Intelligence* (Turing, A., 1950, p.456).

“Rather little mechanism, and lots of blank sheets”, which adopts exactly Locke’s analogy for describing the initial state of pure receptivity of the child’s brain, so that the ideas the child receives are here conceived as direct positive imprints from a reality *pre-given* to sensory experience. It is significant in this context that Turing apparently ignores the fact that the brain of a child must already be replete with highly sophisticated mechanisms for controlling its diverse bodily functions, even from a point in time well before its birth; yet at the same time, he supposes, it is mysteriously *vacant* of mechanisms with respect to its intellectual faculties.

Turing clearly thinks it expedient to the acquisitive task of ‘obtaining the adult brain’²⁹ to ignore the “rather little mechanism” (*whatever* that might consist of), while actually having no idea of its complexity or its structure. This attitude arises from the premise that the mind can be conceived as a collection of procedures (‘mechanisms’) which develop as *a posteriori* solutions to problems in the adjudication of sense data; so that the mind tends to be interpreted as an assemblage of ad hoc puzzle solutions. But if this is how we are to understand the mind develops, what then gives to it the organisational *will* to derive meaningful solutions out of the disorder of its earliest sense impressions?

To return to the shared analogy of the blank sheets of paper, it matters little which particular analogy one employs here, so long as one bears in mind that even a blank sheet of paper exhibits both a structure and a set of properties, that is to say it is a medium, which therefore *mediates* (rather than simply ‘reflects’) whatever is transcribed upon it. Apparently, neither Locke nor Turing can conceive of a mind without structure and properties prior to its first sense impressions, hence their need for the analogy; yet both are content to assume that understanding (or thinking) owes nothing at all to these factors, and that it arises purely as an accumulation of a set of reflections of the data acquired from unmediated sense impressions.

The problem for Turing, and for the cognitive sciences generally, is that whatever the “rather little mechanism” might consist in, it is not something that might conceivably be investigated materially, or empirically, since there is no way to define the physical limits of a consciousness. In the ensuing technological project to design an inanimate machine with something-akin-to-a-capacity-for-thinking,

29. The task of ‘obtaining the adult brain’ would also require, seventeen years following Turing’s article, *the illicit procurement of at least one, but probably several children* (including myself, aged five), as sacrificial research subjects in a program of covert neurosurgical experimentation, conducted within the British National Health Service. See URL: <http://www.somr.info>, for my exposition of this medical crime.

the only way to approach the problem is from the somewhat subjective perspective of the contents of thoughts, or of perceptions, followed by an attempt to 'reverse-engineer' the thinking apparatus through the designed complexity of multiply parallel logical procedures (algorithms) performed upon static data values. There is an expectation that some form of analogue of consciousness, or at least the *appearance* of thinking, will simply arise thereof, as a kind of 'accident' of inbuilt complexity. That expectation relies essentially upon a tenuous functional analogy between the contents of human thoughts and data objects stored electronically – an analogy which acquires philosophical precedence most poignantly in Locke's positivist epistemology regarding sense data.

This projection about the appearance of thinking, and how to model it, includes the premise that all meaningful thinking bears a positive, ontological correspondence with objective reality, and can be reduced to statements of propositional logic, for example of the kind: *The Prime Minister is not bald* – i.e., where the meaning is verifiable by some appeal to direct sensory observation; and that logical analysis may be applied to the language of thought to arrive at judgements of truth and falsity in much the same way that pure logic is applied to mathematical statements. Thus, the computational theory of mind addresses the domain of thought exclusively in terms of its functional role in assessing truth statements about the observable world. Logical positivism treats statements of the kind: *There is honour among thieves*, for example, in which there is no logically derivable truth value, as metaphysical pseudo-statements, which are 'meaningless', since there is no entity corresponding to the concept of 'honour' which may be positively identified from sense data. It may be helpful to point out that there is neither an observable entity corresponding to the idea of 'mind' – that which the computational theory of mind nevertheless takes as its object.

The problem of the computational approach therefore is that it takes as its sole domain for the genesis of ideas the data provided from sensory experience, because it is committed to a positivist epistemological perspective (outlined in the previous section) which states categorically that this is the only and original source of all human ideas. Hence it excludes from the constructions of thought the possibility of a synthesis of meaning by elements which have no derivation in sense experience, but whose origin must instead be credited to intuition. The question returns again therefore upon the origin and the relative status of intuition in the hierarchy of cognitive processes.

Intuition is that which Kant refers to as a non-empirical source of knowledge, or rather of understanding. Kant gives it a pre-eminence in the structure of human understanding as a necessary precondition for the internal sense of self, and for the

external perception of objects – one which has tended to be frowned upon for the past 300 years or so, less so in the field of philosophy, but particularly within the empirical sciences. If the faculty of the intuition should have become so deprecated, in spite of the fact that so much of human discourse and reasoning might depend implicitly upon it (even within the sciences themselves), an explanation is likely to be that its priority and place in the order of mental operations was simply inapprehensible to a positivist epistemology prepared to accept as knowledge only that given to it in the form of empirical sense data – while if the intuition is to be understood as Kant would have it, as the very foundation of all empirical knowledge, then by definition it must be in itself both before and beyond the scope of being known empirically.

February 2015

(revised: 29 October 2023)

Select Bibliography

Bacon, F., *Novum Organum, Or True Directions Concerning the Interpretation of Nature* (1620), Constitution Society (URL: http://www.constitution.org/bacon/nov_org.htm - accessed 18/01/2015).

Barker-Plummer, D., *Turing Machines*, *The Stanford Encyclopedia of Philosophy*, Summer 2013 Edition, Edward N. Zalta (ed.) (URL: <http://plato.stanford.edu/archives/sum2013/entries/turing-machine/> - accessed 27/03/2021).

Chomsky, N., *On Certain Formal Properties of Grammars*, MIT, Cambridge, Massachusetts, 1959 (URL: http://somr.info/lib/Chomsky_1959.pdf - accessed 29/10/2023).

Chomsky, N.; Fitch, W. T.; Hauser, M. D., *The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?*, pp.1571-3, *Science*, Vol.298, 22 November 2002 (URL: http://somr.info/lib/Chomsky_et_al_2002.pdf - accessed 29/10/2023).

Chomsky, N., *Three Models for the Description of Language*, MIT, Cambridge, Massachusetts, 1956 (URL: http://somr.info/lib/Chomsky_1956.pdf - accessed 29/10/2023).

Deutsch, D., *Philosophy will be the key that unlocks artificial intelligence*, *The Guardian*, 3 October 2012 (URL: <http://www.theguardian.com/science/2012/oct/03/philosophy-artificial-intelligence> - accessed 27/03/2021).

Heisenberg, W., *Planck's Discovery and the Philosophical Problems of Atomic Physics* (lecture delivered Sept. 4, 1958); published in: *On Modern Physics*, Collier Books, New York, 1962, pp.9-28.

Horst, S., *The Computational Theory of Mind*, 2003 (rev. 2009), Center for the Study of Language and Information (CSLI), Stanford University (URL: <http://plato.stanford.edu/> - accessed 18/01/2015).

Hume, D., *A Treatise of Human Nature* (1739), Project Gutenberg Ebooks (URL: <http://www.gutenberg.org/ebooks/4705> - accessed 27/03/2021).

Jefferson, T., *Letter to Richard Price Paris*, January 8, 1789, in *The Letters of Thomas Jefferson* (URL: <http://www.let.rug.nl/usa/presidents/thomas-jefferson/letters-of-thomas-jefferson/jefl74.php> - accessed 27/03/2021).

Jones, M.S., *Is Artificial Intelligence a Fallacy?* August 2014 (URL: <http://somr.info/xcetera/aif.php> - accessed 29/10/2023).

Jones, M.S., *Radical Affinity and Variant Proportion in Natural Numbers*, November 2011 (URL: http://somr.info/lib/Radical_affinity_and_variant_proportion_in_natural_numbers.pdf - accessed 29/10/2023).

- Jones, M.S.**, *The Limits of Rationality (An important mathematical oversight)*, May 2016 (URL: http://somr.info/lib/Limits_of_Rationality.pdf - accessed 29/10/2023).
- Kant, I.**, *The Critique of Pure Reason* (1787), Meiklejohn, J. M. D. (trans.), Project Gutenberg Ebooks (URL: <http://www.gutenberg.org/ebooks/4280> - accessed 27/03/2021).
- Katz, Y.**, *Noam Chomsky on Where Artificial Intelligence Went Wrong – An extended conversation with the legendary linguist*, *The Atlantic*, 01/11/2012 (URL: <https://www.theatlantic.com/technology/archive/2012/11/noam-chomsky-on-where-artificial-intelligence-went-wrong/261637/> - accessed 27/03/2021).
- Kuhn, T. S.**, *The Structure of Scientific Revolutions*, Chicago UP, 1996.
- Locke, J.**, *An Essay Concerning Human Understanding* (1690), Penn State's Electronic Classics Series (URL: <http://www2.hn.psu.edu/faculty/jmanis/locke/humanund.pdf> - accessed 18/01/2015).
- Penrose, R.**, *The Emperor's New Mind*, OUP, 1989.
- Rose, S.**, *We are moving ever closer to the era of mind control*, *The Observer*, 5 February 2006 (URL: <http://www.theguardian.com/science/2006/feb/05/comment.themilitary> - accessed 27/03/2021).
- Russell, B.**, *The Philosophy of Logical Atomism* (1918), Routledge, 2010.
- Spinoza, B.**, *Concerning the Nature and Origin of the Mind*, in his *Ethics*, Boyle, A. (tr.); Parkinson, G.H.R. (ed.), J.M. Dent & Sons, 1989.
- Thagard, P., Litt, A., Eliasmith, C., Kroon, F.W., Weinstein, S.**, *Is the Brain a Quantum Computer?*, July 2005, University of Waterloo, Waterloo, Ontario, Canada (URL: https://onlinelibrary.wiley.com/doi/pdf/10.1207/s15516709cog0000_59 - accessed 27/03/2021).
- Turing, A.**, *Computing Machinery and Intelligence*, in *Mind* LIX (236), October 1950 (URL: <http://somr.info/lib/Mind-1950-TURING-433-60.pdf> - accessed 29/10/2023).
- Turing, A.**, *On Computable Numbers, with an Application to the Entscheidungsproblem*; *Proceedings of the London Mathematical Society*, 2 (1937) 42: pp.230-65 (URL: http://somr.info/lib/Turing_paper_1936.pdf - accessed 29/10/2023).