How a *New Mathematics* may with advantage be applied in science

J Gerard Wolff[*]

June 16, 2020

Abstract

This paper is about a proposed *New Mathematics* (NM) and its potential in science. The NM is proposed as an amalgamation of mathematics with the *SP System*—meaning the *SP Theory of Intelligence* and its realisation in the *SP Computer Model*—both now and as they may be developed in the future. A key part of the structure and workings of the SP System and the proposed NM is the compression of information via the matching and unification of patterns.

A preamble includes: brief notes about solipsism in science; an introduction to the SP System with pointers to where fuller information may be found; an outline of how the NM may be developed; and a discussion of several aspects of information compression;

In sections that follow: 1) A summary of some of the potential benefits of the NM in science, including: adding an AI dimension to mathematics; facilitating the integration of mathematics, logic, and computing; development of the NM as a *universal framework for the representation and processing of diverse kinds of knowledge* (UFK); a new perspective on statistics; and new concepts of proof, theorem, and so on. 2) A discussion of mathematical and non-mathematical means of representing and processing scientific knowledge. 3) A discussion of how the NM may help overcome the known problems with infinity in physics. 4) how the NM may help in modelling of the quantum mechanics concepts of 'superposition' and 'qubits' via analogies with concepts in stochastic computational linguistics and ordinary mathematics. 5) Likewise, how the NM may prove useful in modelling the quantum mechanics concept of 'nonlocality' and 'entanglement' via an analogy with

---

[*] Dr Gerry Wolff, BA (Cantab), PhD (Wales), CEng, MIEEE, MBCS; CognitionResearch.org, UK; jgw@cognitionresearch.org; +44 (0) 1248 712962; +44 (0) 7746 290775; *Skype*: gerry.wolff; *Web*: www.cognitionresearch.org.

the phenomenon of discontinuous dependencies in natural languages. 6) How the NM, with the SP System, provides alternative, and arguably more plausible, interpretations of such concepts as the 'Mathematical Universe Hypothesis' and the 'Many Worlds' interpretation of quantum mechanics, as described in Max Tegmark's book *Our Mathematical Universe*. Two appendices including: A) a tentative 'tsunami' interpretation of the concept of 'wave-particle duality' in quantum mechanics; and B) A discussion of the possibility of interference fringes with real tsunamis.

Keywords: information compression, SP-multiple-alignment, mathematics, quantum mechanics, relativity, representation of knowledge, artificial intelligence.

1 Introduction

This is a DRAFT. Please do not quote.

In another paper [54], it is argued that much of mathematics, perhaps all of it, may be understood in terms of a set of techniques for the compression of information, and their application (Section 2.4.5). The paper also argues that much the same may be said about logic and computing.

The thinking in that paper derives from two main sources: evidence that much of human learning, perception and cognition may be understood in terms of information compression (IC) (Section 2.4.4); and the *SP System* (meaning the *SP Theory of Intelligence* and its realisation in the *SP Computer Model*) in which IC is centre stage (Section 2.2).

On the strength of those ideas, it is suggested that there is potential for the augmentation and adaptation of mathematics with concepts and mechanisms from the SP System to create a *New Mathematics* (NM) [54, Section 9.2]. Potential benefits of that development are outlined in Section 2.3, below.

In the light of evidence that, in a broad view, science may be regarded as a process of gathering and compressing empirical information (Section 2.4.11), the main aim in this paper is to see whether or how the proposed NM would have benefits for science. The paper also describes associated thinking.

1.1 In this paper, the status of quotes from and references to non-specialist sources

Some readers may object that, in an academic paper, it is not appropriate to include quotes from books or articles that have been written for non-specialist readers, or to make reference to such sources. The reasons that sources like that are quoted from or referenced in this paper are:

- Since key concepts in areas such as quantum mechanics have been established for many years, recent specialist papers or books in those areas normally assume that the reader already knows them.
- With some honorable exceptions, the pioneers of such concepts rarely described them as clearly can now be done by good communicators. And of course the pioneers of such subjects as quantum mechanics could not be fully up to date with twenty-first century research.
- Care has been taken to ensure that the authors of sources that are quoted from or referenced are either established academics and authorities in their own right, or have demonstrated their expertise in their writings. And care has been taken to ensure that what is quoted from or referenced is either up to date or has not been superseded by anything more recent.

1.2 Abbreviations

The abbreviations used in this paper are:

- *Information compression*: shortened to 'IC';
- *Information compression via the matching and unification of patterns* (Section 2.4.3): shortened to 'ICMUP';
- *Human learning, perception, and cognition*: shortened to 'HLPC';
- *Artificial intelligence* shortened to 'AI';
- *New Mathematics* (Section 2.3): shortened to 'NM';
- *Commonsense reasoning and commonsense knowledge* (Section 4) shortened to 'CSRK';
- *'Shut up and calculate'* (Section 4) shortened to 'SUAC';
- *Computable Universe Hypothesis* (Section 8) shortened to 'CUH';
- *Mathematical Universe Hypothesis* (Section 8) shortened to 'MUH';
- *Spooky Action at a Distance* (Section 7.1) shortened to 'SAD'.
- *Minimum Length Encoding* (Section 2.4.10) shortened to 'MLE'.

1.3 Presentation

A ‘preamble’, Section 2, next, describes some ideas that will be needed later in the paper. Section 3 describes several of the potential benefits of the NM, treating each one relatively briefly. Then, in Sections 4 to 8, other topics are considered more fully.

2 Preamble

This section describes some ideas which underpin the main sections of the article that come later. They provide points of reference for later discussions, where appropriate.

2.1 Solipsism and science

The solipsism which underlies science is that, whilst we may believe in a reality that is external to our senses and brains, the way we perceive and understand that external reality, including both observations and theories, is necessarily coloured by the workings of our senses and brains.[1] As Stephen Hawking and Leonard Mlodinow write:

> “According to the idea of model-dependent realism ..., our brains interpret the input from our sensory organs by making a model of the outside world. We form mental concepts of our home, trees, other people, the electricity that flows from wall sockets, atoms, molecules and other universes. *These mental concepts are the only reality we can know.*” ([19, p. 171], emphasis added.

Those mental concepts include those derived via artificial devices such as sensors and computers. For example, we believe in ultraviolet light and radio waves even though our eyes are not sensitive to them, because we can see or hear devices that are sensitive to them and we have a place for them in our concept of a spectrum of electromagnetic waves.

In much of science, we can make observations and devise theories without worrying about solipsism of any kind. But it seems that in some areas of science, especially physics as it has developed in the 20th and 21st centuries, it has become increasingly important to pay attention to the workings of our senses and our brains—as will be indicated at appropriate points later.

---

[1] In his book a *Critique of Pure Reason* (1781), Immanuel Kant called the world as it *is* the *noumenal* world, distinct from the *phenomenal* world as we understand it via our the senses and our brains.

To anticipate a little, the SP System, outlined in Section 2.2, seems to have some useful things to say about the nature of concepts in physics, and perhaps in other areas of science.

2.2 Introduction to the SP System

The *SP System*, meaning the *SP theory of intelligence* and its realisation in the *SP Computer Model*, is the product of a lengthy programme of research, seeking to simplify and integrate observations and concepts across AI, mainstream computing, mathematics, and HLPC, *with IC as a unifying theme.*

The SP System and some of its applications is described most fully in the book *Unifying Computing and Cognition* [44], it is described quite fully in the paper [46], and more briefly in the paper [54, Section 3]. Since the SP concepts provide a foundation for the proposals and discussion in this paper, readers may find it useful to read at least one of those descriptions in conjunction with this paper. This and the other two papers are open access and available to everyone via the internet.

Key publications in this programme of research, including several about potential applications of the SP System are detailed with download links on bit.ly/37Y0NcI. There is more detail on bit.ly/2Gxici2.

Since people often ask what the name "SP" stands for, it is short for *Simplicity* and *Power*, as described in Section 2.4.2.

A bare-bones description of the SP System is here:

1. *Distinctive features and advantages of the SP System.* Since the SP System is a unique and distinctive product of a programme of research seeking to simplify and integrate observations and concepts across a broad canvass, it naturally has points of resemblance to many other systems. This has sometimes been construed to mean that it is "nothing but" X or Y etc. With the aim of reducing the chances of misunderstandings in this area, distinctive features and advantages of the SP System are described in [51].

2. *SP as a brain-like system* . The SP System is conceived as a brain-like system that receives *New* information from its environment via its senses and stores some or all of it in compressed form in its 'brain' as *Old* information. This is illustrated schematically in Figure 1.

3. *The central importance of IC.* IC is a unifying principle in the SP System because of the substantial evidence that exists for the importance of IC in HLPC (Section 2.4.4).

4. *The SP Computer Model* is a computer program which gives expression to the elements of the SP Theory of Intelligence, which guards against vagueness in those ideas, and which provides a means of demonstrating how they work. It has been developed over a long period with many versions and the testing of many tentative ideas, most of which failed this hurdle and were discarded.

5. *SP-patterns and SP-symbols.* All information in the SP System is stored and processed as *SP-patterns*, where an SP-pattern is an array of atomic *SP-symbols* in one or two dimensions. An SP-symbol is simply a mark in an alphabet of mutually-distinctive marks that can be matched with any other SP-symbol in an all-or-nothing manner. One-dimensional SP-patterns are normally shown with round brackets ('`(`' and '`)`') at each end.

6. *IC via SP-multiple-alignment.* Compression of information is achieved largely via the building of *SP-multiple-alignments*, a powerful concept which has been borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics. An example of an SP-multiple-alignment is shown in Figure 3.

7. *How IC is achieved.* An SP-multiple-alignment provides for the economical encoding of a New SP-pattern in row 0 (sometimes more than one) in terms of one or more Old SP-patterns, one per row, in the other rows of the SP-multiple-alignment. For a given SP-multiple-alignment, the amount of compression of the New SP-pattern that is achieved is calculated as described in [46, Sections 4.1] and in [44, Section 3.5].

8. *Heuristic search.* For reasons explained in [53, Section 2.2.1 and 11.2], the building of SP-multiple-alignments normally requires the use of heuristic search, trading accuracy for speed (Section 2.4.3).

9. *SP and probabilities.* With each SP-multiple-alignment, there are associated probabilities ([46, Section 4.4], [44, Section 3.7]): the *absolute probability* of the encoded version of the New SP-pattern, and the very much more useful *relative probability* which facilitates the comparison of one SP-multiple-alignment with another. Also, for such things as probabilistic reasoning ([46, Section 10], [44, Chapter 7]), the SP Computer Model can, with two or more SP-multiple-alignments, calculate the relative probabilities of SP-patterns and SP-symbols ([46, Section 4.4.4], [44, Section 3.7.3]).

10. *Versatility via SP-multiple-alignment.* The concept of SP-multiple-alignment is central in the workings of the SP System and provides most of the versatility of the SP System, summarised below. *The concept of SP-multiple-alignment has the potential to be as significant for an understanding of 'in-*

telligence' as is DNA for biological sciences. It may prove to be the "double helix" of intelligence.

11. *Unsupervised learning* in the SP System means the intake of New information and its storage in compressed form as Old information, as described in point 1. Compression of information is achieved both via the creation of SP-multiple-alignments, and via the creation of *SP-grammars*, where an SP-grammar is a set of SP-patterns that have been shown to be effective, collectively, in the compression of New information. How much compression may be achieved via a given SP-grammar is calculated as described in [46, Section 5.1.2] and in [44, Section 9.2.2].

12. *The paradox of decompression via compression.* An apparent contradiction to the idea that all mental or computer processing may be achieved via IC is that, in speaking or writing, it is trivially easy to create repeated and thus redundant copies of any word, phrase or sentence, and likewise via the use of a simple computer program. How the SP System may overcome the paradox of *decompression via compression* is explained in [44, Section 3.8]. In brief, it means providing enough residual redundancy in any compressed body of information so that, via further compression, the effect of decompression may be achieved.

13. *Versatility of the SP System.* The versatility and potential of the SP System are summarised in [54, Section 3.7], and that summary is reproduced here:

    - *Versatility in aspects of intelligence* including: unsupervised learning; the analysis and production of natural language; pattern recognition that is robust in the face of errors; pattern recognition at multiple levels of abstraction; computer vision; best-match and semantic kinds of information retrieval; several kinds of reasoning (next item); planning; and problem solving. There is more detail in [46] and [44].

    - *Versatility in reasoning* including: one-step 'deductive' reasoning; chains of reasoning; abductive reasoning; reasoning with probabilistic networks and trees; reasoning with 'rules'; nonmonotonic reasoning and reasoning with default values; a non-Bayesian alternative to Bayesian reasoning with 'explaining away'; causal reasoning; reasoning that is not supported by evidence; the inheritance of attributes in class hierarchies; and inheritance of contexts in part-whole hierarchies. There is more detail in [46, Section 10] and [44, Chapter 7]. There is also potential for spatial reasoning [48, Section IV-F.1] and what-if reasoning [48, Section IV-F.2].

- *Versatility in the representation of diverse kinds of knowledge* including: the syntax of natural languages; class-inclusion hierarchies (with or without cross classification); part-whole hierarchies; discrimination networks and trees; if-then rules; entity-relationship structures; relational tuples; and concepts in mathematics, logic, and computing, such as 'function', 'variable', 'value', 'set', and 'type definition. The addition of two-dimensional SP-patterns to the SP Computer Model is likely to expand the representational repertoire of the SP System to structures in two-dimensions and three-dimensions, and the representation of procedural knowledge with parallel processing. There is more detail in [49, Section III-B], and in [46, 44].
- *Seamless integration.* Because of the versatility of the SP System as outlined above, and because this versatility is largely due to the central role of SP-multiple-alignment, there is clear potential for the seamless integration of diverse aspects of intelligence and diverse kinds of knowledge, in any combination. It appears that that kind of seamless integration is *essential* in any artificial system that aspires to the fluidity, versatility and adaptability of the human mind.
- In the light of the SP System's versatility in aspects of intelligence including several kinds of reasoning, versatility in the representation of diverse kinds of knowledge, and the seamless integration of those diverse aspects of intelligence and divers kinds of knowledge in any combination. In general, the SP System has potential as a *universal framework for the representation and processing of diverse kinds of knowledge* (UFK) [49, Section III].

Figure 2 shows schematically how the SP System, with SP-multiple-alignment at centre stage, exhibits versatility in its capabilities and their seamless integration.

14. *SP-Neural* is a version of the SP System in which concepts such as SP-pattern, SP-symbol, and SP-multiple-alignment are expressed in terms of neurons with their interconnections and intercommunications [50].

15. . It is intended that the *SP Machine* will be derived from the SP Computer Model with the application of high levels of parallel processing and improvements in its user interface. It is intended that the SP Machine will be a vehicle for further research and development, by individual researchers and groups, towards an industrial-strength device, with guidance from a 'roadmap' presented in [26]. This development is shown schematically in Figure 4.

16. *The SP System has clear potential to solve several problems in the quest for human-like general AI*,[55], at least 10 of them described by leading researchers in AI in interviews with the writer Martin Ford and presented in his book *Architects of Intelligence*,[17], and 3 of them described in [55].

    This is amongst the strongest pieces of evidence that the SP System captures many of the essentials of human intelligence. *The SP System provides a firmer foundation for the development of human-level general AI than any alternative, including deep neural networks.*

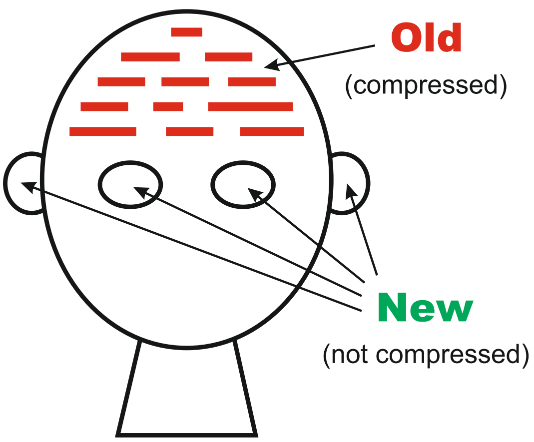

Figure 1: A schematic representation of the SP System. Adapted from Figure 1 in [46], with permission.

2.3 Towards the development of a *New Mathematics*

As mentioned in the Introduction (Section 1), the creation of a New Mathematics was first proposed in [54, Section 9.2]. In that paper it is is suggested that "there is potential for the augmentation and adaptation of mathematics with concepts and mechanisms from the SP System, especially SP-multiple-alignment and unsupervised learning via the building of SP-grammars" and that "those concepts, with associated ideas, may provide the basis of a *New Mathematics*." (*ibid.*).

In brief, the proposed New Mathematics (NM) would be developed like this:

- In broad terms, the NM would be an amalgamation of mathematics and the SP System, both now and as they may evolve in the future.
- More specifically, mathematics would be augmented with structures and mechanisms from the SP Computer Model or SP Machine, including IC

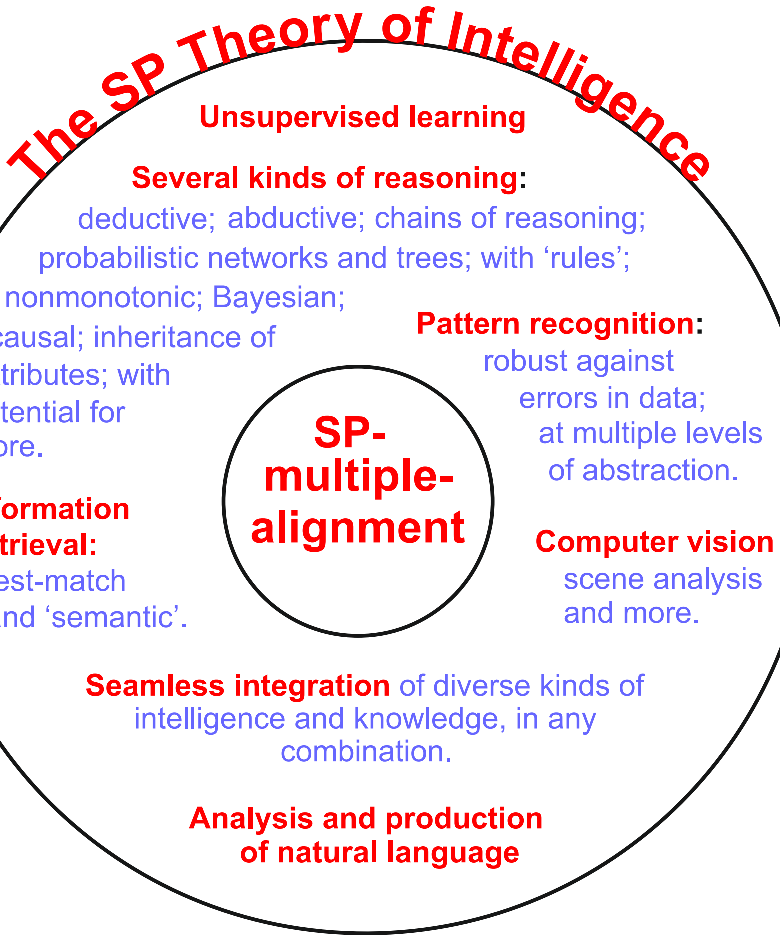

Figure 2: A schematic representation of versatility and integration in the SP System, with SP-multiple-alignment at centre stage. Adapted from Figure 6 in [53], with permission.

```
0                          t w o              k i t t e n     s                    p l a y            0
                           | | |              | | | | | |     |                    | | | |
1                          | | |         Nr 5 k i t t e n #Nr |                    | | | |            1
                           | | |         |                 |  |                    | | | |
2                          | | |    N Np Nr                #Nr s #N                 | | | |            2
                           | | |    | |                          |                  | | | |
3                   D Dp 4 t w o #D | |                          |                  | | | |            3
                    | |          |  | |                          |                  | | | |
4            NP NPp D Dp         #D N Np                         #N #NP             | | | |            4
             |   |                                                  |              | | | |
5            |   |                                                  |         Vr 1 p l a y #Vr        5
             |   |                                                  |         |             |
6            |   |                                                  |  VP VPp Vr           #Vr #VP    6
             |   |                                                  |  |   |                    |
7 S Num    ; NP  |                                                 #NP VP  |                   #VP #S 7
     |     |     |                                                         |
8   Num PL ;    NPp                                                       VPp                         8
```

Figure 3: The best SP-multiple-alignment created by the SP Computer Model with a store of Old SP-patterns like those in rows 1 to 8 (representing grammatical structures, including words) and a New SP-pattern, '`(t w o k i t t e n s p l a y)`', shown in row 0 (representing a sentence to be parsed). Adapted from Figure 1 in [45], with permission.

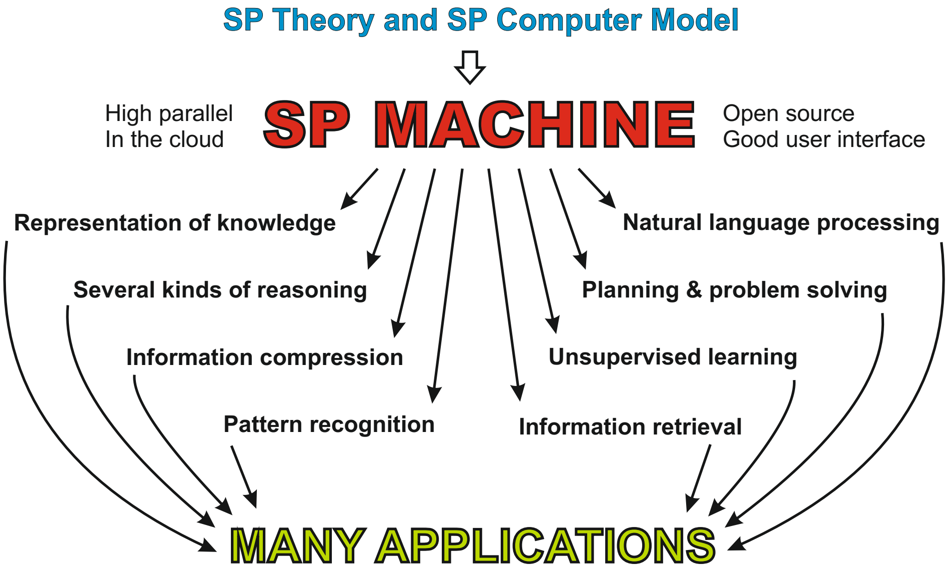

Figure 4: Schematic representation of the development and application of the SP machine. Reproduced from Figure 2 in [46], with permission.

via the building of SP-multiple-alignments and IC via the creation of SP-grammars, as outlined in Section 2.2, with more detail in the sources referenced there.

- Structures in mathematics would be expressed as SP-patterns and SP-symbols. This means that those structures, as with all other knowledge in the SP system, would be processed via ICMUP, and more specifically via the building of SP-multiple-alignments and the creation of SP grammars.
- The vehicle for the NM's representation and processing of data would be the SP Machine, as outlined in Section 2.2 and illustrated schematically in Figure 4.

2.4 Aspects of information compression

This section pulls together some aspects of IC that are relevant to the rest of the paper. Readers may wish to skip most of this and come back to it where necessary.

2.4.1 Concepts associated with information compression

Since IC is a recurring idea in this paper, it needs a definition. The main concepts associated with IC are shown as a diagram in Figure 5, with the meanings of labels expanded in Table 1. The whole rectangle represents the body of information, **I**, which is to be compressed. Of course the *redundancy* in **I**—meaning unnecessary repetition of information in **I**—would normally be distributed throughout **I**, not all in one part as shown in the figure. Likewise for the other categories of information shown in the figure.

In connection with these ideas, it is pertinent to mention that, with any realistically large **I**, the process of identifying all the redundancy in **I** is normally too complex to be achieved by exhaustive search. This means that heuristic methods are needed and it is likely that some redundancy will be missed. Thus it is not normally possible to be sure that all the redundancy has been found. For those reasons, the sizes of several of the elements in Figure 5 can normally only be estimated. These are the ones labelled on the top side of the figure: Reduced RDD (RDD_R), Lost NRI (NRI_L), Reduced NRI (NRI_R), Redundancy in **I** (RDD_O), and Non-redundant information in **I** (NRI_O).

2.4.2 *Simplicity* and *Power*

Compression of a body of information, **I**, may be understood as "maximising the *Simplicity* (with size S) of **I** by reducing, as much as possible, repetition of

Label on figure	*Variables*	*The meaning of the label*
Reduced RDD	RDD_R	The redundancy that remains in **I** after compression of I.
Lost NRI	NRI_L	The portion of the non-redundant information in **I** that has been lost, either by design as in lossy IC or because of weaknesses in the compression algorithm in the case of lossless IC.
Reduced NRI	NRI_R	What is left of the non-redundant information in **I** after the Lost NRI has been subtracted.
Redundancy in **I**	RDD_O	The original amount of redundancy in **I** before compression.
Non-redundant information in **I**	NRI_O	The original non-redundant parts of **I** before compression and loss of non-redundant information.
Redundancy extracted from **I**	RDD_E	The amount of redundancy in **I** that has been extracted by the compression algorithm.
Reduced **I**	I_R	The reduced size of **I** after it has been compressed. Since, as mentioned in the caption to Figure 5, the Lost NRI is not counted.
Original **I**	I_O	The original size of the given body of information, **I**, before it is compressed.

Table 1: The labels in Figure 5, with the names of variables representing their sizes in bits, and with fuller descriptions of their meanings.

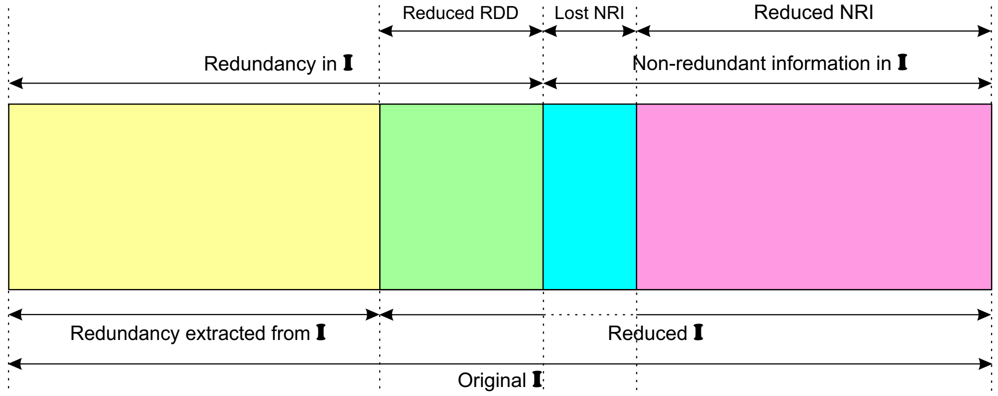

Figure 5: A diagram to illustrate IC-related concepts discussed in the text. The large rectangle represents a given body of information, **I**, processed by an IC algorithm. Table 1 gives a fuller description of each element in this figure, with relatively short names for variables representing the size of each element in bits. Notice that the broken line within the line showing the Reduced **I** means that any Lost NRI is not counted.

information (meaning *redundancy*) in **I**, whilst retaining as much as possible of its non-redundant descriptive or explanatory *Power* (with size P)." [54, Section 3.1].

This may be seen to be equivalent to Ockham's razor—"Entities should not be multiplied without necessity"—if "should not be multiplied" is interpreted as "maximise Simplicity in **I**" and "without necessity" is interpreted as "make sure that the entities in your theory are actually necessary."

In the SP programme of research, the concepts of Simplicity and Power apply to the process of developing the SP Theory itself, and to the IC which is central to the workings of the SP Computer Model. Hence the name "SP".

In general, the SP Computer Model is designed to yield lossless IC. No attempt has yet been made to explore the possible pros and cons of lossy IC. This simplifies the analysis here because we can ignore the 'Lost NRI' (NRI_L) in Figure 5.

In this research:

- *Simplicity* is a feature of any given theory or encoding of **I**. It is defined here as $S = I_O - I_R$, where I_T is the size in bits of **I** when it has been encoded by the compression algorithm. It can be convenient to calculate a normalised value for S as $S_N = I_O/(I_O - I_R)$.

  Normally, I_R would be smaller than I_O, but, as we shall see Section 5.2.3, I_R can be as big or bigger than I_O, so there would be a negative value for IC.

- *Power* is a feature of any given body of information, **I**. It is the size of the non-redundant information in **I**: $P = NRI_O$ which is, with lossless IC, the same as the original size of **I** minus the original size of the redundancy in **I**: $P = I_O - RDD_O$.

But RDD_O can be difficult to measure, and even when a value is found it can be difficult to know whether or not it is the best possible value, because, with most bodies of information, there is a relatively high computational complexity of the process of discovering redundancy in a given **I**.

In general, if we are comparing two or more more algorithms for IC, it is safest to use the same **I** in all cases, and to make comparisons using values of Simplicity. But if one wishes to make comparisons across two or more different **I**s, it is necessary to find some method for making reasonably accurate estimates of RDD_O in each **I**, and to take account of those values.

2.4.3 Information compression via the matching and unification of patterns

An important idea in the SP programme of research is that the process of compressing information may often be understood as a search for patterns that match each other and the merging or 'unification' of patterns that are the same—with the qualification that, depending on the frequency of occurrence of any given pattern, there is a minimum size for compression to be achieved [44, Section 2.2.8.3]. This idea may be expressed more briefly as "information compression via the matching and unification of patterns", which may itself be abbreviated as 'ICMUP'.

In the ICMUP perspective, the process of searching for matches between patterns includes the possibility of finding 'coherent' matches, such as the match between '`A B C`' and '`A B C`', 'discontinuous' matches, such as the match between '`S T A U V W B X C Y Z`' and '`D E F A G H B I J K L C`', and 'partial' matches, such as the match between '`A B C`' and '`B C D`', where those patterns may also be discontinuous.

There are seven main variants of ICMUP ([54, Section 5]. The seventh of those variants, *SP-multiple-alignment*, described briefly in Section 2.2 and more fully in [54, Section 3.3], and illustrated in Figure 3, may be seen as a generalisation of the other six variants of ICMUP [54, Section 5.7].

An important feature of all these variants of ICMUP is that, almost invariably, the search for 'good' matches between patterns, meaning those that yield relatively high levels of IC, is too complex to be achieved effectively via exhaustive search. Almost without exception, it is necessary to use heuristic techniques, trading accuracy for speed: searching occurs in stages, and at the end of each stage, only the best partial solutions are selected to be carried forward to the next stage. With heuristic search, it is not normally possible to guarantee that the best possible solution will be found, but it is normally possible to achieve results that are "reasonably good."

The ICMUP approach to IC which has been adopted in this research contrasts with mathematically-oriented techniques such as arithmetic coding, wavelet compression, transform coding, and the like (see, for example, [33]). This is:

- Partly because of the central importance in the SP programme of research (Section 2.2) of the concept of the afore-mentioned SP-multiple-alignment.
- But it is also because, in connection with the idea that mathematics may be seen as structures and processes for the compression of information (Section 2.4.5), it would not be appropriate for the argument to depend on mathematics. This means that it has been necessary to reach down below the mathematics of other approaches, to focus on the relatively simple, 'primitive' idea that IC may be understood in terms of the matching and unification of patterns.

2.4.4 Information compression as a unifying principle in human learning, perception, and cognition

As mentioned in the Introduction (Section 1), there is substantial evidence for the importance of IC in HLPC [53]. Here are two examples:

- *Unifying 'before' and 'after' views.* "If, when we are looking at something, we close our eyes for a moment and open them again, what do we see? Normally, it is the same as what we saw before. But creating a single view out of the before and after views, means unifying the two patterns to make one and thus compressing the information, as shown schematically in Figure 6." [53, pp. 16–17].
- *Chunking-with-codes in natural language.* A widely-used technique for compressing information is, for any pattern that repeats two or more times (a 'chunk'), assign that pattern to some kind of dictionary and give it a relatively short name or 'code'. Then use the code instead of the pattern wherever the pattern appears. Thus 'New York' is a relatively short code for the very complex city that has that name, 'table' is a short code for the relatively complex object that has that name, 'running' is a short code for the relatively complex activity that has that name. And so on through the vast numbers of nouns, verbs, adjectives, and adverbs, in any natural language. Any such language is a very effective means of expressing information in a very compressed form. [53, pp. 16–17].

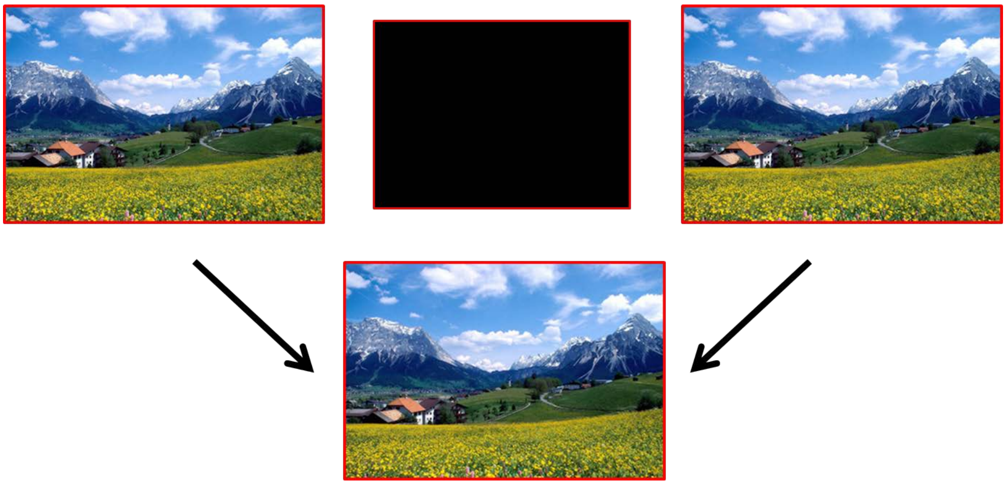

Figure 6: A schematic view of how, if we close our eyes for a moment and open them again, we normally merge the before and after views to make one. The landscape here is from Wallpapers Buzz (www.wallpapersbuzz.com), reproduced with permission.

2.4.5 Mathematics as information compression via the matching and unification of patterns

As mentioned in the Introduction (Section 1), it is argued in another paper [54] that much of mathematics, perhaps all of it, may be understood as a set of techniques for the compression of information, and their application.

In brief, the argument is that there are seven main variants of ICMUP, described in [54, Section 5], and that several of them may be seen in the structure and workings of mathematics [54, Section 6].

One example is the chunking-with-codes technique mentioned in Section 2.4.4. This may be seen in ordinary digits, where 7 is a code for the unary number '///////'; where 27 contains the code 2 for the number of 10s, and 7 is a code for the number of 1s, as before; and so on. Chunking-with-codes may also be seen in any named function such as $sqrt(x)$ or $factorial(x)$. Here, the name is the code and the relatively complex program required to execute the function is the chunk.

Another example is 'run-length coding'—the idea that if two or more copies of a pattern are in a sequence with no intervening patterns, the sequence may be reduced to one instance of the pattern, with something to show that it repeats. Thus 'INFORMATIONINFORMATIONINFORMATIONINFORMATIONINFORMATION' may be reduced to something like 'INFORMATION (x5)' which shows the number of instances of 'INFORMATION', or 'INFORMATION *' with a star, which shows that 'INFORMATION' repeats but does not specify the length of the sequence.

Run-length coding may be seen in the arithmetic operation of addition, where for example, $+7$ within $3 + 7$ may be seen as a shorthand for seven repetitions of the operation of augmenting the unary number '///' with the unary number '/'. In a similar way, run-length coding may be seen in multiplication, where, for example, 3×10 is a shorthand for ten repetitions of the operation $x + 3$, where x starts with the value 0.

As mentioned in Section 3.2, it is argued in [54, Section 7] that similar principles may be seen in the structures and workings of logic and computing.

2.4.6 Information compression and concepts of probability

Due mainly to the development of Algorithmic Probability Theory (APT) by Ray Solomonoff [35, 36], it has been recognised for some time that there is a very close relationship between concepts of IC and concepts of probability.

In view of that intimate connection between IC and concepts of probability, the NM, like SP, would be fundamentally probabilistic. This should not be seen as a problem since it has been recognised for some time that the same is true of mathematics:

> "I have recently been able to take a further step along the path laid out by Gödel and Turing. By translating a particular computer program into an algebraic equation of a type that was familiar even to the ancient Greeks, I have shown that there is randomness in the branch of pure mathematics known as number theory. My work indicates that—to borrow Einstein's metaphor—God sometimes plays dice with whole numbers." [7, p. 80].

As indicated in this quotation, randomness in number theory is closely related to Gödel's incompleteness theorems. These are themselves closely related to the phenomenon of recursion, a feature of many formal systems (including the SP System), many of Escher's pictures, and much of Bach's music, as described in some detail by Douglas Hofstadter in *Gödel, Escher, Bach: An Eternal Golden Braid* [21].

In terms of ICMUP, that close relation between IC and concepts of probability makes sense because, for each unification of two or more copies of a pattern, the number of copies of the pattern is recorded as a frequency, and from frequency values it is relatively straightforward to calculate absolute and relative probabilities for SP-multiple-alignments ([46, Section 4.4], [44, Section 3.7]).

2.4.7 Asymmetry between IC and concepts of probability

The close IC/probability relation (Section 2.4.6) may suggest that, in areas of study such as the development of AI systems or computer models of human cognition, it makes no difference whether the development is based on probabilities or values of IC.

But, for reasons described in [54, Section 8.2], it appears that there is an asymmetry between concepts of IC and concepts of probability. The former are more fundamental than the latter, meaning that IC provides a relatively firm foundation for the development of theory. In brief:

1. *Loss of information in the derivation of probabilities.* As argued in [54, Section 8.2.1], absolute and conditional probabilities may be derived via ICMUP, but the reverse is not true. Values for probability, in themselves, have lost information about the matches and unifications that led to their creation.

2. *SP as an alternative to Bayesian reasoning.* Although Ray Solomonoff, writing in [36, Section 2.4], expresses a positive view of Bayesian concepts of probability, the ICMUP/SP-multiple-alignment approach to probability is largely incompatible with Bayesian concepts, and with at least one advantage over Bayesian concepts (next bullet point). But as noted in Section 3.4, the SP system has been shown, on occasion, to be an effective alternative to Bayesian reasoning. Of course this does not mean that Thomas Bayes was wrong. It simply means that that approach does not marry easily with the SP Theory of Intelligence.

3. *The potential for creating new structures via ICMUP.* With Bayesian and some other probabilistic approach to AI, it is assumed that all of the conceptual entities in a probabilistic analysis have been created already, and there is nothing about how they may be formed [54, Section 8.2.2]. By contrast, ICMUP in the SP System opens up the possibility of isolating words as discrete entities in speech, and likewise for phrases [53, Section 15]. And it can provide a basis for the building of three-dimensional models of entities, as outlined in [47, Sections 6.1 and 6.2].

4. *The scope of frequency information may be extended via ICMUP.* In statistics and other calculations with probabilities, it is often assumed that high frequencies are needed to ensure that the results are statistically significant. In [54, Section 8.2.3] it is argued that such assumptions can be invalid with the matching of patterns in ICMUP. In that kind of situation, frequencies as low as 2 can be highly significant.

5. *Probability, causation, and structure.* Judea Pearl argues [27, 29] that any satisfactory account of causation requires a description of relevant structures,

not merely numbers and equations. ICMUP in the SP System has the potential, via unsupervised learning, to create the kinds of structure that are needed for a comprehensive causal analysis (see also Section 4).

2.4.8 Why is the world so compressible?

While ‘mathematics as information compression’ provides part of the reason why mathematics is so effective in science [54, Section 1], a further question is “Why should the world be so compressible?”[2,3] Possible answers here include:

- In terms of one version of the anthropic principle: if the world were to be entirely incompressible, then everything in the world would be a soup of randomness including mathematicians, scientists, mathematics, and more.
- The world is often not very compressible: “... one must ask why maths is often so unreasonably *ineffective* in the human and social sciences of behaviour, psychology, economics, and the study of life and consciousness. ... Some complex sciences contain unpredictabilities in principle (not just in practice): predicting the economy changes the economy whereas predicting the weather doesn’t change the weather.” (Part of a quote from John Barrow in [54, Section 1], emphasis added).
- The compressibility of parts of the world may be regarded as an observation: an empirical fact with no more or less explanation than many other observations that we make.

2.4.9 How can inductive reasoning be justified?

A question which is quite closely related to the intimate connection between IC and concepts of probability (Section 2.4.6), is “What is the rational basis for inductive reasoning?”, where inductive reasoning takes the general form: “We expect X because it has occurred frequently in the past”, eg “We expect sunrise tomorrow because it has always followed nighttime in the past”.

It is no good saying that inductive reasoning is justified because it has always worked in the past, because that simply invokes the principle that one is trying to justify.

An alternative line of reasoning, summarised in [54, Appendix B], is similar to the reasoning that led some people to buy London bomb sites during the second world war: “If the war were to be lost, the money saved by not making the

---

[2] Thanks to Roger Penrose for raising this point in a personal communication, 2017-05-06.

[3] Here, as elsewhere in science, ‘the world’ is shorthand for ‘everything in the observable universe’.

investment would, in an uncomfortable and uncertain future, probably not be much use anyway."

2.4.10 Minimum Length Encoding

The interrelated concepts of Algorithmic Probability Theory, Algorithmic Information Theory, Minimum Description Length, and Kolmogorov Complexity Theory, described in [23], and Minimum Message Length, described, for example, in [40], revolve around the idea that the information content of a string of symbols is the length of the shortest computer program (in some kind of universal programming language) that can be created or discovered that will produce the given string. So the string '`6 1 3 4 5 8 8 8 5 7 5 3 2 0 4 8 6 7 2 1 5 9 6 1 3 6 7 4 9 4 1 7 2 9 2`' (from a table of random numbers) contains more AIT-information than the equally long string '`2 0 4 8 6 2 0 4 8 6 2 0 4 8 6 2 0 4 8 6 2 0 4 8 6 2 0 4 8 6 2 0 4 8 6`' because the second string can be reduced to a simple 'program', something like this: {'`2 0 4 8 6`' × 7}, but the first string cannot be reduced in that way.

In this area of research, which may be referred to as 'Minimum Length Encoding' (MLE), IC is a central idea, as it is in the SP programme of research. But otherwise the two areas are largely independent, for the following main reasons:

- Research in MLE is founded on the assumption that 'computing' is defined in terms of the concept of a 'universal Turing machine' (or concepts that are recognised as equivalent such as the concept of a 'Post canonical system' [30]). By contrast, the SP System is itself a theory of computing [51, Section III].

- By contrast with MLE research, there is a central role in the SP theory for ICMUP and, more specifically, the concept of SP-multiple-alignment [51, Section III].

- There are several other distinctive features of the SP programme of research, described in [51]. In particular, the overarching goal of the SP research is the simplification and integration of observations and concepts across a broad canvass (Section 2.2).

2.4.11 IC as a unifying theme across several areas

The aim in this section is, in the light of what has been said in preceding sections, to expand a little on the way in which IC may be a unifying theme across several areas:

- *AI and HLPC.* In view of the overarching goal of the SP programme of research—to simplify and integrate observations and concepts across AI, mainstream computing, mathematics, and HLPC (Section 2.2)—the SP Theory of Intelligence, with the SP Computer Model, is as much a theory of HLPC as it is of AI.

  Evidence for IC as a unifying theme in HLPC is presented in [53]. There is less direct evidence for the same idea in [50] and [47].

- *Mathematics.* Since mathematics was developed largely by human brains, and as an aid to human thinking, it should not be surprising, in the light of evidence for the importance of IC (more specifically ICMUP) in HLPC [53], that much of mathematics may be understood as ICMUP, as described in [54].

- *Logic and computing.* Similar arguments apply to logic and computing [54, Section 7].

- *Mainstream computing.* How the SP System may illuminate issues in software engineering is described in [52].

- *Science.* It seems that in science, similar principles apply. The importance of Ockham's razor is widely recognised amongst scientists: John Barrow writes that "Science is, at root, just the search for compression in the world" [4, p. 247]; Ming Li and Paul Vitányi write that "Science may be regarded as the art of data compression" [23, Section 8.9.2]; Albert Einstein writes that "A theory is more impressive the greater the simplicity of its premises, the more different things it relates, and the more expanded its area of application."; and there are more quotations with a similar theme in [54, Section 9.3].

How these ideas may illuminate problems and issues in aspects of science is the main theme of this paper.

3 Some of the potential benefits of the NM for science, in brief

Development of the NM may seem like a monstrous upsetting of the mathematical applecart that has worked well for hundreds of years. But there are many potential benefits from such an NM, most of which would be helpful in scientific research. In subsections below, this section summarises some of them, drawing on and adapting what is described in [54, Section 9.2].

3.1 Adding an AI dimension to mathematics

Since the SP System has been developed with its main focus on the simplification and integration of concepts in AI and HLPC, it has strengths in several aspects of AI/HLPC, as summarised in Section 2.2. How this potential may be developed is described in [26].

The versatility of the SP Computer Model in aspects of AI is summarised in Section 2.2. Also mentioned there is evidence that the SP System has clear potential to solve 19 problems in the quest for human-like general AI [55].

With regard to science: 1) The SP System is well placed to serve as a foundation for the development of general, human-like AI; 2) To the extent that that is successful, there is potential for those AI capabilities to help solve problems in science.

3.2 Facilitating the simplification and integration of mathematics, logic, and computing

While [54, Section 6] shows, with evidence, that much of mathematics, perhaps all of it, may be understood in terms of a set of techniques for IC, Section [54, Section 7] provides evidence that similar principles may be seen in the structures and workings of logic and computing. Taken together, these two bodies of evidence suggest the possibility that mathematics, logic, and computing, may be simplified and integrated as a single system, with IC as a unifying theme.

That kind of integration in the SP Machine has potential to supercharge the system as a tool for scientific research.

3.3 Development of the NM as a UFK

As noted in Section 2.2, the SP System has potential to be developed into a *universal framework for the representation and processing of diverse kinds of knowledge*(UFK). Those arguments apply even more strongly to the NM, drawing as it would on the resources of both the SP System and mathematics (Section 2.3), and with the integration of mathematics, logic, and computing (Section 3.2). To the extent that the NM may achieve the status of a UFK, the potential benefits include:

- *Expanding the role of mathematics in the representation of scientific knowledge.* The NM may expand the role of mathematics in the representation and processing of scientific knowledge. Instead of being confined to the representation of a few brief formulae, the NM may serve in the representation of processing of knowledge that may otherwise be represented with pictures, diagrams, and even verbal descriptions.

- *Helping the integration of related scientific theories.* The NM may be helpful in attempts to integrate separate but related theories such as quantum mechanics and general relativity.

- *Unsupervised learning and the development of scientific theories.* Although the SP Computer Model can demonstrate unsupervised learning ([44, Chapter 9]), it needs further development as outlined in [46, Section 3.3]. However, there is clear potential for much greater capabilities.

  Since science may be seen as a process of gathering information and compressing it (Section 2.4.11), there is potential for the automatic or semi-automatic creation of scientific theories via unsupervised learning.

- *Quantitative evaluation of scientific theories.* In scientific research as it has been up to now, the evaluation of rival scientific theories has been done via more-or-less informal debate, and it seems likely that this will be true for some time to come. There is potential for the NM to achieve quantitative evaluation of scientific theories in terms of IC, as outlined in Section 2.4.

3.4 A new perspective on statistics

Within mathematics, statistical theory is well established and has proved its worth in many applications in science and other areas. But of course there is always room for improvement.

The potential of the NM in this area includes:

- *The probabilistic nature of mathematics.* Since mathematics is probabilistic at its core (Section 2.4.6), there is potential for corresponding developments in any new perspective on statistics.

- *Strengths of the SP System with probabilities.* The strengths of the SP System in the making of inferences and the calculation of associated probabilities (Section 2.2 flow directly from the central role of information compression in the workings of the SP System, and because of the intimate relation between information compression and concepts of probability (Section 2.4.6).

  It appears that, in effect, compression of incoming data via unsupervised learning in the SP System achieves a thorough statistical analysis of those data. Because probabilities are derived from IC and not the other way round, there are potential advantages as outlined in Section 2.4.7.

- *Making good use of small frequencies.* As noted in Section 2.4.7, it is possible with ICMUP to exploit situations where frequencies as low as 2 can be statistically significant. Since ICMUP in the building of SP-multiple-alignments, and since that process lies at the heart of the SP System, the NM would inherit that capability.

- *SP-multiple-alignment and probabilistic reasoning.* The SP-multiple-alignment concept has proved to be a powerful vehicle for several kinds of probabilistic reasoning ([46, Section 10], [44, Chapter 7]), and for their seamless integration in any combination (Section 2.2). Collectively, these several kinds of reasoning, working together, have potential as a powerful aid to statistical inference.

- *Modelling Bayesian networks via the SP System.* The SP System has proved to be an effective alternative to Bayesian reasoning, including reasoning in Bayesian networks and an alternative to Pearl's analysis [27, Section 2.2.4] of the phenomenon "explaining away" ([46, Section 10.2], [44, Section 7.8]).

- *Learning structures via probabilistic associations.* In addition to its strengths in learning patterns of association, the SP System, via unsupervised learning, has strengths and potential to learn entire structures, including the potential to learn 3D structures ([47, Sections 6.1 and 6.2]). In its strengths and potential for the learning of structures, it goes beyond mainstream statistics.

3.5 New approaches to concepts of "proof," "theorem," and related ideas

As noted in [54, Section 9.2.1], there is potential in the development of the NM for the creation of new concepts of proof, theorem, and related ideas. Because IC is central in the workings of the SP System, it is likely that such developments will incorporate IC, and corresponding measures of probability, as indicators of success. There is also potential for the integration of such concepts with concepts of probabilistic reasoning, as described in [46, Section 10] and [44, Chapter 7].

New concepts like these have many potential applications in science.

4 Mathematical and non-mathematical means of representing and processing scientific knowledge

There is quite a lot of evidence that, although mainstream mathematics is often useful for the representation and processing of scientific laws and theories, it may not be so good for the thinking behind those laws and theories. It seems that leading scientists, and probably others, often think in a medium that may be roughly characterised as a 'visual' or 'non-mathematical' medium which is not the same as branches of mathematics such as geometry or topology. Evidence in support of these points, including some from [54, Section 9.2.2], is summarised here:

- *"I've never been good at mathematics"*. Many people say, sometimes with pride, that they have never been good at mathematics—and that includes people who have been successful in careers that require high intelligence, with the implication that they are using non-mathematical structures and processes in their thinking.

- *"Every equation will halve your sales"*. As is well known, Stephen Hawking was told by the publisher of the first edition of his book, *A Brief History of Time*, that "... every equation will halve your sales." [43, location 3227], with the same implication as before.

- *The mind's eye.* Carlo Rovelli writes that "Einstein had a unique capacity to imagine how the world might be constructed, to 'see' it in his mind." [32, location 1025]. In that connection:

  - Walter Isaacson writes that Albert Einstein "repeatedly said that his path toward the theory of relativity began with his thought experiment at age 16 about what it would be like to ride at the speed of light alongside a light beam. This produced a 'paradox,' he said, and it troubled him for the next ten years." [22, location 2207].[4]
  - Here is another example of non-mathematical simplicity in Einstein's thinking:

    > "It is not clear what is to be understood ... by 'position' and 'space'. I stand at the window of a railway carriage which is travelling uniformly, and drop a stone on the embankment, without throwing it. Then, disregarding the influence of the air resistance, I see the stone descend in a straight line. A pedestrian who observes the misdeed from the footpath notices that the stone falls to earth in a parabolic curve. I now ask: Do the 'positions' traversed by the stone lie 'in reality' on a straight line or on a parabola? Moreover, what is meant here by motion 'in space'?" [14, p. 8].

---

[4] The puzzling question was whether, when he was travelling at the speed of light, he would perceive the speed of light as zero or its normal speed. In the end, he decided that the speed of light (in a vacuum) would be the same, however fast one was travelling.

In case these quotes give the impression that Einstein was not good at mathematics: "In 1935, a rabbi in Princeton showed [Einstein] a clipping of [a newspaper column] with the headline 'Greatest Living Mathematician Failed in Mathematics.' Einstein laughed. 'I never failed in mathematics,' he replied, correctly. 'Before I was fifteen I had mastered differential and integral calculus.' " [22, location 488].

- *Scientists don't always use mathematics.* Although one might expect all scientists to be using mathematics, this is not always the case:
  - The geneticist Professor Steve Jones FRS has said that when he comes to equations in scientific papers, he "hums" them.
  - It appears that Michael Faraday developed his ideas about electricity and magnetism with little or no knowledge of mathematics: "Without knowing mathematics, [Faraday] writes one of the best books of physics ever written, virtually devoid of equations. He sees physics with his mind's eye, and with his mind's eye creates worlds." [32, location 623].
  - Charles Darwin's and Alfred Russel Wallace's theory of evolution by natural selection was developed and published as words and pictures [9], and is still normally presented in that form (but see Gregory Chaitin's proposal for a mathematical theory of evolution by natural selection [8]).
- *Feynman diagrams.* The famous diagrams invented by Nobel-prize-winning-physicist Richard Feynman, outside the realms of conventional mathematics, provided and still provide a means of representing concepts in particle physics which greatly simplify the associated calculations:

  > "In 1948 quantum mechanics entered a new phase. Increasingly precise experimental results required new calculation methods, as the existing methods were hopelessly inadequate to deal with the complications of the theory. Richard Feynman came up with a new method that led to enormous simplifications. The method relied heavily on little drawings, now called Feynman diagrams. For a given situation one would draw a few of these diagrams, and then there were simple rules that provided the calculational answers in connection with them. As these diagrams are moreover very appealing intuitively they have become the universal tools of particle physics." [39, p. 244].

- *Alternatives to "Shut up and calculate"*. In connection with attempts to understand quantum mechanics in non-mathematical terms, Philip Ball writes: "Most users don't worry too much about these puzzles. In the words of the physicist David Mermin of Cornell University, they 'shut up and calculate.' " [3, location 146].[5] That last expression, which may be seen to represent the view that quantum mechanics may be understood entirely in terms of mathematics, may be abbreviated to 'SUAC'.

  While it is widely recognised that the mathematics of quantum mechanics accurately describe relevant phenomena, it is, at the same time, notoriously difficult to understand quantum mechanics in non-mathematical terms. That there has been and still is much frustration at this state of affairs, and there are recurrent attempts to develop non-mathematical interpretations of quantum mechanics, is evidence for a strong feeling amongst physicists and others that mathematics by itself is not wholly satisfactory as a way of understanding the world.

  It is likely that, until a satisfactory answer has been found, there will be recurrent attempts to find non-mathematical interpretations of quantum mechanics.

- *Causal diagrams*. A possible reason why pictures or diagrams seem often to be a key to success in scientific research is that they can provide a means for representing and thinking about causal influences [28, 29]. This contrasts with mainstream statistics:

  > "Every student [of statistics] learns to chant, 'Correlation is not causation.' With good reason! The rooster's crow is highly correlated with the sunrise; yet it does not cause the sunrise. Unfortunately, statistics has fetishized this commonsense observation. It tells us that correlation is not causation, but it does not tell us what causation is." [29, location 127].

  And with regard to an electronic circuit diagram shown in [28, Figure 29, p. 345], Pearl writes:

  > "This diagram is, in fact, one of the greatest marvels of science. It is capable of conveying more information than millions of algebraic equations or probability functions or logical expressions. What

---

[5] He adds in a footnote: "It's commonly but wrongly believed that Feynman said this. The belief was so widespread that at one point even Mermin began to fear his quip might in fact have been unconsciously echoing Feynman. But Feynman was not the only physicist with a smart line in quantum aphorisms ...."

> makes this diagram so much more powerful is the ability to predict not merely how the circuit behaves under normal conditions but also how the circuit will behave under millions of *abnormal* conditions.” ([28, p. 344], emphasis in the original).

- *The flexibility in the representation of processes and structures.* Compared with mainstream mathematics, a potential benefit of the NM (and the SP System) is to provide the flexibility of a computer program in the representation of processes, and the representation of structures. It is anticipated that, for both processes and structures, the NM (and the SP System) will provide the versatility of SP-multiple-alignments [54, Section 5], which encompasses the first six of the techniques for IC described in [54, Section 5]. This will provide for such useful things as class-inclusion hierarchies [54, Section 5.5], the key feature of object-oriented design not normally found in mainstream mathematics. It may also provide for parallel processing, as described in [48, Sections IV-B.4, IV-H, Appendix C].

- *Commonsense reasoning and commonsense knowledge.* It appears that the kind of ‘non-mathematical’ medium described in this section, is essentially what researchers are trying to understand in the important field of AI and cognitive science known as ‘commonsense reasoning and commonsense knowledge’ [11]. For that reason, it will be referred to briefly as ‘CSRK’.

  At present, it is envisaged that the SP Computer Model will, in its future development as the *SP Machine*, be enhanced by the addition of two-dimensional SP-patterns [26, Section 9.1]. This should help to provide for the development of CSRK and capabilities of the SP System in computer vision, as described in [47].

  Thus the NM, incorporating these and other developments in the SP System, has potential to meet the need for some kind of visual or non-mathematical medium for the representation and processing of knowledge, as suggested by the examples above.

The kinds of representation and processing outlined above could be one of the benefits of the NM as a UFK (Section 2.3). Such a medium may be useful for scientists in their thinking, and also for the processing of scientific knowledge by the SP Machine. This seems to be one area where science needs to take account of human thinking, as indicated in Section 2.1.

5 Infinity

A prominent feature of mainstream mathematics and other formal systems (including the SP System), and the subject of highly-original research by Georg Cantor (see, for example, [10]), is the ability to describe things that are infinitely big or infinitely small.

This is a feature of the number system and can be done quite simply with a recursive function like the computer function in C or C++ shown in Figure 7. This function has the potential to print an infinitely long sequence of 1s, provided: it does not exhaust the available memory in the computer that is to process the function; and it does not run out of paper; and the computer does not crash; and so on.

```
void infinity()
{
    printf{"\%d ", ‘1’};
    return infinity();
}
```

Figure 7: A simple recursive function in the C or C++ computer language with the potential to print an infinitely long sequence of 1s.

Although it can be fascinating to consider how infinities may come in different sizes, and paradoxes that arise from the concept of infinity, the notion of infinity can be problematic for science. Here we first quote what some scientists say about such problems, and then we show how the SP System may provide some answers.

5.1 What some scientists say about the concept of infinity in science

Max Tegmark writes:

> “The assumption that something truly infinite exists in nature underlies every physics course I’ve ever taught at MIT—and, indeed, all of modern physics. But it’s an untested assumption, which begs the question: is it actually true?” [37, location 804].

He goes on to say:

> “There are in fact two separate assumptions: ‘infinitely big’ and ‘infinitely small.’ By infinitely big, I mean that space can have infinite volume, that time can continue forever, and that there can be infinitely

many physical objects. By infinitely small, I mean the continuum—the idea that even a liter of space contains an infinite number of points, that space can be stretched out indefinitely without anything bad happening, and that there are quantities in nature that can vary continuously. The two assmptions are closely related, because inflation, the most popular explanation of our Big Bang, can create an infinite volume by stretching continuous space indefinitely.

"The theory of inflation has been spectacularly successful and is a leading contender for a Nobel Prize. It explains how a subatomic speck of matter transformed into a massive Big Bang, creating a huge, flat, uniform universe, with tiny density fluctuations that eventually grew into today's galaxies and cosmic large-scale structure—all in beautiful agreement with precision measurements from experiments such as the *Planck* and the BI-CEP2 experiments. But by predicting that space isn't just big but truly infinite, inflation has also brought about the so-called measure problem, which I view as the greatest crisis facing modern physics. Physics is all about predicting the future from the past, but inflation seems to sabotage this. When we try to predict the probability that something particular will happen, inflation always gives the same useless answer: infinity divided by infinity. The problem is that whatever experiment you make, inflation predicts there will be infinitely many copies of you, far away in our infinite space, obtaining each physically possible outcome; and despite years of teeth-grinding in the cosmology community, no consensus has emerged on how to extract sensible answers from these infinities. So, strictly speaking, we physicists can no longer predict anything at all!

"This means that today's best theories need a major shakeup by retiring an incorrect assumption. Which one? Here's my prime suspect: ∞.

"A rubber band can't be stretched indefinitely, because although it seems smooth and continuous, that's merely a convenient approximation. It's really made of atoms, and if you stretch it too far, it snaps. If we similarly retire the idea that space itself is an infinitely stretchy continuum, then a big snap of sorts stops inflation from producing an infinitely big space and the measure problem goes away. Without the infinitely small, inflation can't make the infinitely big, so you get rid of both infinities in one fell swoop—together with many other problems plaguing modern physics, such as infinitely dense black-hole singularities and infinities popping up when we try to quantize gravity." [37, Locations 804–827].

Tegmark acknowledges that “infinity is an extremely convenient approximation for which we haven’t discovered convenient alternatives.” [37, location 830], but warns that “Despite their seductive allure, we have no direct observational evidence for either the infinitely big or the infinitely small.” [37, location 837].

And Jim Al-Khalili writes:

> “An obvious question, with history as our guide, is whether the electron and quarks are indeed fundamental, or made of yet smaller pieces like Russian dolls. The honest answer is: we don’t know! All we can say is that with the best experiments we are able to do today, there is no hint of deeper structure.” [1, Locations 2478–2484].

In a similar vein, writing about the elimination of the concept of infinity in the theory of quantum gravity, Carlo Rovelli says:

> “The infinitely small no longer exists. The infinities which plague conventional quantum field theory, predicated on the notion of a continuous space, now vanish, because they were generated precisely by the assumption, physically incorrect, of the continuity of space. The singularities which render Einstein’s equations absurd when the gravitational field becomes too strong also disappear: they are only the result of neglecting the quantization of the field.” [32, location 2130].

Thus it is possible to develop a scientific theory without the use of infinities. But merely removing infinities from a theory by fiat does not overcome what appears to be a major weakness in mathematics *as a vehicle for scientific knowledge*, but not necessarily in other areas of application.

Since the focus of this paper is on science, Section 5.2 shows how the NM, with the SP System’s basis in IC, can provide principled answers to the problem of infinity in science, preventing inferences that derive from any theory from running too far ahead of their empirical underpinnings.

5.2 How an NM which incorporates the SP System may help to put a brake on the use of infinity in science

As we have seen at the beginning of Section 5, infinity may be expressed in conventional computing via recursive structures like the function shown in Figure 7.

The subsections that follow describe, with examples from the SP Computer Model, variations on how the SP System may capture the same kind of recursion. Section 5.2.1 describes how the SP System, with a small adaptation in its data, may achieve recursive processing without end, like the function shown in Figure 7. Then Sections 5.2.2 and 5.2.3 describe, with examples from the SP Computer Model, how that kind of recursive processing is likely to be misleading in modelling entities or phenomena in the 'world', meaning anything in Earth or beyond.

5.2.1 Case 1: How a potentially infinite sequence of 1s may be created recursively by the SP Computer Model

In the example described here, there is one New SP-pattern, '`(P Q R)`', and two Old SP-patterns: '`(X a P Q R #X)`' and '`(X b X #X 1 #X)`', shown in Figure 8.

```
New SP-pattern

     (P Q R)

Old SP-patterns

     (X a P Q R #X)
     (X b X \#X 1 #X)
```

Figure 8: New and Old SP-patterns as discussed in the text.

Figure 9 shows two of many similar SP-multiple-alignments created recursively by the SP Computer Model with the SP-patterns in Figure 8. Each of those many similar SP-multiple-alignments has zero or more 1s on the right-hand side. Here are some points that may help to clarify the example:

- Recursion is achieved because the SP-symbols '`X`' and '`#X`' at the beginning and end of both of the Old SP-patterns ('`(X a P Q R #X)`' and '`(X b X #X 1 #X)`') match the same two symbols in the body of the SP-pattern '`(X b X #X 1 #X)`'.

- Although there is only one instance of the SP-pattern '`(X b X #X 1 #X)`' in the repository of Old SP-patterns, it can appear one, two or more times in any SP-multiple-alignment. That any one SP-pattern may appear two or more times in any SP-multiple-alignment is an important feature of how SP-multiple-alignments may be formed, as described in [44, Section 3.4.6].

- In the SP System there is a distinction between 'ID' SP-symbols, and 'C' SP-symbols:
  - The former, 'ID', SP-symbols, which are the SP-symbols '`X`', '`#X`', '`a`' and '`b`' in Figure 9, serve to identify or classify the SP-patterns in which

they appear. The ID SP-symbols play a supporting role and may be ignored by anyone whose main interest is in the results of a computation and not how it is done. ID SP-symbols are like such elements of the C or C++ programming languages as ‘`void`’, ‘`printf`’, ‘’, ‘’, and ‘`;`’.

  - The latter, ‘C’, SP-symbols, which are the SP-symbols ‘`P`’, ‘`Q`’, ‘`R`’, and ‘`1`’, in Figure 9, serve to describe the ‘contents’ which is being processed by the system. In the interpretation of the examples in Figure 9, it is mainly the C SP-symbols (‘`P`’, ‘`Q`’, ‘`R`’, and ‘`1`’) that are of interest.

- For this example to work as described below, a small adaptation, mentioned above, is needed in its data. Normally, the ID SP-symbols ‘`a`’ and ‘`b`’ would be assigned a non-zero positive information value in bits which would become the information ‘cost’ of the code derived from any SP-multiple-alignment. But in this example, that information value is set to zero so that any and all SP-multiple-alignments produced by the SP Computer Model yield an encoding cost which is zero.

In principle, this kind of recursive processing can proceed without limits, but with the kinds practical limitations mentioned in connection with Figure 7. However, as we shall see in Sections 5.2.2 and 5.2.3, there are constraints that are likely to apply in any example that aims to model one or more aspects of the world.

In this example, the C SP-symbols ‘`P`’, ‘`Q`’, and ‘`R`’, may be seen to represent established knowledge, meaning things that have been directly observed. The C SP-symbol ‘`1`’, within the SP-pattern ‘`(X b X #X 1 #X)`’, may be seen to represent an inference that may be made recursively from the established knowledge to create such sequences of C SP-symbols as ‘`P Q R 1 1`’, ‘`P Q R 1 1 1 1 1 1`’, and so on.

5.2.2 Case 2: The effect of more complexity in stored knowledge

This subsection and the one that follows describe reasons in principle (not merely practical limitations) why the kind of recursive processing without limits that is illustrated by the SP-multiple-alignments (a) and (b) in Figure 9, may be unrealistic in modelling aspects of the world.

The argument in this subsection is that the world is rarely as precise as mathematics may suggest. This is accepted with things like Boyle's Law—$PV = k$ where P is the pressure of a gas in a container, V is its volume, k is a constant, and the temperature is constant—which is seen to approximate what is happening with the multitude of molecules in the gas. But there is perhaps a need to bear in mind the possibility that, beneath the apparent precision of the mathematics

```
0             P Q R              0
              | | |
1         X a P Q R #X           1
          |         |
2     X b X         #X 1 #X      2
      |                  |
3 X b X                  #X 1 #X 3
```

(a)

```
0                             P Q R                                  0
                              | | |
1                         X a P Q R #X                               1
                          |         |
2                     X b X         #X 1 #X                          2
                      |                  |
3                 X b X                  #X 1 #X                     3
                  |                           |
4             X b X                           #X 1 #X                4
              |                                    |
5         X b X                                    #X 1 #X           5
          |                                             |
6     X b X                                             #X 1 #X      6
      |                                                      |
7 X b X                                                      #X 1 #X 7
```

(b)

Figure 9: Two of many similar SP-multiple-alignments that may be created by the SP Computer Model as described in the text. They show how the SP Computer Model may potentially create an infinite sequence of 1s.

that describes such things as quarks[6], there may be something more complex and messy.

As an example with the SP System, our Case 1 example (Section 5.2.1) may be modified to show how a little more complexity in the repository of Old SP-patterns can disturb the purity of its recursive production of a potentially infinite sequence of 1s.

Thus, if the Old SP-pattern '`(X c X #X 2 #X)`' is added to the Old SP-patterns of Case 1 ('`(X a P Q R #X)`' and '`(X b X #X 1 #X)`' in Figure 8), and if the SP Computer Model is run with the same New SP-pattern as before ('`(P Q R)`'), the result is SP-multiple-alignments exhibiting haphazard sequences of 1s and 2s such as '`2 2 1 2 1`', '`2 2 2 1 2`', '`1 1 2 2 2`', and the like.

In short, a small amount of additional information in the repository of Old SP-patterns may easily add complexity to the sequence of '`1`s' in Case 1 (Section 5.2.1). This added complexity may disturb the neat inferences that things may be infinitely big or infinitely small, as can happen with over-simplistic applications of mathematics.

5.2.3 Case 3: The effect of taking full account of the compression of information

As described in Section 5.2.1, to create an infinite sequence of SP-multiple-alignments like those shown in Figure 9, it is necessary for the ID SP-symbol '`a`' and '`b`' to have zero information values.

If we correct the data so that the information costs of the '`a`' and '`b`' SP-symbols are greater than zero, the picture is different, as can be seen in Table 2.

Here, each row shows variables from an SP-multiple-alignment like the two shown in Figure 9:

1. The first column, headed L, shows, in each row, the length of the sequence of 1s in the corresponding SP-multiple-alignment. Columns 2 and 3 show values as described in in Table 1, and the variables S and S_N are described in Section 2.4.2.
2. The figures in the second column (headed I_O) are all the same because they are the total size, in bits, of the SP-pattern '`P Q R`' in Figure 8, and this SP-pattern is the same in all the 9 SP-multiple-alignments for which data is shown in Table 2.

---

[6] so much so that Max Tegmark has argued in his book *Our Mathematical Universe* that "our physical world is a giant mathematical object" [38, p. 246]

L	I_O	I_R	$S = I_O - I_R$	$S_N = S/I_O$
1	87.21	12.46	74.75	0.857
2	87.21	16.36	70.85	0.812
3	87.21	20.27	66.94	0.768
4	87.21	24.18	63.03	0.723
5	87.21	28.09	59.12	0.678
6	87.21	31.99	55.22	0.633
7	87.21	35.90	51.31	0.588
8	87.21	39.81	47.40	0.544
9	87.21	43.71	43.50	0.499

Table 2: Each row of this table shows the values of variables associated with an SP-multiple-alignment like those in Figure 8. *Key to headings:* L is the length of the sequence of 1s in the given SP-multiple-alignment; I_O is short for the original size of the given body of information **I** before it has been compressed, as described in Table 1; likewise, I_R is short for the reduced size of **I** after it has been compressed, and S is short for 'Simplicity', which is the same as RDD_E as in Section 2.4.2; S_N is a normalised value for S, calculated as $S_N = S/I_O$. The sizes in columns 2, 3, and 4, are all in bits.

3. The progressively increasing figures in the third column (headed I_R) are, in each row, the size of **I** after it has been compressed via the SP-multiple-alignment for that row. Thus, there is a progressive weakening of support for the the recursive inference of what one may call "yet another '1'.

4. In keeping with the figures in the third column, the decreasing values for S_N in the fourth column, calculated as $S = I_O - I_R$, and the decreasing values for a normalised value for Simplicity in the fifth column, calculated as $S_N = I_O/(I_O - I_R)$, show a progressive weakening of support for that recursive inference of 1s.

In general, this example shows how, in the proposed NM, recursive inferences that something would be infinitely big or infinitely small, would be constrained by the way in which the IC which may be achieved via such inferences would decrease progressively with the increasing length of the inferential chain. In view of the linear decrease of S_N with the length of the sequence of 1s (Figure 10), the value of S_N would reach zero when the length of that sequence is about 15, and well before any notion that the sequence might be infinitely long. In practice, there would probably be an even shorter cutoff since the inference under discussion would normally be competing with several other inferences, and many of these might well have values for S_N that would be above zero.

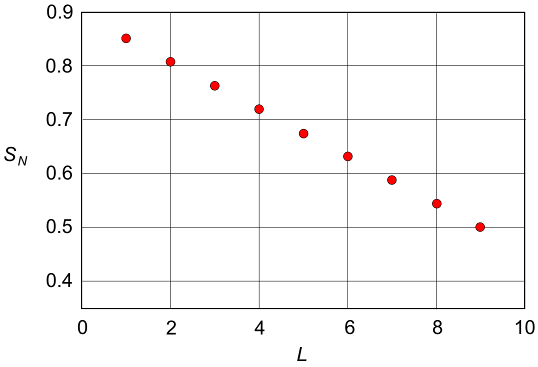

Figure 10: A plot of S_N against L from Table 2.

5.3 Comment

The evidence and arguments presented in Sections 5.2.2 and 5.2.3 are intended to show, with examples, how the SP System, within the proposed NM, would naturally tend to inhibit inferences that anything in nature was infinitely big or infinitely small. But, in keeping with Al-Khalili's previously-quoted remarks (Section 5.1), failure to demonstrate infinity in nature says nothing about whether or not it exists. Frustrating as that may be, we should be content to accept an open verdict, pending further evidence and arguments.

6 Superposition

This section describes what appear to be some parallels between the concept of 'superposition' in quantum mechanics and concepts in theoretical linguistics and in mathematics and computing.

But before we get on to those ideas in quantum mechanics, it seems necessary first to describe some apparent misconceptions associated with quantum mechanics.

6.1 Some apparent misconceptions associated with quantum mechanics

As it's title suggests, this subsection, which draws on parts of Philip Ball's book *Beyond Weird* [3], describes some apparent problems in quantum mechanics.

6.1.1 The concept of a ‘wavefunction’

It seems that a source of confusion in quantum mechanics is when people believe that a wavefunction describes a physical entity. But it is merely an abstraction as described by Ball:

> “The wave in Schrödinger’s equation isn’t a wave of electron charge density. In fact it’s not a wave that corresponds to any concrete physical property. It is just a mathematical abstraction—for which reason it is not really a wave at all, but is called a *wavefunction*.” ([3, Location 435], emphasis in the original).

In this connection, Ball quotes from an article by Berthold-Georg Englert:

> “... were there not the widespread habit of the debaters to endow the mathematical symbols of the formalism with more meaning than they have. In particular, there is a shared desire to regard the Schrödinger wave function as a physical object itself after forgetting, or refusing to accept, that it is merely a mathematical tool that we use for a description of the physical object.” [15, p. 12].

6.1.2 ‘Measurement’ and ‘collapse of a wavefunction’

There seem also to be problems associated with the concepts of ‘measurement’ and ‘collapse of a wavefunction’.

The process of ‘measurement’ in a quantum system is described by Ball thus:

> “Before measurement, ... the system is fully described by a wavefunction from which one can calculate the various probabilities of the different possible measurement outcomes. Let’s say that the system is in a superposition of possible states A, B and C. Then, according to quantum mechanics, the wavefunction can do nothing except continue evolving in its unitary way, preserving these three possible states.
>
> “But measurement does something else. It ‘collapses’ (Heisenberg’s original word was ‘reduces’) those possibilities, expressed in the wavefunction, to just one. Suppose that, before the measurement, the probabilities of finding some property of a quantum object with the values corresponding to states A, B and C are 10%, 70% and 20% respectively. When we make a single measurement on the object, we might find that we get the result C. What happened to A and B? We are forced to assert that the probabilities have now changed: that for C it is 100%, and for A and B it is zero. What’s more, we can’t get A

> and B back: if we repeat the measurement, we'll keep getting C." [3, Locations 1091–1098].

These features of the Copenhagen interpretation of quantum mechanics are another source of confusion:

- As described in Section 6.1.1, the belief that the wavefunction is a physical entity, not an abstraction.
- Correspondingly, the belief that the 'collapse' of the wavefunction is a physical process that we might observe.
- The rather strange language which is used to describe what are really quite ordinary things. For example, the use of obscure words like 'measurement' or the 'collapse' of a 'wavefunction', instead of more humdrum words like 'detect', 'learn', or 'see'.

In connection with the last point, Freeman Dyson writes:

> "Unfortunately, people writing about quantum mechanics often use the phrase 'collapse of the wave function' to describe what happens when an object is observed. This phrase gives a misleading idea that the wave function itself is a physical object. A physical object can collapse when it bumps into an obstacle. But a wave function cannot be a physical object. A wave function is a description of a probability, and a probability is a statement of ignorance. Ignorance is not a physical object, and neither is a wave function. When new knowledge displaces ignorance, the wave function does not collapse; it merely becomes irrelevant." In "The Collapse Of The Wave Function" [6, p. 73].

6.2 Definitions

Before getting on to the main subject of Section 6, we need to understand the concept of superposition.

The concept is described by Al-Khalili like this:

> "The idea of superposition is not unique to quantum mechanics but is a general property of all waves. Imagine watching someone dive into an empty swimming pool. You will see the ripples travel outwards along the surface of the water as simple undulations all the way to the other end of the pool. This is in stark contrast to the state of the water when the pool is full of people swimming and splashing about.

> The turbulent shape of its surface is now due to the combined effect of many disturbances and is achieved by adding them all together. This process of adding different waves together is known as superposition." [1, location 1025].

This is clear. But it appears that the concept of superposition can suffer from the kinds of misconception described in Section 6.1, especially Section 6.1.1 ("The concept of a 'wavefunction'"). Ball suggests how misconceptions can arise like this:

> "This 'two (or more) states at once' is called a superposition. The terminology conjures up the image of a ghostly double exposure. *But strictly speaking a superposition should be considered only as an abstract mathematical thing.* The expression comes from wave mechanics: we can write the equation for a wave as the sum of equations for two or more other waves." ([3, Locations 683–690], emphasis added).

Thus, physical waves can indeed be in a state of superposition, one with another. But it is wrong to believe that superposition of two or more *wavefunctions* has anything to say about the corresponding physical entities, whether they be seen as physical waves or particles (Sections 6.1.1 and 6.1.2).

As one might expect from one of the pioneers of quantum mechanics, Paul Dirac makes a clear distinction between superposition and the probabilities of physical states that it describes:

> "The non-classical nature of the superposition process is brought out clearly if we consider the superposition of two states, A and B, such that there exists an observation which, when made on the system in state A, is certain to lead to one particular result, a say, and when made on the system in state B is certain to lead to some different result, b say. What will be the result of the observation when made on the system in the superposed state? The answer is that the result will be sometimes a and sometimes b, according to a probability law depending on the relative weights of A and B in the superposition process. It will never be different from both a and b. *The intermediate character of the state formed by superposition thus expresses itself through the probability of a particular result for an observation being intermediate between the corresponding probabilities for the original states, not through the result itself being intermediate between the corresponding results for the original states.*" ([12, Locations 359–367], emphasis in the original).

6.3 Similarities between superposition in quantum mechanics and 1) syntactic classes in natural languages, and 2) type definitions in ordinary mathematics and computing

It appears that there are similarities between the concept of superposition in quantum mechanics and the concept of a syntactic class in theoretical/computational linguistics, and also the concept of a 'type definition' or 'data type' in ordinary mathematics and computing. The apparent similarities are described here with the help of a simple example from the SP Computer Model.

6.3.1 Similarity between 'superposition' in quantum mechanics and 'syntactic class' in stochastic grammars

Figure 11 shows a set of SP-patterns that may be seen as a stochastic grammar for a very simple English-like language.

The number at the end of each SP-pattern, preceded by a '⋆', is a supposed frequency of occurrence of the SP-pattern in some imaginary text or texts. Within every SP-multiple-alignment created by the SP System, the frequency of every participating SP-pattern is used to calculate absolute and relative probabilities for the SP-multiple-alignment, and inferences that may be drawn from the SP-multiple-alignment (Section 2.4.6).

The first SP-pattern in the figure, '`(S s1 D #D N #N V #V #S)`', represents the abstract structure of a sentence. Within that SP-pattern, '`D #D`' may be seen as a 'slot', 'space', or 'variable', for a word of the grammatical category 'determiner', '`N #N`' may be seen as a variable for a word of the category 'noun', and '`V #V`' may be seen as a variable for a word of the category 'verb'.

```
(S s1 D #D N #N V #V #S)*750
(D d1 the #D)*600
(D d2 this #D)*150
(N n1 dog #N)*400
(N n2 cat #N)*350
(V v1 walks #V)*500
(V v2 runs #V)*250
```

Figure 11: SP-patterns representing grammatical structures, as discussed in the text. The number after each SP-pattern, preceded by a '⋆', is a supposed frequency of occurrence in some imaginary text or texts.

The remaining SP-patterns in Figure 11 represent words, each one with its grammatical category:

- The SP-patterns ‘`(D d1 the #D)`’ and ‘`(D d2 this #D)`’ represent word of the grammatical category ‘determiner’. In ordinary mathematics, they may also be seen as the range of possible values, and therefore a type definition of the ‘variable’ ‘`D #D`’.
- The SP-patterns ‘`(N n1 dog #N)`’ and ‘`(N n2 cat #N)`’ represent words of the category ‘noun’. They may also be seen as the ‘type definition’ of the ‘variable’ ‘`N #N`’.
- The SP-patterns ‘`(V v1 walks #V)`’ and ‘`(V v2 runs #V)`’ represent words of the category ‘verb’. They may also be seen as a ‘type definition’ of the ‘variable’ ‘`V #V`’.

When the SP Computer Model is run with the New SP-pattern, ‘`(this dog runs)`’, and with the SP-patterns in Figure 11 as Old SP-patterns, the best SP-multiple-alignment created by the program is the one shown in Figure 12. Here, “best” means that the SP-multiple-alignment in the figure is the one which, via an encoding (as described in [46, Section 4.1] and [44, Section 3.5]), yields the greatest compression of the New SP-pattern. Overall, the SP-multiple-alignment may be seen as a parsing of the sentence ‘`(this dog runs)`’ in terms of its grammatical constituents.

```
0           this         dog         runs       0
             |            |           |
1            |      N n1 dog #N       |         1
             |      |        |        |
2 S s1 D     |   #D N        #N V     |   #V #S 2
       |     |   |              |     |   |
3      D d2 this #D             |     |   |     3
                                |     |   |
4                               V v2 runs #V    4
```

Figure 12: The best SP-multiple-alignment created by the SP Computer Model using the New SP-pattern ‘`(this dog runs)`’ and the Old SP-patterns shown in Figure 11.

This example shows how, in a parsing of a simple sentence, each of the variables ‘`D #D`’, ‘`N #N`’, and ‘`V #V`’, may take on appropriate values. With other sentences, such as ‘`(the cat walks)`’ and so on, the variables would take on different values.

In summary, a superposition is similar to a syntactic class in four respects::

- ‘Superposition’ in quantum mechanics and ‘syntactic class’ in linguistics are both abstractions, without any corresponding physical structure in the world;
- They both represent two or more values;

- With both superposition and syntactic class, there are situations where one value is selected out of the two or more values in the superposition or syntactic class. In quantum mechanics, this happens when a ‘particle’ is detected or ‘measured’. With syntactic classes, this happens when a ‘variable’ is assigned a value from the relevant syntactic class.
- In a superposition and and in a syntactic class in a stochastic grammar, there is a probability or frequency associated with each value. Non-stochastic grammars in linguistics leave out this refinement, but in effect that means that all the values of any given syntactic class have the same probability.

6.3.2 Similarity between ‘superposition’ in quantum mechanics and ‘type definition’ in mathematics and computing

Much the same may be said about the similarity between the concept of superposition in quantum mechanics and the concept of a ‘type definition’ or ‘data type’ in mathematics and computing. It appears that they are similar in three respects, possibly four:

- Both ‘superposition’ in quantum mechanics and ‘type definition’ in ordinary mathematics or computing are abstract constructs without any corresponding physical structure;
- A superposition and a type definition both represent a range of possible values;
- In most applications, there are many instances of ‘variables’ where a specific value from the set of values in the type definition may be realised;
- As in superposition, probabilistic programming on computers assigns a probability or frequency to each of the values of each type definition.

6.4 Quantum computing

This section considers what if anything, elsewhere in this paper, is relevant to quantum computing, and, if so, how?

The main ideas in quantum computing are described by Al-Khalili thus:

> “... in 1985, Oxford physicist David Deutsch published a pioneering paper that showed how [quantum computing] might be achieved in practice. ... Deutsch’s machine would operate according to quantum principles to simulate any physical process. It required a row of quantum systems that could each exist in a superposition of two states, such

> as atoms in superpositions of two energy levels. These quantum systems would then be entangled together to create quantum logic gates that would be made to perform certain operations.
>
> The basic idea is that of the 'quantum bit' or qubit. In a normal digital computer, the basic component is the 'bit', a switch that can be in either of two positions: off or on. These are denoted by the binary symbols of 0 and 1. However, if a quantum system, such as an atom, is used then it could exist in the two states at once. A qubit can thus be both off and on at the same time, just as long as it can be kept isolated from its environment.
>
> Of course a single qubit is not very useful. But if we entangle two or more qubits we can start to see the power of such a set-up. Consider the information content of three classical bits. Each can be either 0 or 1 and so there are eight different combinations of the three (000, 001, 010, 100, 011, 101, 110, 111). But just three entangled qubits allow us to store all eight combinations at once! Each of the three digits is both a 1 and a 0 at the same time.
>
> Adding a fourth qubit would give us 16 combinations and a fifth, 32 and so on. The amount of information stored increases exponentially (as 2^N, where N is the number of qubits). Now imagine carrying out operations in the same way that we would with classical bits. We would be able to perform 2^N computations at once, the ultimate in parallel processing. Certain problems that might take a normal supercomputer years to solve could be cracked in a fraction of a second. [1, locations 3421–3433].

In the light of what is said in Sections 6.1 and 6.2, and in Appendix A.4.3, and in particular to avoid the mistake of viewing a wavefunction as a physical entity (Section 6.1.1), it seems necessary to view the physical substrate of a qubit as a wave, not a bullet-like particle.

6.4.1 Syntactic classes and quantum computing

If the parallels described in Section 6.3 are accepted, then in the SP-pattern '`(S s1 D #D N #N V #V #S)`' in Figure 11, the syntactic variable '`D #D`' is like a superposition of the SP-patterns '`(D d1 the #D)`' and '`(D d2 this #D)`', the syntactic variable '`N #N`' is like a superposition of the SP-patterns '`(N n1 dog #N)`' and '`(N n2 cat #N)`', and the syntactic variable '`V #V`' is like a superposition of the SP-patterns '`(V v1 walks #V)`' and '`(V v2 runs #V)`'.

Viewed in that way, the whole grammatical structure may be seen as having the same general form as the three entangled qubits in the quote in Section 6.4

above. Like those three entangled qubits, the grammatical structure has the potential to create (or 'generate' in the jargon of theoretical linguistics) eight possible sentences like '`(the dog walks)`', '`(the dog runs)`', '`(the cat walks)`', and so on, corresponding to the eight combinations, (000, 001, 010, 100, 011, 101, 110, 111), mentioned in the quote above.

6.4.2 Quantum computing in comparison with ordinary parallel processing

What is said about syntactic classes in Sections 6.3.1 and 6.4.1 suggests that, with a conventional computer working on a linguistic application, parallel processing may be applied very simply by treating each invocation of a syntactic class as an opportunity to process all the members of the class in parallel. Since any member of a syntactic class may be a sequential pattern like 'the', 'dog', or 'walks', there may also be opportunities for the application of pipeline parallel processing.

Since superposition also resembles the concept of a type definition or data type in ordinary mathematics and computing (Section 6.3.2), much the same may be said about qubits in quantum computing (but, almost certainly, pipeline parallelism would not be applicable).

Thus, we may suppose that the anticipated speedup in any quantum computer would be largely because the '`0`' and '`1`' in each qubit would be processed in parallel.

However, that possibility, and several other possibilities are discussed quite fully by Ball in a chapter in [3] headed "Quantum computers don't necessarily perform 'many calculations at once' ". In brief, the possibilities discussed by Ball include:

- That speedup in quantum computers is indeed due to the parallelism that is implicit in the concept of a qubit. [3, Location 3010].
- "Just as we can use quantum theory to correctly predict the outcomes of experiments on double-slit diffraction or Bell-test entanglement yet without being able to say exactly why, so it is clear that quantum computing works in principle but we can't say exactly why." [3, Locations 3017–3024].
- Speedup via processing in "Many Worlds", as advocated by David Deutsch. [3, Location 3024]. But, as discussed in Section 8, there are reasons to doubt the validity of that interpretation of quantum mechanics.
- Speedup via entanglement: "The computation uses the entangled relationships between qubits to manipulate them all together, without having to do many repetitive operations on each qubit individually. That can cut out a

lot of bother, because it means that you can leap between many-qubit states without having to work through intermediate steps that would have been taken on a classical computer.” [3, Locations 3024–3031].

- “Others feel that quantum-computational speed-up is more about the interference that is possible between quantum states: the fact that the probability of two quantum states is not the same as the sum of their individual probabilities.” [3, Location 3031].
- “Another candidate source [of speedup] is contextuality ... the dependence of quantum outcomes on the context of measurement.” [3, Location 3047].
- And more.

In the light of these and other aspects of quantum computing, there are reasons to doubt the feasibility of developing quantum computers that are useful:

- The range of views outlined above demonstrate many uncertainties in current thinking about quantum computing.
- The undoubted technical difficulties in making quantum computers work: “Just as in a classical computer, the 1s and 0s of the input to a quantum algorithm are marshalled into binary digits encoding solutions. The catch is that superpositions are generally very ‘delicate’. They get easily disrupted by disturbances from the surrounding environment, particularly the randomizing effects of heat. ... this doesn’t really mean—as is often implied—that superpositions are destroyed, but rather that the quantum coherence spreads into the environment, so that the original system decoheres.” [3, Location 2822].
- It appears to be generally accepted that: “There isn’t a straightforward way of making use of what quantum mechanics has to offer, and designing good quantum algorithms is a very difficult task.” [3, Location 3084]
- A report in the *Communications of the ACM* describes how a senior honors student, at the University of Texas at Austin, “discovered an algorithm that showed classical computers can indeed tackle predictive recommendations at a speed previously thought possible only with quantum computers.” [18, p. 15].
- A paper by Mikhail Dyakonov [13] argues that the astronomically-large number of “degrees of freedom” in quantum computing means that, in answer to the question “When will we have a quantum computer?” in the title of the paper, “As soon as physicists and engineers learn to control this number of degrees of freedom, which means—never!” [13, p. 4].

In view of these considerations, it seems that there is a case for devoting at least as much effort to the development of non-quantum parallel processing as is currently devoted to the development of quantum computing:

- It seems as likely as not that gains in the anticipated speed of processing, or reductions in computational complexity, in quantum computers could be matched with non-quantum parallel processing applied to knowledge structures with implicit parallelism like syntactic classes and type definitions, perhaps with pipelining applied to sequential structures within those knowledge structures.
- If it turns out that there is something special about bit-level parallel processing (something that does not seem very likely), that kind of parallelism could probably developed with non-quantum computers.
- It seems likely that engineering problems in the advancement of non-quantum computers would be more easily solved than problems such as decoherence that have proved so hard to solve with quantum computers.
- It seems likely that automatic, semi-automatic, or manual programming of non-quantum computers will prove to be less difficult than programming quantum computers.

7 Nonlocality, entanglement, SP-multiple-alignment, and discontinuous dependencies

The interrelated concepts of 'nonlocality' and 'entanglement' are described by Al-Khalili thus:

> "... in this chapter I have discussed two different and quite tricky concepts: superposition and nonlocality. The first states that a quantum particle can be in a combination of two or more states at the same time, while the second says that two quantum particles ... can somehow remain in touch with each other however far apart they are. [Combining these two ideas], the idea of two dice remaining in (nonlocal) contact with each other how ever far apart they are is known as entanglement." [1, Location 1233].

There is little doubt that the phenomena of "nonlocality" and "entanglement" are genuine features of the world and not merely some "weirdness" in quantum mechanics which may, at some stage, be explained away:

> “Today, quantum nonlocality and entanglement are no longer the subject of philosophical debate. They are accepted as crucial features of the quantum world. Indeed, entanglement of many particles could lead to the development of a whole new technology not even dreamed of by the quantum pioneers.” [1, location 1274].

7.1 A potential analogy in the processing of natural language

As with superposition (Section 6), there is a potentially useful analogy for nonlocality and entanglement in the processing of natural language.

The SP-multiple-alignment shown in Figure 3 provides an example (with simplifications of some of the details of English grammar). Here, the sentence ‘`t w o k i t t e n s p l a y`’ is identified as a sentence (defined by the SP-pattern ‘`(S ... #S)`’ in row 7), and parsed into constituents such as a noun phrase (‘`(NP ... #NP)`’ in row 4), a verb phrase (‘`(V ... #V)`’ in row 6), and so on.

For present purposes, the key point of interest is that, within sentences like this, there is a syntactic “dependency” between the ‘subject’ at the beginning (which is the noun-phrase ‘`t w o k i t t e n s`’) and the later main verb-phrase (which is the single word ‘`p l a y`’).

The rule here is that, in English at least, if the ‘subject’ noun-phrase is plural then the main or only verb-phrase must be plural, and if the subject noun-phrase is singular then the main verb-phrase must be singular. Most natural languages have dependencies like that, such as for example, gender dependencies in French which may cut across number dependencies (for more discussion, see [44, Section 5.4]).

This kind of dependency is often described as “discontinuous” because it can jump over intervening structure such as *that were only born yesterday* between the subject noun-phrase, *two kittens*, and the main verb-phrase, *play*, in *two kittens that were only born yesterday play*, and there appears to be no limit on how big that intervening structure may be.

Amongst the several ways in which discontinuous dependencies may be represented in AI systems, one of the simplest is via an SP-pattern within an SP-multiple-alignment, like the SP-patten in row 8 in Figure 3. Here, the SP-symbol ‘`NPp`’ (meaning plural noun-phrase) is aligned with a matching SP-symbol within the SP-pattern for the subject noun-phrase, ‘`(NP NPp D Dp #D N Np #N #NP)`’ in row 4, and the symbol ‘`VPp`’ (meaning plural verb-phrase) is aligned with a matching symbol within the SP-pattern for the main verb-phrase, ‘`(VP VPp Vr #Vr #VP)` in row 6.

The fact that the SP-symbols ‘`NPp`’ and ‘`VPp`’ both appear in one SP-pattern (in row 8) is what marks the dependency between the subject noun-phrase and the main verb-phrase. It is anticipated that, when the SP System is more fully developed, it will be able, via unsupervised learning, to discover that kind of recurrent pattern for itself, instead of it being provided ready-made in the store of Old patterns as is the case with the example in Figure 3.

This example suggests that insights gained in the SP System may have traction in quantum mechanics. It seems possible that a dependency between, for example, two entangled electrons, such that one electron has a clockwise spin while the other electron has a counter-clockwise spin, may be understood in a manner that is similar to our understanding of the phenomenon of syntactic dependencies in natural languages. In both cases:

- There is a correlation between the two elements of the dependency.
- The dependency may bridge arbitrarily large amounts of intervening structure.
- There is a kind of 'instant' communication in the sense that, if we know one element of a dependent pair, we know immediately what the other should be. This effect is what Einstein famously called 'spooky action at a distance' (SAD).

The kind of instant communication just mentioned—something that has been verified in many experiments—looks like communication that is faster than the speed of light and thus incompatible with a basic principle in general relativity, that nothing can travel faster than light. How can that contradiction be resolved?

7.2 How a non-local, entangled pair of entities may be regarded as a single object

A suggested answer to the question at the end of the preceding subsection is that what is normally construed as *two* entangled particles could equally equally well be seen as a single object, in the same way that the SP-pattern '`(Num PL ; NPp VPp)`' is a single object containing the two significant SP-symbols, '`NPp`' and '`VPp`'. In this case, there is no need for any communication at all, SAD or otherwise, because if we have a full knowledge of an SP-pattern, we know its contents.

Ball makes this point in a chapter in [3] headed "There is no 'spooky action at a distance' " [3, location 1973]. Without attempting to discuss all the arguments and counter-arguments that Ball considers, here is one of the more telling examples that he describes:

> “Think of a pair of gloves: one left-handed, the other right-handed. If we were to post one at random to Alice in Aberdeen and the other to Bob in Beijing ..., then the moment Alice opened the parcel and found the left glove (say), she’d know that Bob’s glove is right-handed. This is trivial, because the gloves had that handedness all the time they were in transit—it’s just that Alice and Bob didn’t know which was which until one of them looked.” [3, locations 1838–1845].

Here, the pair of gloves may be seen as a discrete entity rather than two separate entities. This is like the SP-pattern ‘`(Num PL ; NPp VPp)`’ in Figure 3 being regarded as a single entity which provides knowledge of both ‘`NPp`’ and ‘`VPp`’ without the need for communication between them.

In the same vein, a little later, Ball writes that:

> “We can’t regard particle A and particle B [that are entangled] as separate entities, even though they are separated in space. As far as quantum mechanics is concerned, entanglement makes them both parts of a single object.” [3, location 2026].

Here’s another example. Imagine a scene in which a car is partly obscured by the trunk of a tree, with the front part visible. If we see the front part move forwards, or backwards, we can infer instantly that the back of the car will be moving in the same direction and at the same speed. Of course, it could be a stage magician’s car that does something different, so the inference is probabilistic. But in this case the probabilities strongly favour the normal interpretation.

Since this kind of scene is very familiar, it would be strange indeed if people were to speak of nonlocality and entanglement between the front and back of a car! Perhaps we’ll eventually drop that kind of language when speaking or writing about quantum particles.

8 *Our Mathematical Universe*, NM, and the SP System

This section is about ideas described by physicist Max Tegmark in his book *Our Mathematical Universe* [38], and the suggested relevance of the NM and the SP System.

In the book, Tegmark argues for the “... crazy-sounding belief of mine that our physical world not only is *described* by mathematics, but that it *is* mathematics, making us self-aware parts of a giant mathematical object” (the ‘Mathematical Universe Hypothesis’, MUH). ([38, p. 6], emphasis in the original).

In connection with that idea, Tegmark describes the concepts of:

- ‘Our universe’: “The part of physical reality we can in principle observe; quantum complications aside, this is the spherical region of space from which light has had time to reach us during the 14 billion years since our Big Bang.” [38, p. 138];
- ‘Parallel universe’: “A part of physical reality that can in principle be observed from somewhere else but not from here—parallel universes are not a theory, but a prediction of certain theories.” (*ibid.*);
- ‘Multiverse’: “a collection of universes” (*ibid.*);
- ‘Level I multiverse’: “Distant regions of space that are currently but not forever unobservable; they have the same effective laws of physics but may have different histories” (*ibid.*);
- ‘Level II multiverse’: “Distant regions of space that are forever unobservable because space between here and there keeps inflating; they obey the same fundamental laws of physics, but their effective laws of physics may differ.” [38, pp. 138–139];
- ‘Level III multiverse’: “Different parts of quantum Hilbert space ...; same diversity as Level II.” [38, pp. 139];
- ‘Level IV multiverse’: “All mathematical structures ... corresponding to different fundamental laws of physics” (*ibid.*);

And Tegmark describes [38, pp. 185–191] a set of ideas with approval that were originated in 1957 by Hugh Everett III in his PhD thesis at Princeton University [16].

In brief, Everett proposed a radical alternative to Niels Bohr’s ‘Copenhagen’ interpretation of quantum mechanics. Instead of supposing a “collapse” of the wavefunction whenever someone makes an “observation” of a quantum entity, Everett supposed that the wavefunction would never collapse, regardless of any “observation”, and the alternatives that it describes (such as heads or tails for a coin flip) would, in effect, mean a splitting of the universe into two, corresponding to the heads or tails alternatives. Tegmark writes:

> “... parallel-universe splitting is happening constantly, making the number of quantum parallel universes truly dizzying. Since such splitting has been going on ever since our Big Bang, pretty much any version of history that you can imagine has actually played out in a quantum parallel universe, as long as it doesn’t violate any physical laws. This makes vastly more parallel universes than there are grains of sand in

> our Universe. In summary, Everett showed that if the wavefunction never collapses, then the familiar reality that we perceive is merely the tip of an ontological iceberg, constituting a minuscule part of the true quantum reality." [38, p. 190].

And further:

> "Let's call the quantum parallel universes that Everett discovered Level III parallel universes, and the collection of all of them the Level III multiverse. Where are these parallel universes? Whereas the Level I and Level II kinds are far away in our good old three-dimensional space, the Level III ones can be right here as far as these three dimensions are concerned, but separated from us in what mathematicians call Hilbert space, an abstract space with infinitely many dimensions where the wavefunction lives. [38, p. 190].

and "Everett's version of quantum mechanics first began to get popularized by the famous quantum-gravity theorist Bryce DeWitt, who called it the *Many Worlds* interpretation—a name that stuck." [38, p. 190].

What about Level IV multiverses? Tegmark suggests that "mathematical existence and physical existence are equivalent, so that *all* structures that exist mathematically exist physically as well." ([38, p. 321], emphasis in the original) and "there's a fourth level of parallel universes that's vastly larger than the [other three levels], corresponding to different mathematical structures." (*ibid.*).

Tegmark also writes: "According to the CUH [Computible Universe Hypothesis], the mathematical structure that is our physical reality has the attractive property of being computable and hence well defined in the strong sense that all its relations can be computed. There would thus be no physical aspects of our Universe that are uncomputable/undecidable, eliminating the concern that the work of Church, Turing and Gödel somehow makes our world incomplete or inconsistent." [38, p. 333]. In other words, the Level IV multiverse would only contain mathematics that is, metaphorically, kosher.

The subsections that follow describe how the MUH ideas may be interpreted from other viewpoints, including perspectives from the SP System and the proposed NM.

8.1 Fudging the distinction between 'abstract' and 'physical'

Many people, which probably includes many physicists, are likely to agree with Tegmark that MUH is "crazy-sounding", perhaps with the rider that it is indeed crazy.

One apparent problem is that MUH makes no provision for the way in which mathematics is often used to abstract away from any physical reality, so that an abstraction like '4' may be used to say something useful about physical entities such as apples, people, cups of tea, and so on.

What is the 'reality' of 4, 4×5, 4^8, $4!$, and the many other abstract entities in mathematics? The belief that they have some kind of 'real' existence would be much like the 'Platonism' view that mathematical entities are "numinous and transcendent entities, existing independently of both the phenomena they order and the human mind that perceives them" [20, pp. 95–96]. In keeping with that idea, Tegmark writes:

> "[Each mathematical structure] describes a physically real universe—just a different one from the one we happen to inhabit. This can be viewed as a form of radical Platonism, asserting that all the mathematical structures in Plato's 'realm of ideas' exist 'out there' in a physical sense." [38, p. 320–321].

Of course, Plato was a great thinker, but like all great thinkers, he could make mistakes. And he would probably have been the first to admit that he had not entirely 'nailed' the fundamental nature of mathematics. With the passage of time and a great deal more thought, and with the benefit of insights from researchers such as Claude Shannon [34] and Ray Solomonoff [35, 36], we can do better.

Of course, we are free to suppose that abstract structures in mathematics are themselves 'physically real' entities. But this makes nonsense of the ordinary meanings of words like 'abstract' and 'physical'. Those ordinary meanings are rooted in ideas which are in constant use everywhere: with unary arithmetic, the shepherd may count his sheep with a mark on paper or his hand for each one, in essentially the same way that a prisoner may keep track of days that have passed with marks on the wall, and so on through the many uses of numbers. From those fundamentals is built the whole abstract structure of number theory, none of which says anything about numbers of sheep, days that have passed, apples, people, and so on.

There seems to be no good reason for fudging that longstanding and extremely useful distinction between a number and what it applies to, except to salvage the extraordinary idea that mathematics can, by some obscure kind of magic, jump from being a good means of abstracting away from reality to being a reality itself.

8.2 A psychological and biological perspective

An alternative to Tegmark's view of mathematics is a psychological and biological view which derives from the SP programme of research. Here's a summary:

- *Evidence for the importance of IC in HLPC.* There is now a substantial body of evidence for the importance of IC, and more specifically ICMUP, in HLPC. Much of this is described in [53].

- *IC and natural selection.* It is clear that, for many animals including humans, IC has been and remains an important driver for natural selection [53, Section 4]:

  - *Volumes of information and size of storage space.* By allowing a creature to store more information in a given storage space or use less storage space for a given amount of information.
  - *Speed of transmission of information and bandwidth.* By speeding up the transmission of any given volume of information along nerve fibres (thus speeding up reactions), or by reducing the bandwidth needed for the transmission of the same volume of information in a given time.
  - *Inference and probabilities.* Perhaps more important than the impact of IC on the storage or transmission of information is the close connection between IC and concepts of inference and probability (Section 2.4.6), so that any animal may anticipate where it is likely to find food, how it may avoid dangers, and so on.

- *Mathematics, human brains, and IC.* Since mathematics is the product of human brains, and is an aid to human thinking, it should not be surprising to find that mathematics, in exhibiting features of ICMUP [54], chimes with the way in which human brains appear to work [46, 44], especially the importance of IC in human thinking.

In short, the SP/NM perspective on mathematics has deep roots in human psychology and human biology, things that are missing from the MUH.

8.3 Computability

The idea that all mathematical structures in the Computable Universe should be computable means that most AI algorithms, and most computer models of human cognition, would be excluded. This is because most of AI and the modelling of human cognition is about the use of heuristic techniques to find reasonably good approximate answers to problems that are are too complex to be computable (see, for example, [53, Section 2.2.2]). This seems to rule out any kind of intelligent being, including ourselves, in the Computable Universe, so that we cannot be "self-aware parts of a giant mathematical object" (Section 8).

8.4 What about compression of information?

With or without the SP Theory, it is widely accepted that, in the development of any kind of scientific theory, we should steer clear of things that increase complexity rather than reduce it. In those terms, the mind-boggling complexity of creating a whole new universe for each dichotomy that may arise, however small it may be, and this from the beginning of the world (Section 8), seems to be entirely wrong.

Since Ockham's razor is primarily a principle for theories, not physical entities, it matters little that that monstrous complexity is "... separated from us in what mathematicians call Hilbert space, an abstract space with infinitely many dimensions where the wavefunction lives." (*ibid.*). And the reference to infinity should ring alarm bells (Section 5).

It is true that the SP Computer Model, that works entirely via the compression of information, can achieve the paradoxical effect of decompressing information (Section 2.2). But that would normally be when the system is used to retrieve the uncompressed form of information that had previously been compressed. It is quite different from the colossal redundancy that would be generated in the Many Worlds view of quantum mechanics. And that colossal redundancy is totally at odds with the widely-accepted idea that scientific theories should, as far as possible, conform to Ockham's razor.

9 Conclusion

In its essentials, the proposal in this paper is that a *New Mathematics* (NM) may be created as an amalgamation of mathematics as it is today with the SP System as it is today, with future developments in both areas. It is envisaged that the amalgamation would incorporate concepts from the SP System such as SP-multiple-alignment and unsupervised learning, and that concepts in mathematics would generally be interpreted within the SP framework.

In these proposals, the NM, like the SP System, will be fundamentally probabilistic, it will exhibit versatility in the representation of diverse kinds of knowledge and versatility in diverse aspects of intelligence, and it will provide for the seamless integration of diverse kinds of knowledge and diverse aspects of intelligence in any combination.

In Section 3, several potential benefits of the NM are described briefly. They include: adding an AI dimension to mathematics; facilitating the simplification and integration of mathematics, logic, and computing; developing the NM as a *universal framework for the representation and processing of diverse kinds of knowledge*(UFK); the potential for a new perspective on statistics derived from the already recognised probabilistic nature of mathematics, and from the probabilistic nature of the SP System; and the potential for new approaches to concepts of "proof," "theorem," and related ideas.

Other possibilities are described more fully in other sections, one topic per section:

- *Alternative views of knowledge* (Section 4). There is quite a lot of evidence that leading scientists, and others, often think about phenomena in science in a way which is, in some sense, non-mathematical. There is potential for the NM to represent and process knowledge in a similar way, with relatively direct benefits in solving problems, and benefits for scientists in supporting 'non-mathematical' modes of thinking, in addition to the more familiar uses of mathematics in science.

- *Infinity* (Section 5). For some time there has been disquiet amongst scientists, especially physicists, about the way in which mathematics can all too easily suggest that something is infinitely big or infinitely small. Of course it is possible to simply ignore what the mathematics implies, but it would be more satisfactory if the 'disquiet' was backed by some good principles.

  With examples from the SP Computer Model, it is argued that the SP System provides principles that suggest why the kind of recursion that is needed for predictions of infinity, would in practice be reigned in. In brief, the arguments are:

  - Mathematical predictions that something is infinitely big or small only work when the relevant mathematics is artificially 'clean'. Since the real world is rarely as clean and tidy as mathematical abstractions from it, the introduction of some realistic 'dirt' into the calculations are likely to deflect them from any prediction that something in the world might be infinitely big or infinitely small.
  - In the kind of recursive processing that is needed for any inference that something is infinitely big or infinitely small, the gain in terms of IC remains constant throughout the many cycles but the encoding costs increase from cycle to cycle. This means that, after several cycles, there is no net benefit in terms of IC. Thus the recursive cycles are snuffed out before they reach any kind of infinity.

- *Superposition* (Section 6):

  - *Similarity of superposition with syntactic classes and with type definitions in ordinary mathematics and computing.* With an example from the SP Computer Model, it is argued that: 1) a superposition of states is like a syntactic class and also like a type definition or data type for a variable in ordinary mathematics and computing; 2) in both a superposition, a syntactic class, a type definition in probabilistic computing, there may be a probability associated with each of the alternatives for that superposition or syntactic class or type definition.

  - *Quantum computing and non-quantum parallel processing.* With reference to the same example from the SP Computer Model, it is argued that non-quantum parallel processing may be a rival to quantum computing in terms of speed of processing, without decoherence and other problems with quantum computing.

- *Nonlocality and entanglement* (Section 7). With an example from the SP Computer Model it is argued that there are similarities between the quantum mechanics concept of entanglement (with nonlocality) and the phenomenon of discontinuous dependencies in natural language. In both cases: 1) there is a correlation between two entities; 2) the relationship may bridge arbitrarily large amounts of intervening structure; and 3) there is a kind of 'instant' communication between the two entities. The gist of the argument, also made by Philip Ball [3, location 1973–2160], is that if the two entities are regarded as a single entity, worries about 'spooky action at a distance' disappear.

- *"Our Mathematical Universe"* (Section 8). This section argues against Tegmark's "... crazy-sounding belief of mine that our physical world not only is described by mathematics, but that it is mathematics, making us self-aware parts of a giant mathematical object" (MUH) [38, p. 6], and the Many-Worlds interpretation of quantum mechanics proposed by Hugh Everett III [16]. In brief, the arguments are that: 1) merging the idea of a mathematical description with the thing it describes confounds the very useful distinction between abstract and physical; 2) the 'SP' perspective on mathematics provides a psychological and biological foundation for mathematics which is missing from the MUH perspective; 3) There seems to be no place in the MUH perspective for AI or models of human cognition; 4) the Many Worlds interpretation of quantum mechanics seems to conflict with Ockham's razor.

Appendices

These appendices are for ideas that are even more tentative than the main sections, and less well developed.

A A tentative ‘tsunami’ interpretation of ‘wave–particle duality’

This appendix presents a tentative ‘tsunami’ alternative to the way in which the expression ‘wave–particle duality’ is generally understood.

To begin, Ball’s book on quantum mechanics [3] contains a fourth chapter with the provocative title “Quantum objects are neither wave nor particle (but sometimes they might as well be)” in which he says:

> “When we speak of electrons and photons, atoms and molecules, it seems perfectly reasonable to use [the word ‘particle’], and I’ll occasionally do so. Then we might have the image of a tiny little thing, a microscopic ball-bearing all hard and shiny. But probably the most widely known fact of quantum mechanics is that ‘particles can be waves’. What then becomes of our compact little balls?” [3, Location 397].
>
> “Quantum objects are not sometimes particles and sometimes waves, like a football fan changing her team allegiance according to last week’s results.” [3, Location 404].
>
> “... all we can say is that what we measure sometimes looks like what we would expect to see if we were measuring discrete little ball-like entities, while in other experiments it looks like the behaviour expected of waves of the same kind as those of sound travelling in air, or that wrinkle and swell on the sea surface. So the phrase ‘wave–particle duality’ doesn’t really refer to quantum objects at all, but to the interpretation of experiments—which is to say, to our human-scale view of things.” [3, Locations 404–492].

It is that “the interpretation of experiments” where the concept of a ‘tsunami’ may be helpful. To begin, here is a description of a tsunami as it occurs in the sea:

> “[A] **tsunami**, also called [a] **seismic sea wave** or **tidal wave**, ... [is] usually caused by a submarine earthquake, an underwater or coastal landslide, or a volcanic eruption. ....
>
> “**Origin and Development**
>
> “After an earthquake or other generating impulse occurs, a train of simple, progressive oscillatory waves is propagated great distances over the ocean surface in ever-widening circles, much like the waves produced by a pebble falling into a shallow pool. In deep water a tsunami can

> travel as fast as 800 km (500 miles) per hour. The wavelengths are enormous, about 100 to 200 km (60 to 120 miles), but the wave amplitudes (heights) are very small, only about 30 to 60 cm (1 to 2 feet). These long periods, coupled with the extremely low steepness and height of the waves, enables them to be completely obscured in deep water by normal wind waves and swell. A ship on the high seas experiences the passage of a tsunami as an insignificant rise and fall of only half a metre (1.6 feet), lasting from five minutes to an hour or more.
>
> "As the waves approach the coast of a continent, however, friction with the rising sea bottom reduces the velocity of the waves. As the velocity lessens, the wavelengths become shortened and the wave amplitudes (heights) increase. Coastal waters may rise as high as 30 metres (about 100 feet) above normal sea level in 10 to 15 minutes. The continental shelf waters begin to oscillate after the rise in sea level. Between three and five major oscillations generate most of the damage, frequently appearing as powerful "run-ups" of rushing water that uproot trees, pull buildings off their foundations, carry boats far inshore, and wash away entire beaches, peninsulas, and other low-lying coastal formations. Frequently the succeeding outflow of water is just as destructive as the run-up or even more so. In any case, oscillations may continue for several days until the ocean surface reaches equilibrium." ([24, Retrieved 2020-04-15], emphasis in the original).

A.1 How the 'tsunami' interpretation would work

In brief, here is the way in which the concept of a 'tsunami' may help in the interpretation of the expression 'wave-particle duality'.

In the double-slit experiment when only one slit is open:

- Everything that is normally regarded as a subatomic particle, or even an atom or molecule, may be seen to start out like a tsunami when it is out at sea.

- While the 'particle' is like an 'out-at-sea tsunami', it does not create much disturbance, in much the same way that "A ship on the high seas experiences the passage of a tsunami as an insignificant rise and fall of only half a metre (1.6 feet), lasting from five minutes to an hour or more."

- When the 'out-at-sea tsunami' encounters a detecting device or final screen, this is somewhat like a tsunami when it approaches a coast: "... the wavelengths become shortened and the wave amplitudes (heights) increase. Coastal waters may rise as high as 30 metres (about 100 feet) above normal sea level in 10 to 15 minutes. ...” and great damage can be done.

- In this kind of way, the ‘out-at-sea tsunami’ may be detected. As with an insect hitting a car windscreen, the ‘out-at-sea tsunami’ is destroyed leaving only a ‘splat’, which we interpret as being a ‘particle’. In the case of the insect, the ‘splat’ is parts of the insect, but in the case of the ‘out-at-sea tsunami’, the ‘splat’ may be grains of silver in a photographic plate or charges created in an ionization chamber.
- In short, the tsunami analogy suggests that the bullet-like view of subatomic particles, at least in arrangements like the double-slit experiment, is entirely an artifact of how such ‘out-at-sea tsunamis’ are detected.

Of course, what is missing from the analogy here is what would happen to a real tsunami if it were possible to arrange for something like a double slit to be placed far out at sea. In that case, we may guess that something like an interference pattern would appear on beaches further on. But, apart from the technical difficulties, any attempt to arrange such an experiment would clearly have a much lower priority than actions to protect people from the effects of any tsunami, whether it be a single big wave or something more like an interference pattern.

That said, an interference pattern may be less damaging than one big wave and, if the technical difficulties could be overcome, might conceivably provide a justification for the experiment. It is also possible that there may be natural experiments here and there where the arrangement of land and sea has the effect of creating one or more natural double-slits or the equivalent. There is more about such possibilities in Appendix B.

A.2 Some other quantum experiments

Although evidence from real-life tsunamis is not as comprehensive as we might like, the concept may still be useful in the interpretation of other aspects of wave-particle duality.

In connection with the double-slit experiment, an issue which has been the subject of much discussion and research is whether there is a wave, or a bullet-like particle, travelling through the apparatus, or whether that is not determined until a measurement is taken.

A.2.1 The release of particles one at a time

When photons or other particles are released one a time, which might favour a bullet-like view of photons, interference patterns are still observed [3, Locations

768–776]. Despite the fact that it is not feasible to do relevant experiments with real tsunamis, we may say that this result with quantum phenomena is at least consistent with the 'tsunami' view of what is going on.

A.2.2 Delayed choice experiments

'Delayed choice' variants of double-slit experiments have been conceived and performed with the aim of throwing light on forms of quantum entities—'wave' or 'bullet-like particle'—as they pass through the stages of the double-slit experiment. In many variants, the results seem always to favour the view described by John Archibald Wheeler here:

> "The double-slit experiment ... imposes a choice between complementary modes of observation. In each experiment we have found a way to delay that choice of type of phenomenon to be looked for up to the very final stage of development of the phenomenon, *whichever* type we then fix upon. That delay makes no difference in the experimental predictions. On this score everything we find was foreshadowed in that solitary and pregnant sentence of Bohr [5, p. 370], ' ... it ... can make no difference, as regards observable effects obtainable by a definite experimental arrangement, whether our plans for constructing or handling the instruments are fixed beforehand or whether we prefer to postpone the completion of our planning until a later moment when the particle is already on its way from one instrument to another.' " ([41, pp. 39–40], emphasis in original).

or more briefly: "No elementary quantum phenomenon is a phenomenon until it is a registered phenomenon." [42, p. 399].

A.3 'Quantum states determined by the process of detection' compared with 'the tsunami interpretation of quantum phenomena'

The idea that a given state does not exist until we apply a detection or measuring device is somewhat stronger than the 'tsunami' interpretation:

- In the first case we are saying that the phenomenon does not exist until measured or detected.
- In the latter case we are saying that, before the final stage, the phenomenon is like a tsunami that is far out to sea and that, later, anything that looks like a particle is an artifact of the detection device.

The fact that we get interference patterns when particles are released one at a time to pass through the double-slit apparatus, seems to be strong evidence that they are waves, not bullet-like objects. That suggests that the Wheeler-Bohr interpretation is too strong and that the tsunami analogy is a better fit with what is observed.

A.4 Other issues relating to the tsunami interpretation of quantum phenomena

This section considers other issues that relate to the tsunami interpretation described above.

A.4.1 Einstein's view

> "It must have been around 1950. I was accompanying Einstein on a walk from The Institute for Advanced Study to his home, when he suddenly stopped, turned to me [Abraham Pais], and asked me if I really believed that the moon exists only if I look at it.' " [25, p. 5].

What Einstein said to Abraham Pais is very much in keeping with the puzzlement that most people experience when hearing that, in the 'Copenhagen' view of quantum mechanics, a quantum entity does not exist until it is observed.

None of this is relevant to the moon's existence because the way we perceive the moon, or anything that is bigger than quantum size, is fundamentally different from the way a subatomic particle may be detected.

Of course it makes no sense to suggest that the moon exists only if Einstein or anyone else is looking at it. At the same time, it makes a certain amount of sense to say, apart from the conclusions of Section A.3, that a subatomic particle comes into existence when it meets a detection device.

But it makes even more sense to say that: *at all stages throughout a double-slit apparatus except the last, a subatomic particle exists like a tsunami when it is far out to sea, and, finally, it creates a human-detectable artifact that looks like a particle when it encounters a detection device*; and, for particles that are stationary or travelling only slowly (Appendix A.4.3), it seems most plausible that they should be understood primarily as waves.

A.4.2 The detection of particles in motion

In keeping with the 'tsunami' view described above, we may suppose that, in a wire chamber, spark chamber, bubble chamber, cloud chamber, or equivalent device, any 'particle' in the form of an 'out-at-sea tsunami' would initially be travelling at high speed and that, as it enters the detection device, it would gradually slow down, at the same time giving up its energy to form ions or other entities that may be interpreted as the track of a particle.

In terms of the 'tsunami' model of wave-particle duality, this would be somewhat like a tsunami wave coming to a gently-sloping shore, travelling far inland, and destroying weaker trees, less robust buildings, leaving debris, and other things that may mark its path as it goes.

A.4.3 Particles in association with other particles

What about particles such as atoms in more-or-less stable associations with other particles, in such things as molecules, metals in solid form, and crystals?

Here, it appears that 'particles' cannot be artifacts created by hitting a detector or screen in a double-slit apparatus, and they cannot be artifacts created as a 'tsunami' wave travels through a cloud chamber or equivalent device.

If the 'particle-as-artifact' analogy is to be consistent, it seems that the most reasonable interpretation of particles in molecules, solid metals, or crystals, is that they are waves without anything to convert them into particle-like artifacts. This should not be controversial since it is widely recognised that atoms, electrons, protons, and neutrons can behave like waves.

B Interference fringes with real tsunamis?

In connection with the possibility of experiments with real tsunamis (mentioned at the end of Appendix A.1), it is not necessary for there to be literally two 'slits' between bodies of land. A single island would do, corresponding with the single "slip of card" that Thomas Young used to block a beam of sunlight coming through a pinhole in a piece of thick paper covering a hole in a "window-shutter" of a darkened room [2, Location 180].[7]

With Young's arrangement, he demonstrated the wave-like nature of light because interference fringes were seen on the wall behind the slip of card, something that could only happen if the light flowed around the edges of the slip of card as a wave would do, but not if the light was a succession of bullet-like particles.

In the light of these points, a natural experiment was provided by the 2004 Indian Ocean tsunami, created by an earthquake with its epicentre off the west coast of northern Sumatra, Indonesia, and travelling westwards across the Indian ocean. This would be somewhat like Young's beam of light, with Madagascar as

---

[7] Although Anil Ananthaswamy does not give the source of information about Young's early experiments with light, it appears that it was probably from *The Last Man Who Knew Everything* by Andrew Robinson [31].

his “slip of card” and the coast of Mozambique as the wall behind the slip of card. It would be interesting to know whether anything like an interference fringe was seen on that coast.

Some weak evidence here is that the Wikipedia page about “Countries affected by the 2004 Indian Ocean earthquake and tsunami”[8] shows that Madagascar suffered “damage only” but Mozambique is absent from the page suggesting that it was relatively unaffected by the tsunami. Thus Mozambique may have been largely protected by Madagascar, something that is at least consistent with the occurrence of an interference fringe on the coast of Mozambique, if, for example, the waves from such an interference fringe were too small too excite comment.

References

[1] J. Al-Khalili. *Quantum: A Guide for the Perplexed.* Phoenix, London, Kindle edition, 2012.

[2] A. Ananthaswamy. *Through Two Doors at Once.* Penguin Books, London, Kindle edition, 2018.

[3] P. Ball. *Beyond Weird.* Vintage, London, Kindle edition, 2018.

[4] J. D. Barrow. *Pi in the Sky.* Penguin Books, Harmondsworth, 1992.

[5] N. Bohr. Discussion with Einstein on epistemological problems in atomic physics. In P. A. Schilpp, editor, *Albert Einstein: Philosopher-Scientist*, volume VII of *The Library of Living Philosophers*, pages 201–241. Southern Illinois University Carbondale, Evanston, Illinois, 1949.

[6] J. Brockman. *This Idea Must Die.* HarperCollins, London, Kindle edition, 2015.

[7] G. J. Chaitin. Randomness in arithmetic. *Scientific American*, 259(1):80–85, 1988.

[8] G. J. Chaitin. *Proving Darwin: Making Biology Mathematical.* Pantheon Books, New York, Kindle edition, 2012.

[9] C. Darwin. *The Origin of Species (illustrated).* Murray, London, 1872. Republished by Kindle.

[10] J. W. Dauben. *Georg Cantor: His Mathematics and Philosophy of the Infinite.* Princeton University Press, Princeton, New Ed edition, 1990.

---

[8] bit.ly/2KWeMWi, Retrieved 2020-05-05.

[11] E. Davis and G. Marcus. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Communications of the ACM*, 58(9):92–103, 2015.

[12] P. A. M. Dirac. *The Principles of Quantum Mechanics.* Snowball Publishing, Northern Illinois, fourth, Kindle edition, 1947.

[13] M. Dyakonov. The case against quantum computing. *IEEE Spectrum*, 15 Nov 2018, 2018.

[14] A. Einstein. *Relativity: the Special and General Theory, a Popular Exposition.* Methuen & Co. Ltd., London, third edition, 1920.

[15] B-G. Englert. On quantum theory. *arXiv*, page 16, 2013. arXiv:1308.5290. Later published in Englert, B. On quantum theory. *The European Physical Journal D* volume 67, 238 (2013). https://doi.org/10.1140/epjd/e2013-40486-5.

[16] H. Everett. *The Many-Worlds Interpretation of Quantum Mechanics.* Princeton University Press, Princeton, 1973.

[17] M. Ford. *Architects of Intelligence: the Truth About AI From the People Building It.* Packt Publishing, Birmingham, UK, Kindle edition, 2018.

[18] S. Greengard. The algorithm that changed quantum machine learning. *Communications of the ACM*, 62(8):15–17, 2019.

[19] S. Hawking and L. Mlodinow. *The Grand Design.* Transworld Publishers, London, Kindle edition, 1988.

[20] R. Hersh. *What is Mathematics Really?* Vintage, London, 1997.

[21] D. Hofstadter. *Gödel, Escher, Bach: an Eternal Golden Braid.* Penguin Books, Harmondsworth, 1980.

[22] W. Isaacson. *Einstein: His Life and Universe.* Pocket Books, London, Kindle edition, 2007.

[23] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Springer, New York, 4th edition, 2019.

[24] The Editors of the Encyclopaedia Britannica. Tsunami water wave. In *Encyclopaedia Britannica.* Encyclopaedia Britannica, 2020.

[25] A. Pais. *'Subtle is the Lord ... ': The Science and the Life of Albert Einstein.* Oxford University Press, Oxford, 1982.

[26] V. Palade and J. G. Wolff. A roadmap for the development of the 'SP Machine' for artificial intelligence. *The Computer Journal*, 62:1584–1604, 2019. https://doi.org/10.1093/comjnl/bxy126, arXiv:1707.00614, bit.ly/2tWb88M.

[27] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, revised second printing edition, 1997.

[28] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, Cambridge, 2000.

[29] J. Pearl and D. Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Allen Lane, London, Kindle edition, 2018.

[30] E. L. Post. Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics*, 65:197–268, 1943.

[31] A. Robinson. *The Last Man Who Knew Everything*. Plume (Penguin Group), London, 2007.

[32] C. Rovelli. *Reality Is Not What It Seems: The Journey to Quantum Gravity*. Penguin Books, London, Kindle edition, 2016.

[33] K. Sayood. *Introduction to Data Compression*. Morgan Kaufmann, Amsterdam, 2012.

[34] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[35] R. J. Solomonoff. A formal theory of inductive inference. Parts I and II. *Information and Control*, 7:1–22 and 224–254, 1964.

[36] R. J. Solomonoff. The discovery of algorithmic probability. *Journal of Computer and System Sciences*, 55(1):73–88, 1997.

[37] M. Tegmark. Infinity. In J. Brockman, editor, *This Idea Must Die*. HarperCollins, London, Kindle edition, 2015. Locations 801–846.

[38] M. Tegmark. *Our Mathematical Universe*. Penguin Books, London, Kindle edition, 2015.

[39] M. J. G. Veltman. *Facts and Mysteries in Elementary Particle Physics*. World Scientific Publishing, London, Kindle edition, 2003.

[40] C. S. Wallace and D. M. Boulton. An information measure for classification. *Computer Journal*, 11(2):185–195, 1968.

[41] J. A. Wheeler. The “past” and the “delayed-choice” double-slit experiment. In A. R. Marlow, editor, *Mathematical Foundations of Quantum Theory*, pages 9–48. Academic Press, San Diego, Ca., 1978. This chapter has been reprinted is several places, e.g. L. M. Dolling, A. F. Gianelli, and G. N. Statilem, (Eds.), *Readings in the Development of Physical Theory*, p. 486ff.

[42] J. A. Wheeler. “on recognizing ‘law without law,’ ”. *American Journal of Physics*, 51:398–404, 1983.

[43] M. White and J. Gribbin. *Stephen Hawking: A Life in Science*. Abacus, London, Kindle edition, 2002.

[44] J. G. Wolff. *Unifying Computing and Cognition: the SP Theory and Its Applications*. CognitionResearch.org, Menai Bridge, 2006. ISBNs: 0-9550726-0-3 (ebook edition), 0-9550726-1-1 (print edition). Distributors, including Amazon.com, are detailed on bit.ly/WmB1rs.

[45] J. G. Wolff. Towards an intelligent database system founded on the SP theory of computing and cognition. *Data & Knowledge Engineering*, 60:596–624, 2007. arXiv:cs/0311031 [cs.DB], bit.ly/1CUldR6.

[46] J. G. Wolff. The SP Theory of Intelligence: an overview. *Information*, 4(3):283–341, 2013. arXiv:1306.3888 [cs.AI], bit.ly/1NOMJ6l.

[47] J. G. Wolff. Application of the SP Theory of Intelligence to the understanding of natural vision and the development of computer vision. *SpringerPlus*, 3(1):552–570, 2014. arXiv:1303.2071 [cs.CV], bit.ly/2oIpZB6.

[48] J. G. Wolff. Autonomous robots and the SP Theory of Intelligence. *IEEE Access*, 2:1629–1651, 2014. arXiv:1409.8027 [cs.AI], bit.ly/18DxU5K.

[49] J. G. Wolff. Big data and the SP Theory of Intelligence. *IEEE Access*, 2:301–315, 2014. arXiv:1306.3890 [cs.DB], bit.ly/2qfSR3G. This paper, with minor revisions, is reproduced in Fei Hu (Ed.), *Big Data: Storage, Sharing, and Security*, Taylor & Francis LLC, CRC Press, 2016, Chapter 6, pp. 143–170.

[50] J. G. Wolff. Information compression, multiple alignment, and the representation and processing of knowledge in the brain. *Frontiers in Psychology*, 7:1584, 2016. arXiv:1604.05535 [cs.AI], bit.ly/2esmYyt.

[51] J. G. Wolff. The SP Theory of Intelligence: its distinctive features and advantages. *IEEE Access*, 4:216–246, 2016. arXiv:1508.04087 [cs.AI], bit.ly/2qgq5QF.

[52] J. G. Wolff. Software engineering and the SP Theory of Intelligence. Technical report, CognitionResearch.org, 2017. Submitted for publication. arXiv:1708.06665 [cs.SE], bit.ly/2w99Wzq.

[53] J. G. Wolff. Information compression as a unifying principle in human learning, perception, and cognition. *Complexity*, 2019:38 pages, February 2019. Article ID 1879746. viXra:1707.0161v3, hal-01624595 v2.

[54] J. G. Wolff. Mathematics as information compression via the matching and unification of patterns. *Complexity*, 2019:25, 2019. Article ID 6427493, Archives: vixra.org/abs/1912.0100 and hal.archives-ouvertes.fr/hal-02395680.

[55] J. G. Wolff. Unsolved problems in AI, described in the book *architects of intelligence* by Martin Ford, and how they may be solved via the SP System. 2019. Submitted for publication. Details of Ford's book are in [17]. bit.ly/2th7Bze.