# Biochemistry provides inspiration for a new kind of AI

J Gerard Wolff*

May 17, 2021

**Abstract**

This article is about the origin, development, and benefits of the *SP System* (SPS), which means the *SP Theory of Intelligence* and its realisation in the *SP Computer Model* (SPCM). The SPS is radically different from deep neural networks (DNNs), with many advantages compared with DNNs. As will be described, *the SPS provides a promising foundation for the development of human-like broad AI*. The SPS was inspired in part by: evidence for the importance of information compression in human learning, perception, and cognition; and the concept of 'multiple sequence alignment' in biochemistry. That latter concept led to the development of the powerful concept of *SP-multiple-alignment*, a concept which is largely responsible for the intelligence-related versatility of the SPS. The main advantages of the SPS are: 1) The clear potential of the SPS to solve 19 problems in AI research; 2) Versatility of the SPS in aspects of intelligence, including unsupervised learning, and several forms of reasoning; 3) Versatility of the SPS in the representation and processing of knowledge; 4) Seamless integration of diverse aspects of intelligence and diverse forms of knowledge, in any combination, a kind of integration that appears to be necessary in any artificial system that aspires to the fluidity and adaptability of the human mind; 5) Several other potential benefits and applications of the SPS. It is envisaged that the SPCM will provide the basis for the development of a first version of the *SP Machine*, with high levels of parallel processing and a user-friendly user interface. All software in the SP Machine would be open-source so that clones of the SP Machine may be created anywhere by individuals or groups, to facilitate further research and development of the SP System.

---

*Dr Gerry Wolff, BA (Cantab) PhD (Wales) CEng MBCS MIEEE, ORCID ID 0000-0002-4624-8904. CognitionResearch.org, Menai Bridge, UK; jgw@cognitionresearch.org; +44 (0) 1248 712962; +44 (0) 7746 290775; *Twitter*: @gerrywolff65; *Web*: www.cognitionresearch.org.

# Contents

# 1   Introduction

The currently popular 'deep neural networks' (DNNs)—rather loose approximations of real brains—have some impressive achievements but significant shortcomings as foundations for the development of human-like broad AI [9].

This article is about the origin, development, and advantages of the *SP System* (SPS), which means the SP Theory of Development and its realisation in the SP Computer Model (SPCM).

The SPS is radically different from DNNs, with many advantages compared with DNNs. As will be described, *the SPS provides a promising foundation for the development of human-like broad AI.*

A potentially useful spin-off from this research is the discovery that much of mathematics, perhaps all of it, may be understood as information compression (IC) via the matching and unification of patterns [29]. There is potential here for the development of a *New Mathematics* with many potential benefits and applications [29, Section 9].

## 1.1   Presentation

The next section describes some of the thinking that led to the research. After that, the main sections are as follows: how the SPS has been developed (Section 3); a major breakthrough in the development of the SPS (Section 4); the clear potential for the SPS to solve 19 problems in AI research (Section 5); intelligence-related strengths and potential of the SPS (Section 6); some other potential benefits and applications of the SPS (Section 7); and planned future developments of the SPS (Section 8).

The SPS itself is outlined in Appendix A, some distinctive features of the SP research are summarised in Appendix B, and abbreviations used in this paper are detailed in Appendix C.

## 1.2   The name 'SP'

Since people often ask about the origins of the name 'SP', reasons are given in Sections B.2 and 2.3. But it is intended that 'SP' should be treated as a name, without expanding the letters in the name—like 'IBM' or 'BBC'.

## 1.3   Further information

The SPS, with many examples from the SPCM, are described in the book [16].[1]

---

[1]Information about where the printed and electronic versions of *Unifying Computing and Cognition* may be obtained are detailed on bit.ly/WmB1rs. Some book sellers may say that the

A shortened version of the book is in the paper [18].

Peer-reviewed papers about this research may be downloaded via links from www.cognitionresearch.org/sp.htm. There is also a shorter list of "Key publications, with notes" (https://tinyurl.com/4urby74e).

Source code and Windows executable code for the SPCM may be obtained via links under the heading "SOURCE CODE" on the CognitionResearch.org website (tinyurl.com/ytpa8dj9).

# 2 Inspirations for the research

The main inspirations for this research are described in the following subsections.

## 2.1 Information compression as a unifying principle in brains and nervous systems

Developing the SPS has been strongly influenced by lectures from Horace Barlow that I heard when I was a student at Cambridge University—about the idea that much of the workings of brains and nervous systems may be understood as IC.

For example, the optic nerve is too small to cope with the flood of information falling on the retina, and there is evidence that nerves in the retina have the effect of removing unnecessary repetition or *redundancy* from that information, and thus compressing it.

Ideas in this area were pioneered by Attneave [1, 2] and Barlow [3, 4] and others. From that seminal work, research around this general theme has continued up to the present. A recent review of relevant evidence is in [28].

These ideas suggest the possibility that intelligence might be understood as compression of information. I did not know it at the time I started on the SP research but Barlow had written earlier that:

> "... the operations required to find a less redundant code have a rather fascinating similarity to the task of answering an intelligence test, finding an appropriate scientific concept, or other exercises in the use of inductive reasoning. Thus, redundancy reduction may lead one towards understanding something about the organization of memory and intelligence, as well as pattern recognition and discrimination." [4, p. 210].

Because compression of a body of information may be seen as a process of increasing the *Simplicity* of that information whilst retaining as much as possible

---

print edition of the book is out of print, but this is wrong. Book retailers may obtain the print edition via 'print-on-demand' from "INGRAM Lightning Source" and, via that technology, it will never go out of print.

of its descriptive or explanatory *Power*, the name that has been adopted for this programme of research is *SP*.

## 2.2   Models of unsupervised language learning

The development of the SPS has benefitted from earlier research developing computer models of the unsupervised learning of language [14]. In particular, that research benefitted substantially from the idea that, to a large extent, the unsupervised learning of language may be seen as IC.

## 2.3   The potential for simplification and integration across a broad canvass

Working as a software engineer with a software company[2] set me thinking about the possibility of simplifying and integrating ideas across different areas. This relates to an issue in language-learning research: how to integrate syntax and semantics.

So an overarching goal of the SP research has been to see *whether or how it might be possible to simplify and integrate observations and concepts across AI, mainstream computing, mathematics, and human learning, perception, and cognition.*

The broad reach of this goal chimes with the exhortation, attributed to Dwight D. Eisenhower, that "If you can't solve a problem, enlarge it," meaning that putting a problem in a broader context may make it easier to solve. Good solutions to a problem may be hard to see when the problem is viewed through a keyhole but become visible when the door is opened.

In truth, I started with less exalted aims than simplification and integration across a broad canvass, but that aim took shape slowly in my mind, because of similarities amongst observations and concepts in AI, in mainstream computing, in mathematics, and in human learning, perception, and cognition (HLPC).

By simplifying and integrating across those areas one would, in effect, be developing a theory of intelligence. And if the work was done via the development of computer models, one would, in effect, be developing an AI.

Because simplification and integration of observations and concepts may be seen as a process of increasing the *Simplicity* of those observations and concepts whilst retaining as much as possible of their descriptive and explanatory *Power*, this is a second reason for the name *SP*.

---

[2]Praxis Systems in Bath, UK.

## 2.4 Biochemistry provides inspiration for new thinking about artificial intelligence

The bioinformatics technique for finding good full or partial matches between two sequences of symbols led me to look at a related technique that biochemists use: how to find good full or partial alignments between two *or more* sequences.

An example of that kind of 'multiple sequence alignment', with five DNA sequences, is shown in Figure 4 in Appendix A.3.1. This is considered to be a 'good' multiple sequence alignment because it contains a relatively large number of hits between matching pairs of symbols.

The concept of multiple sequence alignment led to a major discovery, described in Section 4.

# 3 Development of the SP Computer Model

The SP Theory of Intelligence has been developed largely via the development of the SPCM. This has not only provided a means of reducing vagueness in the theory and demonstrating what the theory can do, but, via the creation of many versions of the SPCM, it has been a powerful means of fleshing out the theory itself.

Following a period of brainstorming to settle on a path to follow, it has taken several years to bring the SPCM, and the SP Theory, to their relative maturity now. Without me wishing to claim the status of 'genius', it seems appropriate here to quote Thomas Edison's saying that "Genius is one percent inspiration and ninety-nine percent perspiration."

Why was there so much 'perspiration' in the SP research? Most of the work was developing *hundreds* of different versions of the SPCM, each one developing a possible idea for how the SPCM might work, applying relevant tests, taking notes on the results, and exploring ideas for what might be tried next.

At later stages, when the shape of the SPCM had settled down, there was more work in exploring how the SPCM might exhibit different aspects of intelligence. Later again, there was the exploration of other potential benefits and applications of the SPS, including several outside the realm of AI.

# 4 A major breakthrough: the discovery of the powerful concept of SP-multiple-alignment

Looking at multiple sequence alignments like the one in Figure 4 led to the discovery of the powerful concept of *SP-multiple-alignment* (SPMA).

This discovery began with the realisation that *a modified version of the multiple sequence alignment technique could serve to model several different aspects of intelligence.*

## 4.1 Development

Although that insight only took a few hours to take shape, a much longer period of development—about 17 years—has been needed to develop the SPMA concept within the SPCM to its relatively mature state now, to explore its range of intelligence-related applications, and to write a book about the research (*Unifying Computing and Cognition* [16]).

Without me wishing to claim the status of 'genius', it seems appropriate here to quote Thomas Edison's saying that "Genius is one percent inspiration and ninety-nine percent perspiration."

Why was there so much 'perspiration' in the SP research? Most of the work was developing *hundreds* of different versions of the SPCM, each one developing a possible idea for how the SPMA might work within the SPCM, applying relevant tests, taking notes on the results, and exploring ideas for what might be tried next.

At later stages, when the shape of the SPCM had settled down, there was a lot more work in exploring how the SPCM might exhibit different aspects of intelligence, and writing the book. Later again, after a break of nearly seven years, there was more work exploring potential benefits and applications of the SPS, including some outside the realm of AI.

## 4.2 Example

The SPMA concept is described in Appendix A.3. Figure 5 in that appendix shows an example.

In the figure, row 0 shows an *SP-pattern* (a sequence of *SP-symbols*) that is designated 'New' because we imagine that it has just been received from its environment via the system's 'senses', as illustrated in Figure 3. Clearly, in Figure 5, it corresponds with the sentence "Two kittens play".

Each of rows 1 to 8 in the SPMA shows a single SP-pattern representing a grammatical structure which is designated 'Old' because it comes from the system's store of Old SP-patterns in the system's 'head' (Figure 3).

The whole structure has been built by the SPCM via heuristic search in several stages, weeding out the bad partial structures at the end of each stage and retaining the good structures.

The structure in Figure 5 is considered 'good' because it enables the New SP-pattern to be encoded economically in terms of the Old SP-patterns in the

alignment. How that encoding is done is described in [16, Section 3.5] and [18, Section 4.1].

In Figure 5, the whole SPMA may be seen to achieve the effect of parsing the sentence into its parts and subparts, much as is done in linguistics, except that linguists use trees for parsing, not SP-multiple-alignments.

The way that the SPMA in Figure 5 imitates the kind of tree used by linguists is just one of many examples of the versatility of SPMAs in modelling other structures and their workings (Section 6).

In future work, it is intended that the concept of 'SP-pattern' will be generalised to include two-dimensional arrays of SP-symbols. This should extend the scope of the SPS to the representation and processing of two-dimensional pictures and diagrams and, as noted in Appendix A.1, three-dimensional structures as well.

# 5 The potential of the SPS to solve 19 problems in AI research

Strong support for the SPS arose from a reading of the book *Architects of Intelligence* by science writer Martin Ford [6]. To prepare for the book, he interviewed several leading researchers and entrepreneurs in AI to hear their views about AI research, including opportunities and problems in the field:

> "The purpose of this book is to illuminate the field of artificial intelligence—
> as well as the opportunities and risks associated with it—by having
> a series of deep, wide-ranging conversations with some of the world's
> most prominent AI research scientists and entrepreneurs." Martin Ford
> [6, p. 2].

In the book, Ford reports what the AI experts say, giving them the opportunity to correct errors he may have made so that the text is a reliable description of their thinking.

This source of information has proved to be very useful in defining problems in AI research that experts in AI deemed to be in need of solutions. This has been significant from the SP perspective because, with 17 of those problems and two others—19 in all—there is clear potential for the SPS to provide a solution. This is not a 'discovery' in the ordinary sense of that word but it demonstrates that the SPS deserves attention.

The paper [31] describes those 17 problems, with two others, and how the SPS may solve them.

# 6 Strengths and potential of the SPS in AI

The strengths and potential of the SPS in AI-related functions and structures are summarised in the subsections that follow. Further information may be found in [18, Sections 5 to 12], [16, Chapters 5 to 9], and in other sources referenced in the subsections that follow.

## 6.1 Versatility in aspects of intelligence

The SPS has strengths and potential in the following aspects of intelligence: unsupervised learning; the analysis and production of natural language; pattern recognition that is robust in the face of errors in data; pattern recognition at multiple levels of abstraction; computer vision [19]; best-match and semantic kinds of information retrieval; several kinds of reasoning (next subsection); planning; and problem solving. It seems likely that there are more insights to come.

## 6.2 Versatility in reasoning

Kinds of reasoning exhibited by the SPS include: 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; 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 ([16, Chapter 7], [18, Section 10]).

There is also potential in the system for spatial reasoning [20, Section IV-F.1], and for what-if reasoning [20, Section IV-F.2].

As noted in Appendix A.5, the probabilistic nature of the SPS makes it relatively straightforward to calculate absolute or conditional probabilities for results from the system, as for example in its several kinds of reasoning, most of which would naturally be classed as probabilistic. In that connection, probabilities for inferences may be calculated as described in [16, Section 3.7] and [18, Section 4.4].

## 6.3 Versatility in the representation and processing of knowledge

Although SP-patterns are not very expressive in themselves, they come to life in the SPMA framework. Within that framework, they provide relevant knowledge for each aspect of intelligence mentioned in Section 6.1, and for each kind of reasoning mentioned in Section 6.2, and more. More specifically, they may serve in the representation and processing of such things as: 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 [17, Sections 3 and 4]; relational tuples (*ibid.*, Section 3), and concepts in mathematics, logic, and computing, such as 'function', 'variable', 'value', 'set', and 'type definition' ([16, Chapter 10], [22, Section 6.6.1], [25, Section 2]).

As previously noted (Section A.1), the addition of two-dimensional SP patterns to the SPCM is likely to expand the capabilities of the SPS to the representation and processing of structures in two-dimensions and three-dimensions, and the representation of procedural knowledge with parallel processing.

## 6.4 The seamless integration of diverse aspects of intelligence, and diverse kinds of knowledge, in any combination

An important additional feature of the SPS, alongside its versatility in aspects of intelligence, including diverse forms of reasoning, and its versatility in the representation and processing of diverse kinds of knowledge, is that *there is clear potential for the SPS to provide for the seamless integration of diverse aspects of intelligence and diverse forms of knowledge, in any combination.* This is because those several aspects of intelligence and several kinds of knowledge all flow from a single coherent and relatively simple source: the SPMA framework.

It appears that this kind of seamless integration is necessary in any artificial system that aspires to human-like broad intelligence.

Figure 1 shows schematically how the SPS, with SPMA centre stage, exhibits versatility and integration.
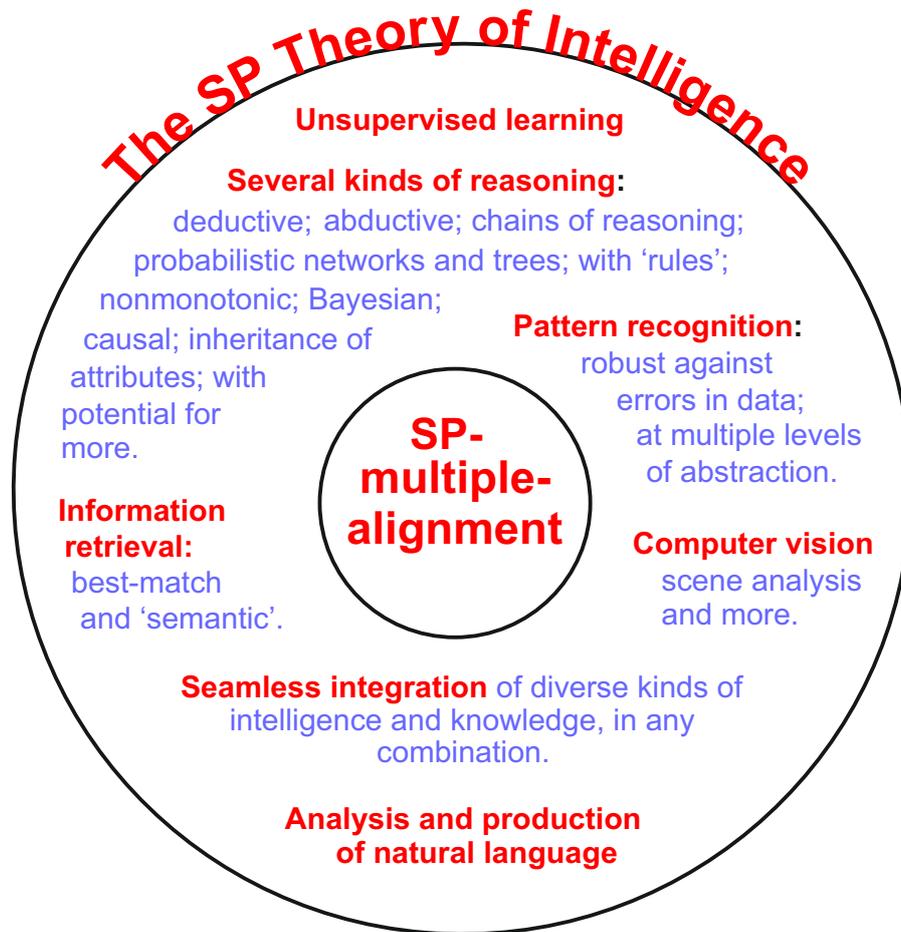
Figure 1: A schematic representation of versatility and integration in the SPS, with the SPMA concept centre stage.

# 7 Other potential benefits and applications of the SPS

Apart from its strengths and potential in intelligence-related functions (Section 6), the SPS has several other potential benefits and applications, several of them not closely related to AI. Peer-reviewed publications about those potential benefits and application (marked 'PRP'), and some others, are listed here:

- *Overview of potential benefits and applications.* The SPS has potential in several other areas of application including those described in [22, PRP]:

the simplification and integration of computing systems; best-match and semantic forms of information retrieval; software engineering [25]; the representation of knowledge, reasoning, and the semantic web; information compression; bioinformatics; the detection of computer viruses; and data fusion.

- *Autonomous robots.* The SPS opens up a radically new approach to the development of intelligence in autonomous robots [20, PRP].

- *Big data.* Somewhat unexpectedly, it has been discovered that the SPS has potential to help solve nine significant problems associated with big data [21, PRP]. These are: overcoming the problem of variety in big data; the unsupervised learning of structures and relationships in big data; interpretation of big data via pattern recognition, natural language processing; the analysis of streaming data; compression of big data; model-based coding for the efficient transmission of big data; potential gains in computational and energy efficiency in the analysis of big data; managing errors and uncertainties in data; and visualisation of structure in big data and providing an audit trail in the processing of big data.

- *Commonsense reasoning and commonsense knowledge.* Largely because of research by Ernest Davis and Gary Marcus (see, for example, [5]), the challenges in this area of AI research are now better known. Preliminary work shows that the SP System has promise in this area [26, 27].

- *An intelligent database system.* The SPS has potential in the development of an intelligent database system with several advantages compared with traditional database systems [17, PRP].

  In this connection, the SPS has potential to add several kinds of reasoning and other aspects of intelligence to the 'database' represented by the World Wide Web, especially if the SP Machine were to be supercharged by replacing the search mechanisms in the foundations of the SP Machine with the high-parallel search mechanisms of any of the leading search engines.

- *Mathematics.* The concept of ICMUP provides an entirely novel interpretation of mathematics [29, PRP]. This interpretation is quite unlike anything described in existing writings about the philosophy of mathematics.

  This new interpretation may provide the foundation for the development of a *New Mathematics* with many potential benefits and applications [29, Section 9].

- *Medical diagnosis.* The SPS may serve as a vehicle for medical knowledge and to assist practitioners in medical diagnosis, with potential for the automatic or semi-automatic learning of new knowledge [15, PRP].

13

- *Natural language processing.* The SP System has considerable strengths and potential in the processing of natural language ([18, Section 8, PRP], [16, Chapter 5], [26]).

- *Software engineering.* The SP System has potential in several aspects of software engineering [25].

- *SP-Neural.* As outlined in Appendix A.7, abstract concepts in the SP Theory of Intelligence map quite well into concepts expressed in terms of neurons and their interconnections in a version of the theory called *SP-Neural* [23, PRP]. This has potential to illuminate aspects of neuroscience and to suggest new avenues for investigation.

- *Sustainability.* The SP System has clear potential for substantial reductions in the very large demands for energy of standard deep neural networks, and applications that need to manage huge quantities of data such as the Square Kilometre Array [30, PRP]. Where those demands are met by the burning of fossil fuels, there would be corresponding reductions in the emissions of $CO_2$.

- *Transparency in computing.* By contrast with applications with deep neural networks, the SP System provides a very full and detailed audit trail of all its processing, and all its knowledge may be viewed. Also, there are reasons to believe that, when the system is more fully developed, its knowledge will often be structured in forms that are familiar such as class-inclusion hierarchies, part-whole hierarchies, run-length coding, and more. Strengths and potential of the SP System in these area are described in [32, PRP].

- *Vision, both artificial and natural.* The SPS opens up a new approach to the development of computer vision and its integration with other aspects of intelligence. It also throws light on several aspects of natural vision [19, PRP].

# 8  Future developments

It is envisaged that the SPCM will provide the foundation for the development of a first version of an *SP Machine*, with high levels of parallel processing and with an improved user interface.

A 'roadmap' for the development of the SP Machine, including some 'unfinished business' outlined in [18, Section 3.3], is described in [10].

The SP Machine and how it may be developed and applied is shown schematically in Figure 2.

Figure 2: Schematic representation of the development and application of the SP Machine. Reproduced from Figure 2 in [18], with permission.

All software would be open-source so that clones of the SP Machine may be set up, each one a vehicle for further research by individuals or groups of researchers, anywhere in the world.

# 9   Conclusion

*Four features of the SPS make it a strong candidate as a foundation for the development of human-like broad AI.* In brief, they are:

- The potential of the SPS to solve 19 problems in AI research (Section 5);

- The versatility of the SPS in aspects of intelligence (Section 6.1), including several forms of reasoning (Section 6.2), and the versatility of the SPS in the representation and processing of knowledge (Section 6.3);

- Because the strengths of the SP System in aspects of intelligence, and in the representation and processing of knowledge, all flow from the SPMA framework, there is clear potential for their seamless integration, in any combination (Section 6.4). That kind of integration appears to be a necessary feature in any artificial system that aspires to human-like broad intelligence.

- Other potential benefits and applications of the SPS, several of them not closely related to AI (Section 7);

It is envisaged that the SPCM will provide a foundation for the development of an SP Machine (Section 8). Clones of the machine may be created by individuals or groups anywhere to facilitate further research and development of the SP System.

# Acknowledgements

# A     Outline of the SP Theory of Intelligence and its realisation in the SP Computer Model

The *SP System* (SPS)—meaning the *SP Theory of Intelligence* and its realisation in the *SP Computer Model* (SPCM)—is the product of a lengthy programme of research, from about 1987 to 2021 with a break between early 2006 and late 2012. This programme of research has included the creation and testing of many versions of the SPCM.

A major discovery has been the concept of *SP-multiple-alignment* (SPMA) and its versatility in many aspects of intelligence (Section 6 and Appendix A.3).

## A.1     High level view of the SPS

In broad terms, the SPS is a brain-like system that takes in *New* information through its senses and stores some or all of it as *Old* information that is compressed, as shown schematically in Figure 3.

In the SPS, all kinds of knowledge are represented with *SP-patterns*, where each such SP-pattern is an array of atomic *SP-symbols* in one or two dimensions. An SP-symbol is simply a 'mark' that can be matched with any other SP-symbol to determine whether it is the same or different.

At present, the SPCM works only with one-dimensional SP-patterns but it is envisaged that it will be generalised to work with two-dimensional SP-patterns as well. This should facilitate:

- The representation and processing of pictures and diagrams;

- For reasons explained in [19, Sections 6.1 and 6.2], the representation and processing of three-dimensional structures as well.
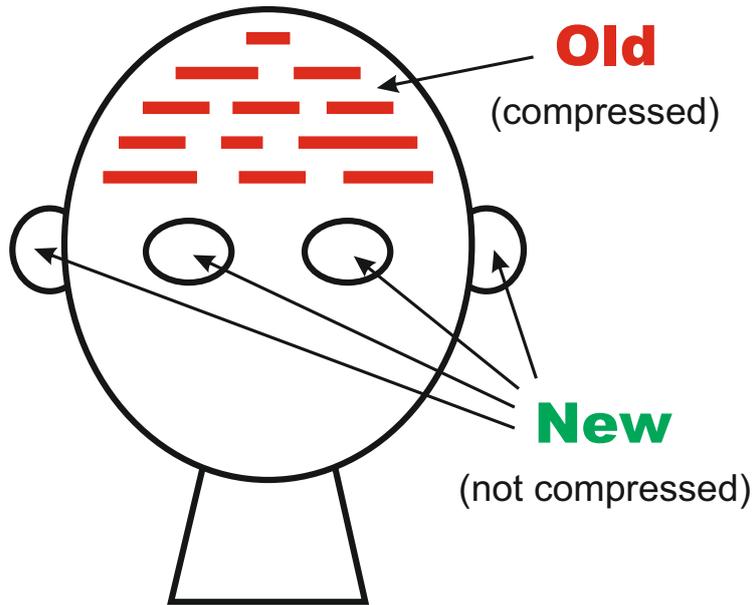
Figure 3: Schematic representation of the SPS from an 'input' perspective. Reproduced, with permission, from Figure 1 in [18].

- The representation of two or more procedures in parallel [20, Sections V-G, V-H, V-I, and Appendix C].

## A.2 Information compression in the SPS via the matching and unification of patterns

In accordance with Section B.2, a central idea in the SPS is that all kinds of processing would be achieved via IC. As noted earlier, evidence for the importance of IC in HLPC is described in papers by pioneers in this area including [1, 2, 3, 4], and in the more recent review in [28].

In the development of the SPS, it has proved useful to understand IC as a process of searching for patterns that match each other and the merging or 'unifying' patterns that are the same. The expression 'Information Compression via the Matching and Unification of Patterns' may be abbreviated as 'ICMUP'.

More specifically IC in the SPS is achieved largely via the creation of SPMAs (Appendix A.3), and the creation of SP-grammars via unsupervised learning, with SPMAs playing an important role (Appendix A.4).

In terms of the SP Theory of Intelligence, the emphasis on IC in the SPS accords with research in the tradition of Minimum Length Encoding (see, for example, [8]),

with the qualification that most research relating to MLE assumes that the concept of a universal Turing machine provides the foundation for theorising, whereas the SPS is itself a theory of computing founded on concepts of ICMUP and SPMA [24, Section II-C].

For the time being, the SPCM has been and now is being developed for lossless IC, but there may be a case in the future for exploring the possible benefits of lossy IC.

## A.3  SP-multiple-alignment

A central idea in the SPS, is the simple but powerful concept of SPMA, borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics.

SPMA is the last of seven variants of ICMUP described in [29, Section 5]. It may be seen to be a generalisation of the other six variants [29, Section 5.7].

Within the SPS, the SPMA concept is largely responsible for the intelligence-related strengths and potential of the SPS as outlined in Section 6.

The versatility of the SPS may also be seen in its several potential areas of application summarised in Section 7.

Bearing in mind that it is just as bad to underplay the strengths and potential of a system as it is to oversell its strengths and potential, it seems fair to say that *the concept of SP-multiple-alignment may prove to be as significant for an understanding of 'intelligence' as is DNA for biological sciences. It may prove to be the 'double helix' of intelligence.*

### A.3.1  Multiple sequence alignments

As mentioned above, the SPMA concept has been borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics. Figure 4 shows an example. Here, there are five DNA sequences which have been arranged alongside each other, and then, by judicious 'stretching' of one or more of the sequences in a computer, symbols that match each other across two or more sequences have been brought into line.

A 'good' multiple sequence alignment, like the one shown, is one with a relatively large number of matching symbols from row to row. The process of discovering a good multiple sequence alignment is normally too complex to be done by exhaustive search, so heuristic methods are needed, building multiple sequence alignments in stages and, at each stage, selecting the best partial structures for further processing.

Some people may argue that the combinational explosion with this kind of problem, and the corresponding computational complexity, is so large that there is no practical way of dealing with it. In answer to that objection, there are several

```
G G A       G     C A G G G A G G A     T G     G   G G A
| | |       |     | | | | | | | | |     | |     |   | | |
G G | G   G C C C A G G G A G G A     | G G C G   G G A
| | |     | | | | | | | | | | | | |     | |     |   | | |
A | G A C T G C C C A G G G | G G | G C T G     G A | G A
| | |             | | | | | | | | |   |   |     |   | | |
G G A A     | A G G G A G G A   | A G     G   G G A
| | |   |     | | | | | | | |     |   |     |   | | |
G G C A       C A G G G A G G     C   G     G   G G A
```

Figure 4: A 'good' multiple sequence alignment amongst five DNA sequences.

multiple sequence alignment programs used in bioinformatics—such as 'Clustal Omega', 'Kalign', and 'MAFFT'[3]—which produce results that are good enough for practical purposes.

This relative success is achieved via the use of heuristic methods that conduct the search for good structures in stages, discarding all but the best alignments at the end of each stage. With these kinds of methods, reasonably good results may be achieved but normally they cannot guarantee that the best possible result has been found.

### A.3.2  How SP-multiple-alignments are created

Figure 5 shows an example of an SPMA, superficially similar to the one in Figure 4, except that the sequences are called *SP-patterns*, the SP-pattern in row 0 is New information and each of the remaining SP-patterns, one per row, is an Old SP-pattern, selected from a relatively large pool of such Old SP-patterns. A 'good' SPMA is one which allows the New SP-pattern to be encoded economically in terms of the Old SP-patterns.

---

[3]Provided as online services by the European Bioinformatics Institute (see https://www.ebi.ac.uk/Tools/msa/).

```
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 D             #D N |                 #N #NP              | | | |           4
           |                    |                       |            | | | |
5          |                    |                       |      Vr 1 p l a y #Vr         5
           |                    |                       |      |              |
6          |                    |                       | V Vp Vr          #Vr #V       6
           |                    |                       | | |                |
7 S Num    ; NP                 |                    #NP V |                    #V #S 7
     |       |                  |                       |
8   Num PL ;                    Np                      Vp                              8
```

Figure 5: The best SP-multiple-alignment created by the SPCM with the sentence 't w o k i t t e n s p l a y' as the New SP-pattern in row 0, and a pool of Old SP-patterns representing grammatical structures, including those shown in rows 1 to 8, one SP-pattern per row.

In this example, the New SP-pattern (in row 0) is a sentence and each of the remaining SP-patterns represents a grammatical category, where 'grammatical categories' include words. The overall effect of the SPMA in this example is the parsing of a sentence ('f o r t u n e f a v o u r s t h e b r a v e') into its grammatical parts and sub-parts.

As with multiple sequence alignments, it is almost always necessary to use heuristic methods to achieve useful results without undue computational demands. The use of heuristic methods helps to ensure that computational complexities in the SPS are within reasonable bounds [16, Sections A.4, 3.10.6 and 9.3.1]. Each SP-multiple-alignment is built up progressively, starting with a process of finding good alignments between pairs of SP-patterns. At the end of each stage, SP-multiple-alignments that score well in terms of IC are retained and the rest are discarded. There is more detail in [18, Section 4] and [16, Sections 3.4 and 3.5].

In the SPCM, the size of the memory available for searching may be varied, which means in effect that the scope for backtracking can be varied. When the scope for backtracking is increased, the chance of the program getting stuck on a 'local peak' (or 'local minimum') in the search space is reduced.

Contrary to the impression that may be given by Figure 5, the SPMA concept is very versatile and is largely responsible for the strengths and potential of the SPS, as described in Section 6.

### A.3.3   The SPS is *quite different* from a deep neural network

The several levels in an SPMA may give the impression that the SPS in its structure and workings is simply a variant of the structure and workings of a DNN. This is entirely false.

In DNNs, the layers are provided at the beginning of processing and do not change except in the strengthening of links between neurons. By contrast, the SPS stores its knowledge in the form of SP-patterns, and those SP-patterns become the rows in each of a multitude of different SPMAs, each of which contains its own distinctive array of SP-patterns, normally unique but sometimes two or more sets are the same but with different alignments. Also, IC is of central importance in the SPS by contrast with most DNNs in which IC has little or no role [11].

Distinctive features and advantages of the SPS are described more fully in [24].

## A.4   Unsupervised learning in the SPS

In the SPS, learning is 'unsupervised', deriving structures from incoming sensory information without the need for any kind of 'teacher', or anything equivalent (*cf.* [7]).

Unsupervised learning in the SPS is quite different from 'Hebbian' learning via the gradual strengthening or weakening of neural connections (Section challenge-of-unsupervised-learning-section), variants of which are the mainstay of learning in DNNs. In the SPS, unsupervised learning incorporates the building of SPMAs but there are other processes as well.

### A.4.1   The creation of Old SP-patterns

In brief, the system creates Old SP-patterns from complete New SP-patterns and also from partial matches between New and Old SP-patterns. All learning in the SPS starts with the taking in of information from the environment:

- If that information is the same as one or more Old SP-patterns, then the frequency of the one SP-pattern is increased, or frequencies of the two or more SP-patterns are increased.

- If that information is entirely new, 'ID' SP-symbols[4] are added at the beginning and end of the pattern so that it becomes an SP-pattern. Then it is added directly to the store of Old SP-patterns.

- If partial matches can be made between the newly-received information and one or more of the stored Old SP-patterns, then each of the parts that match, and each of the parts that don't match, are made into SP-patterns by the addition of ID SP-symbols at the beginning and end, and the newly-created SP-patterns are added to the store of Old SP-patterns.

### A.4.2   The creation of SP-grammars

With a given body of New SP-patterns, the system processes them as just sketched, and then searches for one or two 'good' *SP-grammars*, where an SP-grammar is a collection of Old SP-patterns, and it is 'good' if it is effective in the economical encoding of the original set of New SP-patterns, where that economical encoding is achieved via SPMA.

As with the building of SPMAs, the process of creating good grammars is normally too complex to be done by exhaustive search so heuristic methods are needed. This means that the system builds SP-grammars incrementally and, at each stage, it discards all but the best SP-grammars.

As with the building of SPMAs, the use of heuristic methods helps to ensure that computational complexities in the SPS are within reasonable bounds [16, Sections A.4, 3.10.6 and 9.3.1].

---

[4]Appendix A.1

The SPCM has already demonstrated an ability to learn generative grammars from unsegmented samples of English-like artificial languages, including segmental structures, classes of structure, and abstract patterns, and to do this in an 'unsupervised' manner ([18, Section 5], [16, Chapter 9]).

But there are (at least) two shortcomings with unsupervised learning in the SPS, outlined in [18, Section 3.3].

## A.5 The probabilistic nature of the SPS

Owing to the intimate relation that is known to exist between IC and concepts of probability [12, 13], and owing to the fundamental role of IC in the workings of the SPS, the system is inherently probabilistic ([18, Section 4.4], [16, Section 3.7]).

That said, it appears to be possible to imitate the all-nothing-nature of conventional computing systems via the use of data where all the probabilities yielded by the system are at or close to 0 or 1.

Because of the probabilistic nature of the SPS, it lends itself to the modelling of HLPC because of the prevalence of uncertainties in that domain. Also, the SPS sits comfortably within AI because of the probabilistic nature of most systems in AI, at least in more recent work in that area.

An advantage of the SPS in those areas is that it is relatively straightforward to calculate absolute or conditional probabilities for results obtained in, for example, different kinds of reasoning (Section 6.2).

The very close connection that exists between IC and concepts of probability may suggest that there is nothing to choose between them. But [29, Section 8.2] argues that, in research on aspects of AI and HLPC, there are reasons to regard IC as more fundamental than probability and a better starting point for theorising.

## A.6 Two high-level mechanisms for IC in the SPS, and their functions

The two main mechanisms for IC in the SPS, both of which incorporate ICMUP (Appendix A.2), are as follows, each one with details of its function or functions:

1. *The building of SP-multiple-alignments.* The process of building SPMAs achieves compression of New information. At the same time it may achieve any or all of the following functions described in [16, Chapters 5 to 8] and [18, Sections 7 to 12], with potential for more:

    (a) The parsing of natural language (which is quite well developed); and understanding of natural language (which is only at a preliminary stage of development).

(b) Pattern recognition which is robust in the face of errors of omission, commission, or substitution; and pattern recognition at multiple levels of abstraction.

(c) Information retrieval which is robust in the face of errors of omission, commission, or substitution.

(d) Several kinds of probabilistic reasoning, as summarised in Section 6.2.

(e) Planning such as, for example, finding a flying route between London and Beijing.

(f) Problem solving such as solving the kinds of puzzle that are popular in IQ tests.

The building of SPMAs is also part of the process of unsupervised learning, next.

2. *Unsupervised learning.* Unsupervised learning, outlined in Appendix A.4, means the creation of one or two *SP-grammars* which are collections of SP-patterns which are effective in the economical encoding of a given set of New SP-patterns.

## A.7 SP-Neural

A potentially useful feature of the SPS is that it is possible to see how abstract constructs and processes in the system may be realised in terms of neurons and their interconnections. This is the basis for *SP-Neural*, a 'neural' version of the SPS, described in [23].

The concept of an SP-symbol may realised as a *neural symbol* comprising a single neuron or, more likely, a small cluster of neurons. An SP-pattern maps quite well on to the concept of a *pattern assembly* comprising a group of inter-connected SP-symbols. And an SPMA may be realised in terms of pattern assemblies and their interconnections, as illustrated in Figure 6.

In this connection, it is relevant to mention that the SPS, in both its abstract and neural forms, is quite different from DNNs [11] and has substantial advantages compared with such systems, as described in [31] and in [24, Section V].
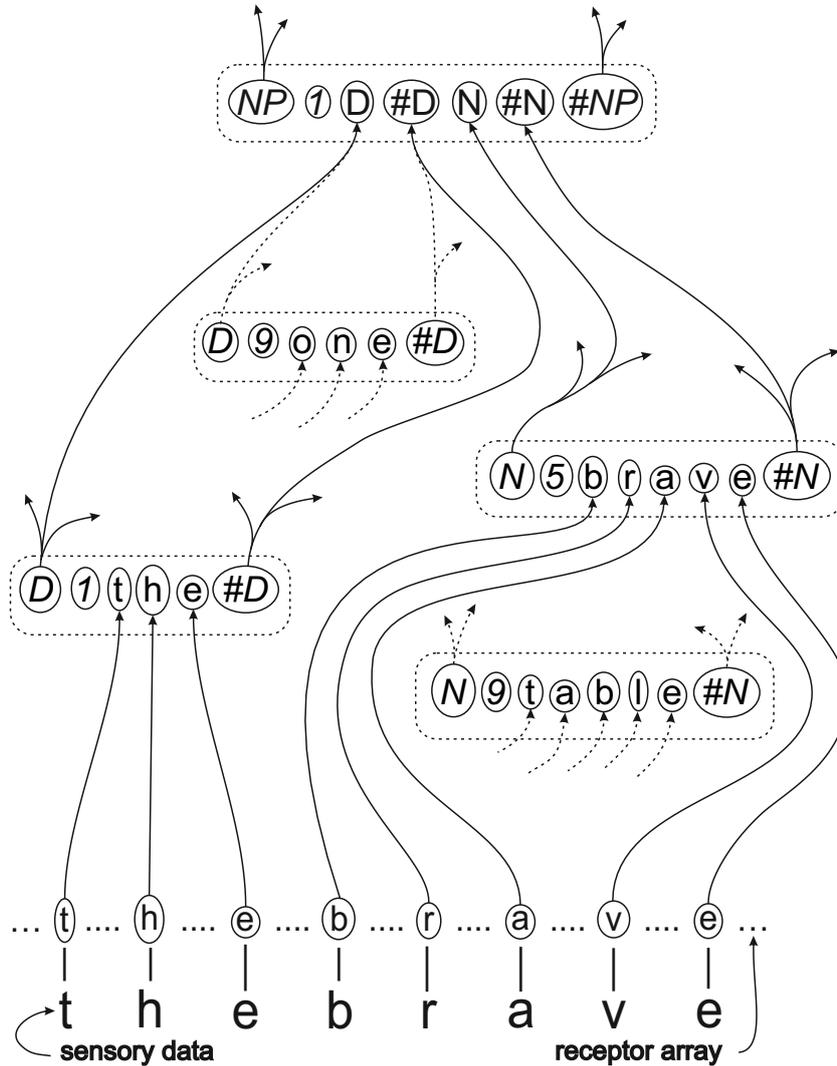
Figure 6: A schematic representation of a partial SPMA in SP-Neural, as discussed in [23, Section 4]. Each broken-line rectangle with rounded corners represents a *pattern assembly*—corresponding to an SP-pattern in the main SP Theory of Intelligence; each character or group of characters enclosed in a solid-line ellipse represents a *neural symbol* corresponding to an SP-symbol in the main SP Theory of Intelligence; the lines between pattern assemblies represent nerve fibres with arrows showing the direction in which impulses travel; neural symbols are mainly symbols from linguistics such as 'NP' meaning 'noun phrase, 'D' meaning a 'determiner', '#D' meaning the end of a determiner, '#NP' meaning the end of a noun phrase, and so on.

25

# B  Distinctive features of the SP research

This appendix picks out features of the SP research that are particularly noteworthy, memorable, or significant.

## B.1  Biochemistry provides an inspiration for new thinking about artificial intelligence

*Section 4.*  The concept of 'multiple sequence alignment', used by biochemists for the analysis of sequences of DNA bases and amino-acid residue, provides the inspiration for new thinking about artificial intelligence.

## B.2  The importance of information compression in HLPC and AI

*Section B.2.*  The SPS is unique amongst AIs in its strict adherence to the principle that IC should be central in AI, derived from the known importance of information compression in the workings of brains and nervous systems [28]. In the SP System:

- *All* kinds of processing is achieved via information compression;

- information compression is the key principle in the structuring of knowledge via unsupervised learning.

## B.3  Information compression may be understood as ICMUP

*Appendix A.2.*  In the SP research, information compression is understood as a search for patterns that match each other and the merging or 'unification' of patterns that are the same. The expression 'Information Compression via the Matching and Unification of Patterns' may be abbreviated as 'ICMUP'.

## B.4  The discovery of the SP-multiple-alignment concept

*Section 4.*  A major discovery in the SP research is the powerful concept of *SP-multiple-alignment*, borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics.

A bonus in this discovery is that the SP-multiple-alignment concept may be understood in terms of ICMUP.

## B.5   Versatility in the SP System

*Section 6.* Thanks largely to the concept of SP-multiple-alignment, the SP System demonstrates versatility across several aspects of intelligence, including unsupervised learning and several kinds of reasoning, and also versatility in the representation of knowledge.

## B.6   Seamless integration in the SP System

*Section 6.4.* Because the strengths of the SPS in aspects of intelligence and the representation of knowledge all flow from the SP-multiple-alignment concept, the SPS provides for the seamless integration of diverse aspects of intelligence and diverse forms of knowledge, in any combination. That kind of seamless integration appears to be a necessary feature of human-like broad AI.

## B.7   Development of the SP System, with associated work, took about 17 years

*Section 4.* Developing computer model for the SPS, exploring its range of AI-related applications, and writing a book about the research (*Unifying Computing and Cognition* [16]), took about 17 years.

## B.8   Hundreds of different versions of the SP Computer Model

*Section 3.* Most of the work of developing the SP System, especially the concept of SP-multiple-alignment, was developing *hundreds* of different versions of the SP Computer Model, each one developing a possible idea for how the model might work, applying relevant tests, taking notes on the results, and exploring ideas for what might be tried next.

## B.9   Mathematics as information compression

*Section 7.* A potentially useful spin-off from this research is the discovery that much of mathematics, perhaps all of it, may be understood as information compression via the matching and unification of patterns [29].

There is potential in this interpretation for the creation of a *New Mathematics* with many potential benefits and applications [29, Section 9].

## B.10 Simplification and integration of observations and concepts across a broad canvass

*Section 2.3.* The SP programme of research is uniquely ambitious in attempting to simplify and integrate observations and concepts across AI, mainstream computing, mathematics, and human learning, perception, and cognition.

Although lip service is often paid to the importance of Ockham's razor, it is rare for researchers to follow through on that principle by attempting to simplify and integrate observations and concepts across a broad canvass.

## B.11 The potential of the SP System to solve 19 problems in AI research

*Section 5.* There is clear potential for the SP System to solve 19 problems in AI research, all but two of them described by influential researchers in AI, in interviews with science writer Martin Ford [6].

## B.12 Intelligence-related strengths and potential of the SP System

*Section 6.* The SP Computer Model demonstrates versatility in aspects of intelligence, including several kinds of reasoning, and versatility in the representation of knowledge.

## B.13 Seamless integration of diverse aspects of intelligence and diverse kinds of knowledge, in any combination

*Section 6.4.* Because versatility of the SPS in aspects of intelligence, and its versatility in the representation of diverse kinds of knowledge all flow from the SPMA concept, there is likely to be seamless integration of diverse aspects of intelligence and diverse forms of knowledge, in any combination. This kind of seamless integration appears to be a necessary feature of any artificial system that aspires to human-like broad AI.

## B.14 Potential benefits and applications of the SP System beyond AI

*Section 7.* Apart from its AI-related strengths and potential, the SP System has potential benefits and applications in other areas of IT including: the management

of big data; transparency in the processing and representation of knowledge; computer vision and the understanding of natural vision; medical diagnosis; providing a radically new interpretation for concepts in mathematics; and more.

# C    Abbreviations

Abbreviations used in this paper are detailed below.

- *Artificial Intelligence*: 'AI'.

- *Deep Neural Network*: 'DNN'.

- *Human Learning, Perception, and Cognition*: 'HLPC'.

- *Information Compression*: 'IC'.

- *Information Compression via the Matching and Unification of Patterns*: 'ICMUP'.

- *SP Computer Model*: 'SPCM'.

- *SP-multiple-alignment*: 'SPMA'.

# References

[1] F. Attneave. Some informational aspects of visual perception. *Psychological Review*, 61:183–193, 1954.

[2] F. Attneave. *Applications of Information Theory to Psychology*. Holt, Rinehart and Winston, New York, 1959.

[3] H. B. Barlow. Sensory mechanisms, the reduction of redundancy, and intelligence. In HMSO, editor, *The Mechanisation of Thought Processes*, pages 535–559. Her Majesty's Stationery Office, London, 1959.

[4] H. B. Barlow. Trigger features, adaptation and economy of impulses. In K. N. Leibovic, editor, *Information Processes in the Nervous System*, pages 209–230. Springer, New York, 1969.

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

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

[7]  M. Gold. Language identification in the limit. *Information and Control,* 10:447–474, 1967.

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

[9]  G. Marcus. Deep learning: a critical appraisal. Technical report, New York University, 2018. arXiv:1801.00631 [cs.AI].

[10]  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.

[11]  J. Schmidhuber. Deep learning in neural networks: an overview. *Neural Networks,* 61:85–117, 2015.

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

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

[14]  J. G. Wolff. Learning syntax and meanings through optimization and distributional analysis. In Y. Levy, I. M. Schlesinger, and M. D. S. Braine, editors, *Categories and Processes in Language Acquisition,* pages 179–215. Lawrence Erlbaum, Hillsdale, NJ, 1988. bit.ly/ZIGjyc.

[15]  J. G. Wolff. Medical diagnosis as pattern recognition in a framework of information compression by multiple alignment, unification and search. *Decision Support Systems,* 42:608–625, 2006. arXiv:1409.8053 [cs.AI], bit.ly/1F36o7.

[16]  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. The print version is produced via print-on-demand from "INGRAM Lightning Source" and, via that technology, will never go out of print.

[17]  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.

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

[19] 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.

[20] 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.

[21] 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, Boca Raton, Florida, 2016, Chapter 6, pp. 143–170.

[22] J. G. Wolff. The SP Theory of Intelligence: benefits and applications. *Information*, 5(1):1–27, 2014. arXiv:1307.0845 [cs.AI], bit.ly/1FRYwew.

[23] 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.

[24] 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.

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

[26] J. G. Wolff. Interpreting Winograd Schemas via the SP Theory of Intelligence and its realisation in the SP Computer Model. Technical report, CognitionResearch.org, 2018. Submitted for publication. bit.ly/2ME8DOA.

[27] J. G. Wolff. Commonsense reasoning, commonsense knowledge, and the sp theory of intelligence. Technical report, CognitionResearch.org, 2019. In preparation. viXra:1901.0051v2, hal-01970147 version 3, bit.ly/2RESeut.

[28] 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.

[29] 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.

[30] J. G. Wolff. How the sp system may promote sustainability in energy consumption in it systems. *Sustainability*, 13(8), 2021. Article-number = 4565.

[31] J. G. Wolff. Problems in AI research and how the SP System may solve them. Technical report, CognitionResearch.org, 2021. Download: arXiv:2009.09079v3 [cs.CY], viXra:2009.0012, hal-02928784; tinyurl.com/5buj8swp. Submitted for publication.

[32] J. G. Wolff. Transparency and granularity in the SP Theory of Intelligence and its realisation in the SP Computer Model. In Witold Pedrycz and Shyi-Ming Chen, editors, *Interpretable Artificial Intelligence: A Perspective of Granular Computing*. Springer-Verlag, Heidelberg, 2021. Published in the book *Interpretable Artificial Intelligence: A Perspective of Granular Computing*, Witold Pedrycz and Shyi-Ming Chen (editors), Springer: Heidelberg, 2021, ISBN 978-3-030-64948-7, DOI: 10.1007/978-3-030-64949-4. Archive of chapter: arXiv:2009.06370 [cs.AI].