Multiple alignment and intelligence

J Gerard Wolff*

August 18, 2016

Abstract

The *SP theory of intelligence* aims to simplify and integrate observations and ideas across artificial intelligence, mainstream computing, mathematics, and human perception and cognition, with information compression as a unifying theme. Compression of information is achieved by searching for patterns that match each other and merging or unifying patterns that are the same. More specifically, information compression is achieved via the powerful concept of *multiple alignment*, borrowed and adapted from bioinformatics. The SP theory, which is realised in the form of a computer model, provides an interpretation for concepts and phenomena in several areas including unsupervised learning, 'computing', aspects of mathematics and logic, the representation of knowledge, natural language processing, pattern recognition, several kinds of reasoning, information storage and retrieval, planning and problem solving, information compression, neuroscience, and human perception and cognition.

1 Introduction

The SP theory of intelligence, and its realisation in the SP computer model, is a unique attempt, in accordance with Occam's Razor, to simplify and integrate observations and concepts across artificial intelligence, mainstream computing, mathematics, and human perception and cognition.[1] Developing an answer to this problem has not been easy and has taken several years.

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

[1]Much of the motivation for this project is the observation that AI, mainstream computer science, and research in human perception and cognition, have become fragmented into many subfields with little communication between them, as described

The most comprehensive description of the theory is in a book [12]. A fairly full but shorter overview of the theory is in [14]. Potential benefits and applications of the theory are described in [19] (more later).

Much of the inspiration for the SP theory has been a body of work, pioneered by Fred Attneave, Horace Barlow, and others, showing the importance of information compression in the workings of brains and nervous systems. Another source of inspiration has been my own work, developing computer models of language learning (summarised in [10]),[2] in which the importance of information compression became increasingly clear.

Of course, the relevance of information compression to AI and related topics is recognised in other research (see, for example, [7, 8]) but the approach which has been adopted in this programme of research is quite distinctive and has advantages compared with several AI-related alternatives [20]). Central in the theory is the powerful concept of *multiple alignment*, borrowed and adapted from bioinformatics. Potentially, this is as significant for computing and cognition as the double helix is for biological sciences.

1.1 An informal account of how the SP system works

The SP theory is conceived as an abstract brain-like system that, in an 'input' perspective, may receive *New* 'patterns' via its senses, and compress some or all of it to create *Old* 'patterns', as illustrated schematically in Figure 1.

In the early stages, when there is little or no Old information in store, the system simply stores New patterns directly as Old patterns, except that 'code' symbols are added to each pattern for use later.

After a while, New patterns are received that are fully or partially the same as Old patterns. Then, the system builds multiple alignments as outlined in Section 5.2. From the multiple alignments:

- When a New pattern is exactly the same as a stored Old pattern, the system records the occurrence of that pattern in terms of its code symbols, not the pattern itself. Since the code symbols are normally relatively short compared with their associated pattern, the effect is to store the New pattern in a compressed form.

- When a New pattern is received that is partially the same as one or more Old patterns, the system constructs patterns from the parts of

very well by Pamela McCorduck: "The goals once articulated with debonair intellectual verve by AI pioneers appeared unreachable ... Subfields broke off—vision, robotics, natural language processing, machine learning, decision theory—to pursue singular goals in solitary splendor, without reference to other kinds of intelligent behaviour." [5, p. 417].

[2] See also www.cognitionresearch.org/lang_learn.html.

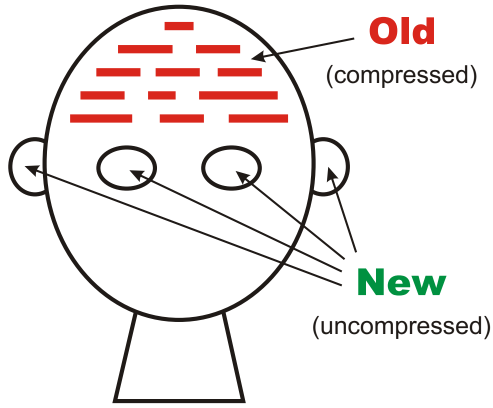

Figure 1: Schematic representation of the SP system from an 'input' perspective.

the New and Old patterns that match each other and from the parts that do not match each other. Each newly-created pattern is assigned its own code symbols and is stored as an Old pattern. The system also constructs 'abstract' patterns that, using relevant code symbols, record the sequential or two-dimensional relationship of the newly-constructed patterns.

- Similar principles apply when the system builds multiple alignments containing previously-constructed abstract patterns.

Many of the Old patterns that are created in this way may be seen as 'good', but normally the system also creates many that people are likely to regard as 'bad'. To get rid of the bad patterns, SP system applies a process for evaluating sets of Old patterns (called *grammars*) in terms of their ability to compress incoming New information. The one or two grammars that yield high levels of compression are retained by the system while residual patterns—those not found in the successful grammars—are discarded.

It is interesting to see that, very often, patterns that are 'good' in terms of information compression are also 'good' in terms of human judgements, and *vice versa*. This correspondence between what people regard as 'natural' and what yields high levels of compression is termed the 'DONSVIC' principle [14, Section 5.2].

The building of multiple alignments provides the means by which New information may be encoded economically in terms of Old information. It also provides much of the versatility of the system in such functions as unsupervised learning, the parsing and production of natural language, pattern recognition, computer vision, information retrieval, several kinds of reasoning, planning, problem solving, and information compression.

2 Foundations and scope of the SP theory

The SP theory is founded on a range of observations suggesting the fundamental importance of information compression via ICMUP in natural and artificial intelligence, in computing, in mathematics, and in neuroscience ([18], [12, Chapter 2]).

Like most theories, the SP theory is narrower in its scope than one might wish. It is certainly not a comprehensive theory of human psychology. For example, it has little to say about emotions and motivations and their impact, in people, on such things as perception, learning, and reasoning (but see [16, Section V-A.2]). At some stage, there is likely to be a case for examining whether or how those kinds of things may be accommodated in the theory.

3 The SP computer model and the SP machine

The SP theory is realised in the form of a computer model, SP71, which may be regarded as a version of the *SP machine*.

An outline of the organisation and workings of the SP computer model works may be found in [12, Section 3.9], with more detail, including pseudocode, in [12, Sections 3.10 and 9.2].[3] Fully commented source code for the SP71 computer model may be downloaded via a link near the bottom of www.cognitionresearch.org/sp.htm, and via "Ancillary files" under www.arxiv.org/abs/1306.3888.

All the multiple alignments shown in this paper are output from the SP computer model.

4 Patterns and symbols

In the SP system, knowledge is represented with arrays of atomic symbols in one or two dimensions called *patterns*. The SP71 model works with one-

[3]These sources describe SP70, a slightly earlier version of the model than SP71 but quite similar to it. The description of SP70 includes a description, in [12, Sections 3.9.1 and 3.10], of a subset of the SP70 model called SP61.

dimensional patterns but it is envisaged that the system will be generalised to work with patterns in two dimensions [14, Section 3.3].

Each SP pattern has an associated frequency of occurrence that has a role in the calculation of probabilities, as outlined in Section 6.

An 'atomic symbol' in the SP system is simply a mark that can be matched with any other symbol to determine whether it is the same or different: no other result is permitted.

Patterns in two dimensions are likely to have a role in the processing of images ([12, Chapter 13], [15]) and also in the processing of sensory or motor streams of information that occur in parallel [16, Sections IV-A.4, IV-H, V-G to V-I, and Section C].

In themselves, SP patterns are not particularly expressive. But within the multiple alignment framework (Section 5.2), they support the representation and processing of a wide variety of kinds of knowledge (Section 8). It appears that the system has potential as a *universal framework for the representation and processing of knowledge* (UFK) [17, Section III].

5 Information compression

In the SP system, all kinds of processing is done by compression of information. This is essentially the principle of *minimum length encoding* (MLE) [7, 9, 6][4] but with qualifications described in [20, Section 3].

The default assumption in the SP theory is that compression of information is always lossless, meaning that all non-redundant information is retained. In particular applications, there may be a case for discarding non-redundant information (see, for example, [17, Section X-B]) but any such discard is reversible.

The name "SP" is short for *Simplicity* and *Power*, because compression of any given body of information, **I**, may be seen as a process of reducing "redundancy" of information in **I** and thus increasing its "simplicity", whilst retaining as much as possible of its non-redundant descriptive and explanatory "power".

Information compression is achieved via the matching and unification of patterns, or parts thereof (Section 5.1, [18]). More specifically, it is achieved via the building of multiple alignments and via the unsupervised learning of grammars. These three things are described briefly in the following three subsections.

[4] MLE is an umbrella term for "minimum message length" encoding (MML), "minimum description length" encoding (MDL), and similar concepts.

5.1 Information compression via the matching and unification of patterns

The basis for information compression in the SP system is a process of searching for patterns that match each other with a process of merging or 'unifying' patterns that are the same: "information compression via the matching and unification of patterns" or "ICMUP" [18].

At the heart of the SP computer model is a method for finding good full and partial matches between sequences, with advantages compared with classical methods [12, Section A].[5]

The emphasis on ICMUP is motivated partly by evidence of the importance of such processes in human perception and cognition, and partly by its potential to cut through much complexity and to achieve a new perspective on artificial intelligence, mainstream computing, and mathematics [18].

Because a goal of the SP theory is to develop a new perspective on AI, computing, and mathematics, without theoretical 'baggage', the theory minimises the use of mathematics [18, Section 2.1].[6]

5.2 Information compression via the building of multiple alignments

The process for finding good full and partial matches between patterns is the foundation for processes that build *multiple alignments* like the one shown in Figure 2. This concept is similar to multiple alignment in bioinformatics but with important differences [12, Section 3.4]. It is a powerful and distinctive feature of the SP system.

In Figure 2, the SP pattern in column 0 is a sentence to be parsed, while each of columns 1 to 12 contains an SP pattern representing a grammatical form (where 'grammatical form' includes words). This example shows the best multiple alignment created by the SP computer model when the New pattern is processed in conjunction with a set of pre-existing Old patterns like those shown in columns 1 to 12.

[5] The main advantages are [12, Section 3.10.3.1]: 1) That it can match arbitrarily long sequences without excessive demands on memory; 2) For any two sequences, it can find a set of alternative matches (each with a measure of how good it is) instead of a single 'best' match; 3) The 'depth' or thoroughness of the searching, which has the effect of controlling the amount of backtracking, can be controlled by parameters.

[6] And, bearing in mind that the SP theory should be consistent with the biological origins of human intelligence, an attempt has been made to ensure that the frequency information that is stored with each SP pattern, and the probability calculations that are performed by the SP computer model, are, potentially, the kinds of things that could be modelled, at least approximately, via analogue processes in biological systems.

```
0        1        2     3    4    5     6     7       8     9      10     11

                                        S
                                        SNG
                  NP ------------------ NP
                  Np1
                  D --- D
                        d2
a --------------------- a
                  #D -- #D
         N ------ N
         Ns --------------------------- Ns
         n6
stitch - stitch
         #N ----- #N
                  #NP ----------------- #NP
                                  Q --- Q
                                  q1
                             P -- P
                             p4
in ------------------------- in
                             #P - #P
                                  NP ------------------------------------ NP
                                                                          Np2
                                                                   N ---- N
                                                                   Ns
                                                                   n7
time ------------------------------------------------------------- time
                                                                   #N --- #N
                                  #NP ----------------------------------- #NP
                                  #Q -- #Q
                                        VP ---------- VP
                                                      Vp1
                                              V ----- V
                                        Vs -- Vs
                                              v2
saves --------------------------------------- saves
                                              #V ---- #V
                                                      N --- N
                                                            Ns
                                                            n8
nine ------------------------------------------------------ nine
                                                      #N -- #N
                                        #VP --------- #VP
                                        #S

0        1        2     3    4    5     6     7       8     9      10     11
```

Figure 2: A multiple alignment illustrating parsing of a sentence, as described in the text.

Here, the 'best' multiple alignment is the one in which the New pattern may be encoded most economically in terms of the Old patterns—and this means a multiple alignment in which there is a relatively large number of symbols that match each other from columns to column, aligned in rows. The way in which an encoding is derived from a multiple alignment is explained in [12, Section 3.5] and [14, Section 4.1]. Like all other kinds of knowledge in the SP system, encodings derived from multiple alignments are recorded using SP patterns (Section 4).

The overall effect of this multiple alignment is to analyse the sentence into its grammatical parts and sub-parts, an analysis that is, in its essentials, the same as a conventional parsing.

A point of interest is that the pattern for the whole sentence in column 6 marks the grammatical dependency between the singular subject of the sentence ('`a stitch`')—marked with '`Ns`'—and the singular verb ('`saves`')—marked with '`Vs`'. Notice how the dependency neatly bridges the subordinate phrase ('`in time`'). This method of encoding discontinuous dependencies in syntax contributes to the compression that is achieved by this multiple alignment and is, arguably, simpler than existing techniques for encoding such dependencies.

5.3 Information compression via unsupervised learning

As outlined in [12, Section 3.9.2] and [14, Section 5.1], and described more fully in [12, Chapter 9], the SP system may, without assistance from a 'teacher' or anything equivalent, derive one or more grammars from a collection of New patterns, with information compression as the guide. The SP computer model demonstrates unsupervised learning of plausible generative grammars for the syntax of English-like artificial languages, including the learning of segmental structures, classes of structure, and abstract patterns. In that process, multiple alignment has a central role as a source of SP patterns for possible inclusion in any grammar [12, Section 9,2,5], [14, Section 5.1.1].

Although the current model has some shortcomings (Section 10, [14, Section 3.3]), it appears that these may be overcome.

5.4 Heuristic search

Like most problems in artificial intelligence, each of the afore-mentioned problems—finding good full and partial matches between patterns, finding or constructing good multiple alignments, and inferring one or more good grammars from a body of data—is normally too complex to be solved by exhaustive search.

With intractable problems like these, it is often assumed that the goal is to find theoretically ideal solutions. But with these and most other AI problems, "The best is the enemy of the good". By scaling back one's ambitions and searching for "reasonably good" solutions, it is often possible to find solutions that are useful, and without undue computational demands.

As with other AI applications, and as with the building of multiple alignments in bioinformatics, the SP71 model uses heuristic techniques—'hill climbing' or 'descent'—in all three cases mentioned above [12, Section A; Sections 3.9 and 3.10; Chapter 9]. This means searching for solutions in stages, with a pruning of the search tree at every stage, guided by measures of compression, and with backtracking where appropriate to increase the chances of success. With these kinds of techniques, acceptably good approximate solutions can normally be found without excessive computational demands and with "big O" values that are within acceptable limits.

5.5 Grammars and encodings, simplicity and power

In unsupervised learning in the SP system, compression of a body of information, **I**, produces two distinct results: a *grammar* and an *encoding* of **I** in terms of the grammar, both of them expressed as SP patterns. The two together represent a lossless compression of **I**.

The term 'grammar' has been adopted because the SP programme of research derives largely from earlier research on models of language learning and grammatical inference [10] but, because of the versatility of SP patterns in the multiple alignment framework (Section 4), the term is applied, in this research, to any kind of knowledge, not just natural language.

Often but not invariably, there is a trade-off between the size of the grammar and the size of the encoding: as a general rule, small grammars yield large encodings and large grammars yield small encodings. Normally, the greatest overall compression of **I** is obtained with grammars that are not at the extremes of size (small or big), and likewise for encodings. It appears that this means learning that avoids both under-generalisation and over-generalisation [20, Sections 4.8 and 10.1].

From the trade-off we can see that there is a direct relationship between the concepts of 'grammar' and 'encoding' on the one hand, and the concepts of 'simplicity' and 'power' on the other: for a given **I**, there is simplicity in any grammar when the grammar is small, and the grammar has power when the encoding is small. Any reasonably thorough compression of **I** is likely to yield a good balance between the two.[7]

6 Information compression, prediction, and probabilities

Owing to the close connection between information compression and concepts of prediction and probability [4], the SP system is fundamentally probabilistic. As noted in Section 4, each SP pattern has an associated frequency of occurrence. Probabilities may be calculated for each multiple alignment and for any inference that may be drawn from any given multiple alignment [12, Section 3.7].

Although the SP system is fundamentally probabilistic: it can be constrained to answer only those kinds of questions where probabilities are close to 0 or 1; and, via the use of error-reducing redundancy, it can deliver decisions with high levels of confidence. Contrary to what one may suppose, there is no conflict between the use of error-reducing redundancy and the notion that "computing" may be understood as information compression—the two things are independent, as described in [12, Section 2.3.7].

7 SP-neural

Part of the SP theory is the idea, described most fully in [12, Chapter 11], that the abstract concepts of *symbol* and *pattern* in the SP theory may be realised more concretely in the brain with collections of neurons in the cerebral cortex.[8]

The neural equivalent of an SP pattern is called a *pattern assembly*. The word "assembly" has been adopted in this term because the concept is quite similar to Hebb's [2] concept of a *cell assembly*. The main difference is that the concept of pattern assembly is unambiguously explicit in proposing that the sharing of structure between two or more pattern assemblies is achieved by means of 'references' from one structure to another, as described and discussed in [12, Section 11.4.1]). Also, learning in the SP system is quite different from the kind of "Hebbian" learning that is popular in artificial neural networks [20, Sections 4.1 and 4.4].

[7] Here, the qualification, 'reasonably thorough' is quite important. Compression algorithms like the popular LZ algorithms are 'quick and dirty'—they are designed for speed on low-powered computers—they are not very thorough and will normally miss quite large amounts of redundancy.

[8] See also [12, Section 2.3.1].

Figure 3 shows schematically how pattern assemblies may be represented and inter-connected with neurons. Here, each pattern assembly, such as '`< NP < D > < N > >`', is represented by the sequence of atomic symbols of the corresponding SP pattern. Each atomic symbol, such as '`<`' or '`NP`', would be represented in the pattern assembly by one neuron or a small group of inter-connected neurons. Apart from the inter-connections amongst pattern assemblies, the cortex in SP-neural is somewhat like a sheet of paper on which knowledge may be written in the form of neurons.

The hierarchical relations that can be seen in Figure 3 may be seen to be broadly in keeping with the hierarchical relations between 'simple' and 'complex' cells discovered by Hubel and Wiesel [3].

It is envisaged that any pattern assembly may be 'recognised' if it receives more excitatory inputs than rival pattern assemblies, perhaps via a winner-takes-all mechanism [12, Section 11.3.4]. And, once recognised, any pattern assembly may itself be a source of excitatory signals leading to the recognition of higher-level pattern assemblies.

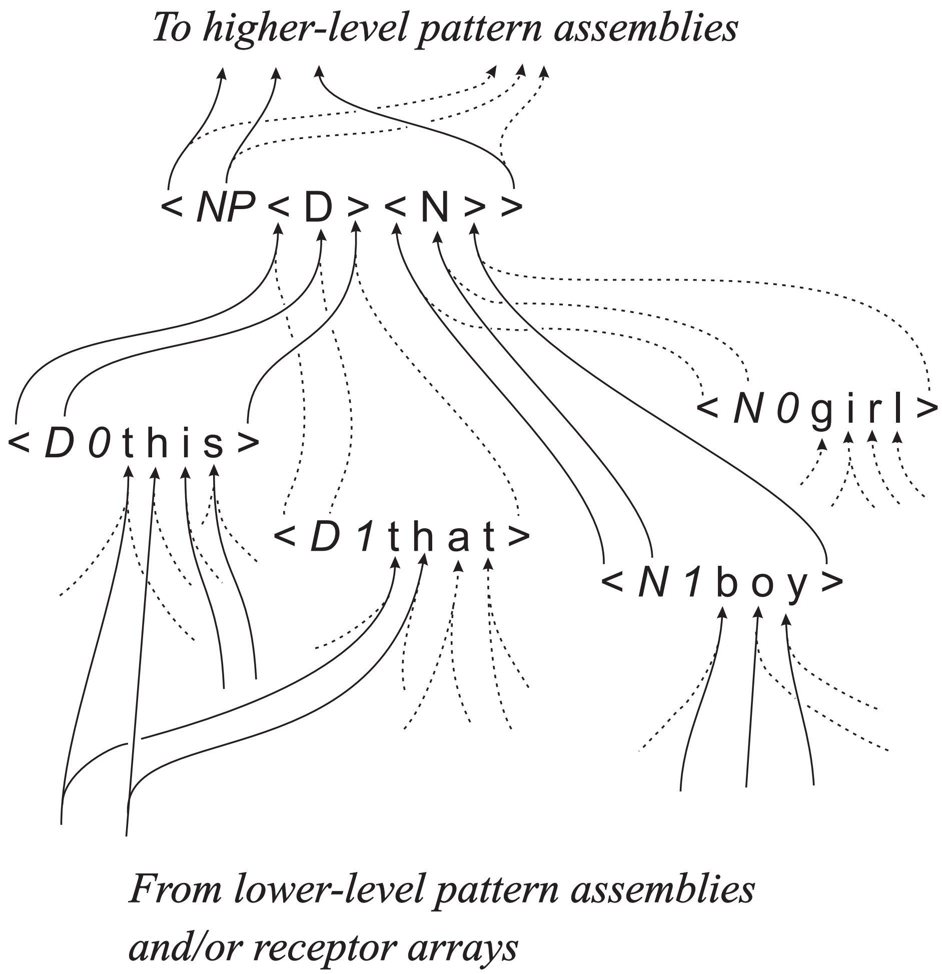

Figure 3: Schematic representation of inter-connections amongst pattern assemblies as described in the text. Not shown in the figure are lateral connections within each pattern assembly, and inhibitory connections elsewhere, as outlined in [12, Sections 11.3.3 and 11.3.4]. Reproduced, with permission, from Figure 11.2 in [12].

8 Empirical and conceptual support for the SP theory

The SP theory has non-trivial things to say about a wide range of observations and concepts in artificial intelligence, mainstream computing, mathematics, and human perception and cognition. These things are described most fully in [12], more briefly in [14], and in extended summaries in [16, Sections IV and V]. In a bare-bones summary, the main strengths of the SP system are in:

- Natural language processing ([12, Chapter 5], [14, Section 8]).
- Pattern recognition and vision ([12, Chapter 6], [14, Section 9], [15]).
- Information storage and retrieval ([12, Chapter 6], [14, Section 11], [13]).
- The representation and processing of diverse kinds of knowledge ([14, Section 7], [17, Section III-B] and, more generally, [12, Chapters 5 to 10]).
- Benefits accruing from the seamless integration of diverse kinds of knowledge and diverse aspects of intelligence ([19, Sections 2, 5, and 7]).
- Several kinds of reasoning ([12, Chapter 7], [14, Section 10]).
- Planning and problem solving ([12, Chapter 8], [14, Section 12]).
- Unsupervised learning ([12, Chapter 9], [14, Section 5], [16, Section V]).
- Implications for our understanding of human perception and cognition, including neural processing ([12, Chapters 11 and 12], [15]).
- Implications for our understanding of the nature of mathematics ([12, Chapter 10], [18]).

There is more detail in [12, 14] and other publications referenced throughout this paper.

9 Potential benefits and applications

In summary, potential benefits and applications of the SP system include:

- Helping to solve nine problems associated with big data [17].
- The development of computer vision and pattern recognition, and the interpreation of aspects of natural vision ([15], [14, Section 9]).
- The development of versatility and adaptability in autonomous robots, with potential for gains in computational efficiency [16].
- The system may be developed as a versatile database management system, with intelligence [13].
- The system may serve as a repository for medical knowledge and as an aid for medical diagnosis [11].
- There are several other potential benefits and applications described in [19]: simplification of computing systems, including software; unsupervised learning; the processing of natural language; software engineering; information compression; the semantic web; bioinformatics; the detection of computer viruses; data fusion; new kinds of computer; the development of scientific theories; and the seamless integration of diverse kinds of knowledge and processing.

As describe in Section 10, next, some potential applications may be developed on relatively short timescales.

10 Development of the SP system

Like most scientific theories, the SP system is not complete [14, Section 3.3]. As it is now, the main shortcomings in the SP computer model are:

- The process for finding good full and partial matches between one-dimensional patterns needs to be generalised to patterns in two dimensions;
- A better understanding is needed of how the system may be applied to the discovery and recognition of low-level features in speech and images;
- In unsupervised learning, the model does not learn intermediate levels of abstraction or discontinuous dependencies in data;

- And a better understanding is needed of how the system may be applied in the representation and processing of numbers.

It appears that none of these problems are showstoppers—that all of them are soluble.

Apart from this work, developing the core model, there is much to be done in exploring how the system may be applied in areas such as pattern recognition, natural language processing, and so on (Section 8), and in applications such as medical diagnosis, crime prevention and detection, and more (Section 9).

Since there are many more avenues to be explored than could be tackled by any one research group, it is envisaged that the SP computer model will be the basis for the creation of a high-parallel, open-source version of the SP machine, hosted on an existing high-performance computer [21], as shown schematically in Figure 4. This will be a means for researchers everywhere to explore what can be done with the system, and to create new versions of it.

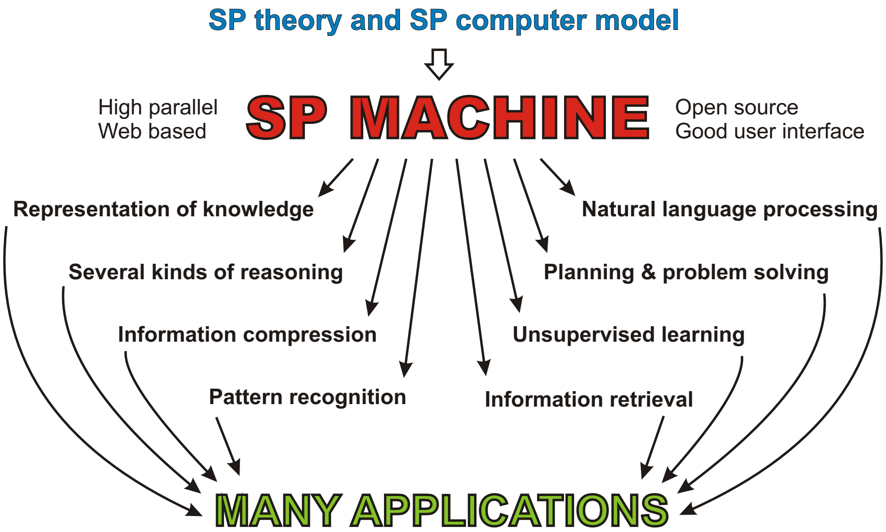

Figure 4: Schematic representation of the development and application of the SP machine. Reproduced from Figure 2 in [14], with permission.

Some potential applications of the SP system may be developed on relatively short timescales using existing high-performance computers or even ordinary computers. These include the SP system as an intelligent database [13], and applications in such areas as medical diagnosis [11], pattern recognition ([12, Chapter 6], [14, Section 9]), information compression [19, Section

6.7], highly-economical transmission of information [17, Section VIII], bioinformatics [19, Section 6.10.2], and natural language processing [19, Section 6.2].

11 Conclusion

As mentioned in the introduction, the concept of *multiple alignment* as it has been developed in this programme of research, may prove to be as significant for computing and cognition as the double helix is for biological sciences. Further development of these ideas is likely to pay dividends.

References

[1] G. C. Deka, editor. *NoSQL: database for storage and retrieval of data in cloud*. Taylor & Francis, Abingdon, 2016.

[2] D. O. Hebb. *The Organization of Behaviour*. John Wiley & Sons, New York, 1949.

[3] D. H. Hubel. Exploration of the primary visual cortex, 1955–78. In M. S. Gazzaniga, editor, *Cognitive Neuroscience: a Reader*, pages 58–80. Blackwell, Oxford, 2000.

[4] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, New York, 3rd edition, 2014.

[5] P. McCorduck. *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*. A. K. Peters Ltd, Natick, MA, second edition, 2004. ISBN: 1-56881-205-1.

[6] J. Rissanen. Modelling by the shortest data description. *Automatica*, 14(5):465–471, 1978.

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

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

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

[10] 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.

[11] 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. bit.ly/XE7pRG, arXiv:1409.8053 [cs.AI].

[12] 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.

[13] J. G. Wolff. Towards an intelligent database system founded on the SP theory of computing and cognition. *Data & Knowledge Engineering*, 60:596–624, 2007. bit.ly/Yg2onp, arXiv:cs/0311031 [cs.DB]. A revised version of this paper is to be published in [1] as "Towards a NoSQL database system with intelligence founded on the SP theory of intelligence".

[14] J. G. Wolff. The SP theory of intelligence: an overview. *Information*, 4(3):283–341, 2013. bit.ly/1hz0lFE.

[15] 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. bit.ly/1scmpV9.

[16] J. G. Wolff. Autonomous robots and the SP theory of intelligence. *IEEE Access*, 2:1629–1651, 2014. bit.ly/1zrSemu.

[17] J. G. Wolff. Big data and the SP theory of intelligence. *IEEE Access*, 2:301–315, 2014. bit.ly/1jGWXDH. This article, with minor revisions, is to be reproduced in Fei Hu (Ed.), *Big Data: Storage, Sharing, and Security (3S)*, Taylor & Francis LLC, CRC Press, 2016, pp. 143–170.

[18] J. G. Wolff. Information compression, intelligence, computing, and mathematics. Technical report, CognitionResearch.org, 2014. bit.ly/1jEoECH, arXiv:1310.8599 [cs.AI]. Submitted for publication.

[19] J. G. Wolff. The SP theory of intelligence: benefits and applications. *Information*, 5(1):1–27, 2014. bit.ly/1lcquWF.

[20] J. G. Wolff. The SP theory of intelligence: its distinctive features and advantages. *IEEE Access*, 4:216–246, 2016. bit.ly/21gv2jT.

[21] J. G. Wolff and V. Palade. Proposal for the creation of a research facility for the development of the SP machine. Technical report, CognitionResearch.org, 2015. bit.ly/1zZjjIs, arXiv:1508.04570 [cs.AI].