# Information compression as a unifying principle in human learning, perception, and cognition, and as a foundation for the SP Theory of Intelligence

*In an extended programme of research, Dr Gerry Wolff of CognitionResearch.org, with colleagues, has been developing the* SP Theory of Intelligence *and its realisation in the* SP Computer Model *– and exploring their potential applications. A unifying principle in this research is that much of human learning, perception, and cognition, may be understood as compression of information. A key development in this research is the powerful concept of* SP-multiple-alignment, *borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics.*

In an extended programme of research, Dr Gerry Wolff, with colleagues, is developing the *SP System*, meaning the *SP Theory of Intelligence* and its realisation in the *SP Computer Model*. An important part of this research is exploring the potential applications of the SP System.

The overarching goal is the simplification and integration of observations and concepts across artificial intelligence, mainstream computing, mathematics, and human learning, perception, and cognition.

A key idea in this research is that much of human learning, perception, and cognition may be understood as compression of information (more below).

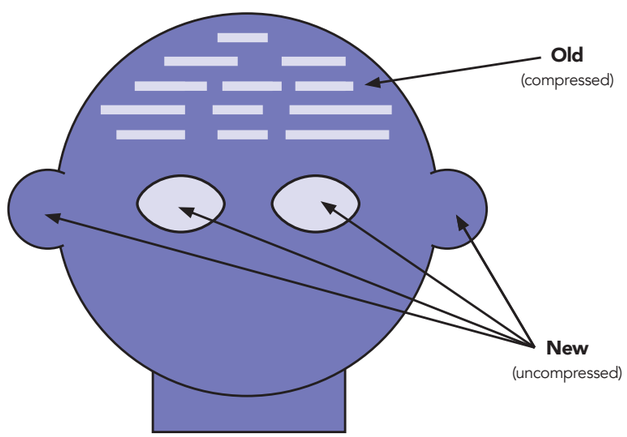



**Figure 1:** Schematic representation of the SP System

The powerful concept of *SP-multiple-alignment*, borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics is an important development in this research.

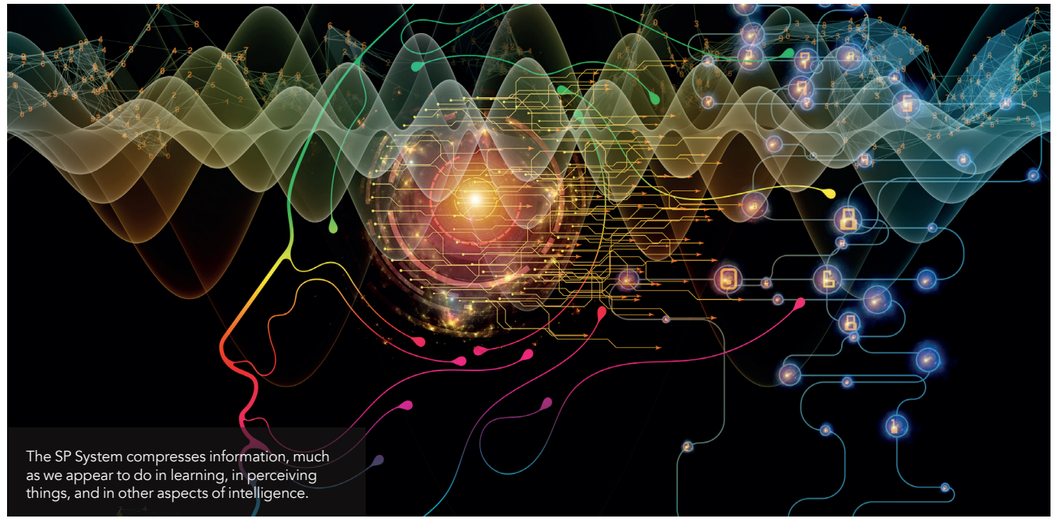
Peer-reviewed papers and other documents about the research are detailed, with download links, on www.cognitionresearch.org/sp.htm.

## THE SP SYSTEM

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

There is abundant evidence, described in the first of the 2019 papers under 'Further reading', below, that much of human learning, perception, and cognition, may be understood as compression of information. And in many cases, that compression of information may be understood as a process of searching for patterns that match each other and then merging or 'unifying' two or more such patterns to make one. The expression "information compression via the matching and unification of patterns" may be shortened to 'ICMUP'.

If, when we are looking at something, we close our eyes for a moment and then open them again, we normally see 'the same' as what we saw before, as shown schematically in *Figure 2*. The merging of the 'before' and 'after' views may be seen as ICMUP.



The SP System compresses information, much as we appear to do in learning, in perceiving things, and in other aspects of intelligence.

A well-known technique for compressing information, called 'chunking-with-codes', means using a relatively short 'code' to stand for and replace each of two or more instances of a relatively large 'chunk' of information. So a relatively short name like 'New York' can stand for the relatively huge 'chunk' of information which is all knowledge of the renowned city itself.[1]

A little reflection shows that every noun, verb, adjective, or adverb in a natural language may be seen to be a short code for a relatively large chunk of information. Every natural language may be seen to be a means of achieving very high levels of information compression.

## SP-MULTIPLE-ALIGNMENT

The previously mentioned concept of SP-multiple-alignment may be seen to provide seamless integration of six main variants of ICMUP, and is thus a powerful means of compressing information.

A large part of the workings of the SP System is a process of creating and selecting 'good' SP-multiple-alignments like the one shown in *Figure 3*, which shows how the system may analyse or 'parse' a sentence (in row 0) into words and other grammatical structures.[2]

Because of the complexity of the search space, heuristic techniques are needed, building SP-multiple-alignments in stages and discarding all but the best partly-developed SP-multiple-alignments at the end of each stage.

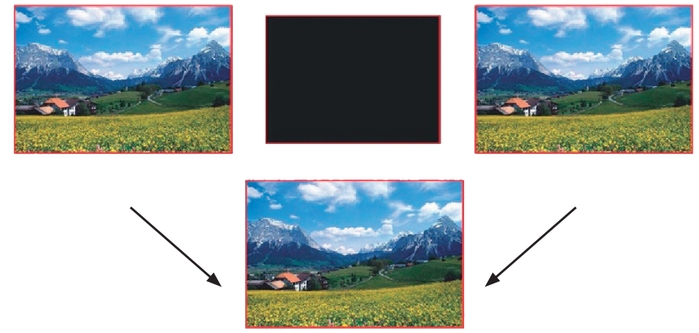


**Figure 2:** A schematic view of how, if we close our eyes for a moment and open them again, we normally merge the 'before' and 'after' views to make one.

***Much of human learning, perception, and cognition may be understood as compression of information.***

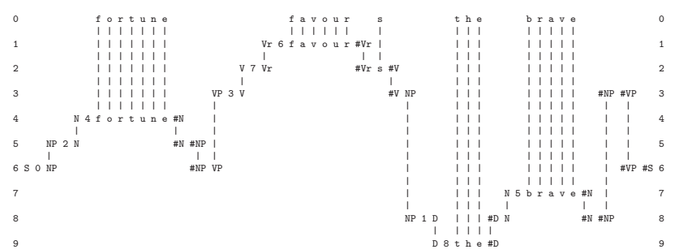



**Figure 3:** An SP-multiple-alignment that represents a parsing of the sentence "Fortune favours the brave".

An SP-multiple-alignment is 'good' if it can yield a relatively large amount of information compression by the merging or 'unification' of patterns that match each other within the SP-multiple-alignment.

SP-multiple-alignments can do much more than parse sentences. They are largely responsible for the versatility of the SP System as summarised in the next section.

The SP-multiple-alignment concept has the potential to be as significant for the understanding of intelligence as is DNA in biological sciences. It could prove to be the 'double helix' of intelligence.

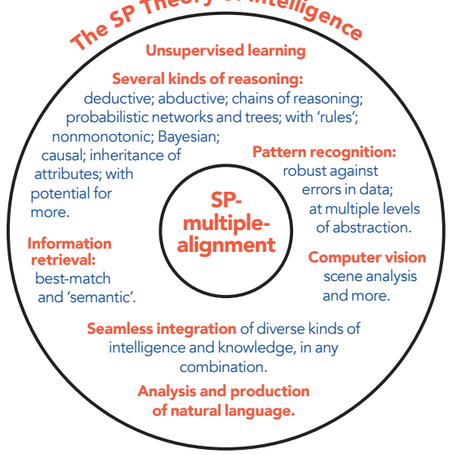



**Figure 4:** Schematic representation of the versatility of the SP System, with SP-multiple-alignment at centre stage.

## STRENGTHS AND POTENTIAL OF THE SP SYSTEM

Largely because of the versatility of the SP-multiple-alignment construct, the SP System has strengths and potential: in several aspects of intelligence, including several kinds of reasoning; in the representation of diverse kinds of knowledge; and in the seamless integration of diverse aspects of intelligence and diverse kinds of knowledge, in any combination. Some of the strengths and potential of the SP System are shown schematically in *Figure 4*.

***Information compression is so much embedded in our thinking and seems so natural and obvious that it is easily overlooked.***

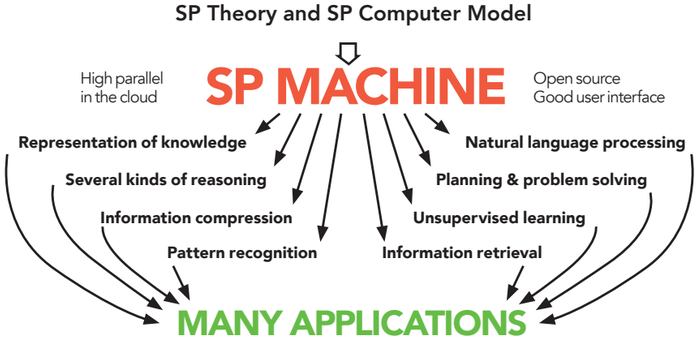



**Figure 5:** Schematic representation of the development and application of the SP Machine.

## POTENTIAL APPLICATIONS OF THE SP SYSTEM

The SP System has potential in several areas of application including: helping to solve problems with big data; helping to develop intelligence in autonomous robots; in the processing of natural language; in the development of an intelligent database system; in medical diagnosis; and more. Peer-reviewed papers and other documents about potential applications of the SP System are detailed, with download links, on www.cognitionresearch.org/sp.htm.

## SP-NEURAL

Abstract concepts in the SP System map quite well into *SP-Neural*, a version of the SP System expressed in terms of neurons and their interconnections, described in the 2016 paper detailed opposite. SP-Neural is quite different from the currently-popular 'deep neural networks' (DNNs).

## THE SP MACHINE

Dr William J Teahan of Bangor University, Dr Vasile Palade of Coventry University, and Dr Gerry Wolff are aiming to develop the SP Computer Model into an industrial-strength *SP Machine*, starting with the application of high levels of parallel processing. A paper by Dr Palade and Dr Wolff that describes a roadmap for the development of the SP Machine is detailed under "Further reading", opposite.

**Acknowledgements**
All the figures are reproduced with permission from Wolff (2019) (in "Further reading"). The landscape in Figure 2 is from Wallpapers Buzz (www.wallpapersbuzz.com), reproduced with permission.

**Endnotes**

[1] Chunking-with-codes may be seen as an example of ICMUP because, in effect, multiple instances of the relatively large chunk of information have been reduced to one.

[2] At present, the SP Computer Model works only with one-dimensional patterns, but it is envisaged that it will be generalised to work with 2D patterns. With relatively large basic symbols such as letters or words, the SP System has the flavour of 'symbolic AI'. With relatively small basic symbols such as pixels in an image, the SP System may be more like 'non-symbolic AI'.



# Behind the Research

## Dr Gerry Wolff

**E:** jgw@cognitionresearch.org **T:** +44 1248 712962
**W:** A download link for the source code of the latest version of the SP Computer Model (SP71), may be found under the heading 'SOURCE CODE' near the bottom of the web page bit.ly/2Gxici2.

## Research Objectives

Dr Wolff's work focuses on simplification and integration of observations and concepts across artificial intelligence, mainstream computing, mathematics, and human learning, perception, and cognition.

## Detail

Dr Gerry Wolff
CognitionResearch.org
Menai Bridge
UK

**Bio**
University of Cambridge (BA), University of Wales Cardiff (PhD), University of Dundee (computer models of language learning), Research Fellowship with IBM, Praxis Systems plc (CEng), School of Computer Science and Electronic Engineering at Bangor University (developing the SP System), CognitionResearch.org (continuing the development of the SP System).

**Funding**
The School of Computer Science and Electronic Engineering at Bangor University kindly provided an office and research facilities for five years, in support of Dr Wolff's research.

**Collaborators**
Dr William J Teahan, Lecturer in Computer Science in the School of Computer Science and Electronic Engineering in Bangor University, and Dr Vasile Palade, Reader in Pervasive Computing in the School of Computing, Electronics and Mathematics in Coventry University, are collaborating with Dr Wolff in the first stages of a project to develop the *SP Machine*, based on the SP Computer Model with the application of high levels of parallel processing. Dr Palade and Dr Wolff are the authors of a peer-reviewed paper, "A roadmap for the development of the 'SP Machine' for artificial intelligence", with details and a download link under "Further reading".

## Further Reading

Wolff, J. G. (submitted for publication). Unsolved problems in AI, described in the book 'Architects of Intelligence' by Martin Ford, and how they may be solved via the SP System, *PDF*, bit.ly/2th7Bze. Most of the problems described in Martin Ford's book are problems with DNNs. The SP System provides solutions or potential solutions to most of those problems, and some others.

Wolff, J. G. (2019). Information compression as a unifying principle in human learning, perception, and cognition, *Complexity*, vol. 2019, Article ID 1879746, 38 pages. *DOI*: doi.org/10.1155/2019/1879746, *PDF*: bit.ly/2GdlqnY.

Palade, V., Wolff, J. G. (2019). A roadmap for the development of the 'SP Machine' for artificial intelligence, *The Computer Journal*, *DOI*: doi.org/10.1093/comjnl/bxy126, *PDF*: bit.ly/2Vu0M9Q.

Wolff, J.G. (2016). Information compression, multiple alignment, and the representation and processing of knowledge in the brain, *Frontiers in Psychology*, 7, 1584, 2016, *PDF*: bit.ly/2esmYyt. This paper describes *SP-Neural*, a version of the SP Theory of Intelligence expressed in terms of neurons and their interconnections.

Wolff, J. G. (2013). The SP theory of intelligence: an overview, *Information*, 4 (3), 283-341, 2013, *PDF*: bit.ly/1NOMJ6l. This provides a fairly full description of the SP System. More detail may be found in the book *Unifying Computing and Cognition* by J Gerard Wolff, available via Amazon and other distributers detailed on bit.ly/2JxhJwO.

## Personal Response

***What prompted you to research and develop the SP System?***

" Information compression as an important principle in the workings of brains and nervous systems was the subject of fascinating lectures given by Dr Horace Barlow (now Professor Barlow FRS) when I was a student at Cambridge University. My later research developing computer models of language learning confirmed the importance of information compression in the learning of language. Later experience in commercial software development suggested the potential of a new kind of computing system with a central role for information compression. Experience in developing the SP System, with information compression at centre stage, has progressively expanded and refined the ideas. "

**What do you envisage as the next development of the SP System?**

" As mentioned earlier, three of us are aiming to create a first version of the *SP Machine*, derived from the SP Computer Model with the application of high levels of parallel processing. A peer-reviewed paper describes a roadmap for the development of an industrial-strength SP Machine. "