Article: Intelligence Via Compression of Information

The SP-Theory of Intelligence, and its realization in the SP Computer Model, can demonstrate

diverse aspects of intelligence, all achieved by compressing information via the powerful

SP-Multiple-Alignment construct. Quite unexpectedly, these ideas have led to an important new insight in the foundations of mathematics, and to new thinking about the concepts of probability and computing.

J Gerard Wolff*

Tuesday 19th December, 2023, 10:28

1 Introduction

As the title of this article suggests, it is about intelligence and how it may be understood as compression of information. More specifically, the article is a relatively short introduction to the book [1], which is itself is about the *SP-Theory of Intelligence* (SPTI), its realisation in the SP *Computer Model* (SPCM), and their potential benefits and applications, and associated ideas. Unless otherwise stated, all statements about the SPTI should be understood to apply also to the SPCM.

Since deep neural networks (DNNs) have yielded some impressive results and are currently the centre of much attention, readers may ask why anyone should pay attention to the SPTI and what it can do. In brief, the SPTI has been developed to create a firm *foundation* for the development of human-level intelligence, aka artificial general intelligence (AGI), and a major part of the theory is a conceptual framework—*SP-Multiple-Alignment*—that can model many more aspects of intelligence than can DNNs. Some of that versatility is described in [2].

The main things described in this article are:

- Some of the substantial evidence for the importance of information compression (IC) in the workings of brains and nervous systems (Section 2). This evidence is the reason for making IC central in the design of the SPTI and SPCM.
- Within a description of the SPTI, an important part is the powerful concept of *SP-Multiple-Alignment*, the key to the versatility of the SPTI in modelling diverse aspects of intelligence (Sections 4.3).
- The article presents several examples showing how the SP-Multiple-Alignment concept may model several aspects of intelligence (Section 5),

^{*}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; X: @gerrywolff65; Web: www.cognitionresearch.org.

- The idea that intelligence may be understood as IC leads naturally to a major discovery in the foundations of mathematics: that mathematics may be understood as a set of techniques for IC, and their application (Section 6).
- On the strength of this insight a *New Mathematics* is proposed as an integration of mathematics with the SPTI, with several potential benefits and applications (Section 7).
- Other features of the SPTI are summarised in a final section (Section 8).

2 Information Compression in Brains and Nervous Systems

A basis for the SPTI is substantial evidence for the importance of information compression (IC) in the workings of brains and nervous systems. This section describes a little of this evidence, drawing on [3], where further information may be found.

Fred Attneave [4], one of the pioneers of this area of research, describes evidence that visual perception may be understood in terms of the distinction between areas in a visual image where there is much redundancy, and boundaries between those areas where non-redundant information is concentrated. He writes [4, p. 184]:

... information is concentrated along contours (i.e., regions where color changes abruptly), and is further concentrated at those points on a contour at which its direction changes most rapidly (i.e., at angles or peaks of curvature).

And he suggests that:

Common objects may be represented with great economy, and fairly striking fidelity, by copying the points at which their contours change direction maximally, and then connecting these points appropriately with a straight edge.

He illustrates the point with a drawing of a sleeping cat reproduced in Figure 1.



Figure 1: Drawing made by abstracting 38 points of maximum curvature from the contours of a sleeping cat, and connecting these points appropriately with a straight edge. Reproduced from Figure 3 in [4], with permission.

Horace Barlow, another pioneer in this area, argues that, in mammals at least, each optic nerve is too small, by a wide margin, to carry anything but a small amount of the very large amounts of visual information reaching the retina. Hence, 'the storage and utilization of this enormous sensory inflow would be made easier if the redundancy of the incoming messages was reduced.' [5, p. 537]. Here 'redundancy' means 'repetition of information', so that reducing redundancy means IC. There is now much evidence that neurons in and near to the retina do indeed compress visual information by the process of 'lateral inhibition'.

As early as 1969, Barlow wrote [6, p. 210]:

... the operations needed 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.

where 'find[ing] a less redundant code' leads to 'redundancy reduction' which means lossless IC.

Perhaps more important than the impact of IC on the storage or transmission of information is the close connection between IC and concepts of probability, first identified by Ray Solomonoff [7, 8]. Compression of information provides a means of predicting the future from the past and estimating probabilities so that, for example, an animal may learn to predict where food may be found, or where there may be dangers, and so on—with the same kinds of benefits for people.

3 Origin of the SP Theory of Intelligence

Arriving at a good framework for the SPTI has not been straightforward. Here are a few words about how things developed. An early task was to develop a program for finding 'good' full and partial matches between two sequences of symbols where matches were evaluated in terms of the level of IC that could be achieved by the merging or 'unification' of those matching patterns.

Apart from its relevance to IC, the idea of searching for matching patterns was adopted because it seemed to be relevant to several aspects of human intelligence.

The next step was to look at 'multiple sequence alignments' programs used in biochemistry for finding good full and partial matches between two *or more* sequences of DNA bases or amino-acid residues.

These programs are designed to analyse two, three, four, or more sequences simultaneously and to show one or more of the best results, each as an arrangement of the two or more sequences next to each other, with judicious 'stretching' of sequences in a computer to align symbols that match each other from one sequence to another.

An example of a good 'multiple sequence alignment' of five DNA sequences is shown in Figure 2.

	G	G	А			G			С	А	G	G	G	А	G	G	А			Т	G			G		G	G	А
	Ι	Ι	Ι			Ι			Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι			Ι	Ι			Ι		Ι	Ι	Ι
	G	G	Ι	G		G	С	С	С	А	G	G	G	А	G	G	А			Ι	G	G	С	G		G	G	А
	Ι	Ι	Ι			Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι			Ι	Ι			Ι		Ι	Ι	Ι
А		G	А	С	Т	G	С	С	С	А	G	G	G		G	G		G	С	Т	G			G	А	Ι	G	А
	Ι	Ι	Ι						Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι		Ι		Ι			Ι		Ι	Ι	Τ
	G	G	А	А					Ι	А	G	G	G	А	G	G	А		Ι	А	G			G		G	G	А
	Ι	Ι		Ι					Ι	Ι	Ι	Ι	Ι	Ι	Ι	Ι			Ι		Ι			Ι		Ι	Ι	Ι
	G	G	С	А					С	А	G	G	G	А	G	G			С		G			G		G	G	А

Figure 2: A good multiple sequence alignment amongst five DNA sequences.

From studying this kind of multiple sequence alignment and creating different examples 'manually' with a word processor, I began to realise that multiple sequence alignments, or something like them, could provide a very versatile model for different aspects of AI.

Although the *SP-Multiple-Alignment* concept as it is now may seem straightforward, its development took several years' work. The process of developing the concept required the creation and testing of *hundreds* of versions of the SPCM.

4 The SP Theory of Intelligence and Its Realisation in the SP Computer Model

The SPTI is conceived as a brain-like system as shown in Figure 3, with *New* information (green) coming in via the senses (eyes and ears in the figure), and with some or all of that information compressed and stored as *Old* information show in red in the brain.

The processing of New information to create Old information is central in all processing by the SPCM.



Figure 3: Schematic representation of the SPTI showing incoming, not compressed 'sensory', information in green, and showing compressed information, stored in the system's brain, in red.

4.1 SP-patterns and SP-symbols

In the SPTI, all kinds of knowledge are represented by *SP-patterns*, where an SP-pattern is an array of *SP-symbols* in one or two dimensions.

An SP-symbol is simply a mark from an alphabet of alternatives where each SP-symbol can be matched in a yes/no manner with any other SP-symbol.

An SP-symbol does not have any hidden meaning, such as 'add' for the symbol '+' in arithmetic, or 'multiply' for the symbol ' \times ', and so on. Any meaning attaching to an SP-symbol is provided by one or more other SP-symbols with which it is associated.

This kind of simplicity, combined with the relatively simple but powerful concept of SP-Multiple-Alignment (Section 4.3) is the key to the seamless integration of diverse aspects of AI and diverse kinds of intelligence-related knowledge, in any combination.

Examples of SP-patterns may be seen in Figure 4, as described in the caption to the figure.

4.2 Two-Dimensional SP-Patterns

At present, the SPCM works only with one-dimensional SP-patterns but it is envisaged that the SPCM will be generalised to work with two-dimensional SP-patterns, as well as one-dimensional SP-patterns. It is envisaged that 2D SP-patterns may be of any shape, not necessarily simple rectangles

The addition of 2D SP-patterns should open up the system for the representation and processing of diagrams and pictures. As described in [9, Sections 6.1 and 6.2)], 2D SP-patterns may serve in the representation of structures in three dimensions. And 2D SP-patterns may also have a role in the representation and processing of parallel streams of information, as described in [10, Sections V-G, V-H, and V-I, and Appendix C]. SP-patterns in one or two dimensions may also serve in representing the time dimension in videos and the like.

4.3 The SP-Multiple-Alignment Concept

The SP-Multiple-Alignment concept is described in outline here. More detail may be found in [1, Section 4.3–4.6], [11, Section 3.4] and [12, Section 4].

The SP-Multiple-Alignment concept is largely responsible for the strengths of the SPTI in modelling aspects of intelligence and some other strengths, summarised in Section ??. Although it is far from being trivially simple, the SP-Multiple-Alignment concept is remarkably simple as the source of most of the strengths of the SPTI. In short, the relative simplicity of the SPTI combined with its high descriptive and explanatory power, is largely due to the SP-Multiple-Alignment concept.

In the light of the foregoing remarks, it is appropriate to say that the SP-Multiple-Alignment concept is a major discovery with the potential to be as significant for an understanding of intelligence as is the concept of DNA for an understanding of biology. It may prove to be the 'double helix' of intelligence!

4.3.1 Organisation of the SP-Multiple-Alignment Concept

As described above, the SP-Multiple-Alignment concept in the SPTI has been borrowed and adapted from the concept of 'multiple sequence alignment' in bioinformatics. An example of an SP-Multiple-Alignment is shown in Figure 4.



Figure 4: The best SP-Multiple-Alignment created by the SPCM that achieves the effect of parsing a sentence ('t h e p l u m s a r e r i p e') into its parts and sub-parts, as described in the text. The sentence in row 0 ia a New SP-pattern, while each of the rows 1 to 9 contains a single Old SP-pattern, drawn from a relatively-large repository of Old SP-patterns in the SP-grammar shown in Figure 5. Reproduced from Figure A2 in [13].

The main components of the SP-Multiple-Alignment concept, illustrated in the Figure 4, are these:

- Row 1 contains one New SP-pattern representing information that has been received recently from the system's environment (Section ??). Sometimes row 1 contains more than one New SP-pattern. In the example shown, the SP-pattern in row 1 is a sentence but in other SP-Multiple-Alignments the New SP-pattern may represent any other kind of information.
- Each of rows 1 to 8 contains a single Old SP-pattern, drawn from a relatively large repository of Old SP-patterns. In this case, that repository of Old SP-patterns is the SP-grammar shown in Figure 5, and each Old SP-pattern in the SP-Multiple-Alignment and in the SP-grammar represents a grammatical structure including words. More generally, Old SP-patterns may represent any other kind of information.

!S Num ; NP #NP VP !#VP #S !NP !O NP #NP Q #Q !#NP !NP !Oa D #D N #N !#NP !VP !1 V #V A #A !#VP !Q !1 P #P NP #NP !#Q !N !Ns !2 i t !#N !N !Ns !3 s h e !#N !N !Ns !4 w e s t !#N !N !Ns !5 o n e !#N !Nrt !7 w i n d !#Nrt !Nrt !6 p l u m !#Nrt !N !Npl !8 s i x !#N !N !Npl !9 s e v e n !#N !N !Npl Nrt #Nrt s !#N !V !Vs !10 d o e s !#V !V !Vs !10a i s !#V !V !Vpl !11 are !#V !V !Vpl !12 p l a y !#V !V !Vpl !13 d o !#V !P !14 w i t h !#P !P !15 f r o m !#P !P !16 o f !#P !D !17 t h e !#D !D !18 a !#D !A !19 s t r o n g !#A !A !20 w e a k !#A !A !21 r i p e !#A !Num !SNG !; Ns Vs !Num !PL !; Npl Vpl

Figure 5: In this SP-grammar, each SP-pattern starts and ends with SP-symbols representing a grammatical category which are called 'ID' SP-symbols. The character '!' in the SP-grammar serves to mark a symbol as being an 'ID-symbol' (Section 4.1) and, in each SP-pattern, the unmarked SP-symbols are 'C' or 'contents' SP-symbols.

4.3.2 How SP-Multiple-Alignments Are Built Up

Here is a summary of how SP-Multiple-Alignments like the one shown in Figure 4 are built up:

1. At the beginning of processing, the SPCM has the afore-mentioned store of Old SP-patterns (Figure 5).

When the SPCM is more fully developed, those Old SP-patterns would have been learned from raw data as outlined in Section ??, but for now they are supplied to the program by the user.

- 2. The next step is to read in the New SP-pattern, 't h e p l u m s a r e r i p e', shown in row 0 of Figure 4.
- 3. Then the program searches for 'good' matches between the New SP-pattern and Old SPpatterns, where 'good' matches are ones that yield relatively high levels of compression of the New SP-pattern in terms of Old SP-pattern(s) with which it has been unified.
- 4. As can be seen in the figure, matches are identified at early stages between (parts of) the New SP-pattern and the Old SP-patterns 'D 17 t h e #D', 'Nrt 6 p l u m #Nrt', 'V Vpl 11 a r e #V', and 'A 21 r i p e #A'. Although this is not shown in this example, the SPCM also searches for matches within the New SP-pattern.
- 5. In SP-Multiple-Alignments, IC may be achieved by collapsing the whole SP-Multiple-Alignment into a single sequence of symbols and thus unifying matching patterns within the SP-Multiple-Alignment, like the match between 't h e' in the New SP-pattern and the same three letters in the Old SP-pattern 'D 17 t h e #D'. In practice, this is not done explicitly but only notionally to achieve a measure of IC for the whole SP-Multiple-Alignment and for each partial SP-Multiple-Alignment created in the course of building the whole SP-Multiple-Alignment.
- 6. The unification of 't h e' with 'D 17 t h e #D' yields the unified SP-pattern 'D 17 t h e #D', with exactly the same sequence of SP-symbols as the second of the two SP-patterns from which it was derived.
- 7. Details of how IC for any one full or partial SP-Multiple-Alignment is calculated are given in Section ??.
- 8. As processing proceeds, similar pair-wise matches and unifications eventually lead to the creation of SP-Multiple-Alignments like that shown in Figure 4. At every stage, all the SP-Multiple-Alignments that have been created are evaluated in terms of IC, and then the best SP-Multiple-Alignments are retained and the remainder are discarded. In this case, the final 'winner' is the SP-Multiple-Alignment shown in Figure 4.
- 9. This process of searching for good SP-Multiple-Alignments in stages, with selection of good partial solutions at each stage, is an example of heuristic search. This kind of search is necessary because there are too many possibilities for anything useful to be achieved by exhaustive search. By contrast, heuristic search can normally deliver results that are reasonably good within a reasonable computational complexity—but it cannot guarantee that the best possible solution has been found.

10. A simple but important detail is that, within an SP-grammar, each SP-pattern occurs only once, but within an SP-Multiple-Alignment, and SP-pattern may appear two or more times. This may be seen in Figure ??, in which the SP=pattern 'NP D #D N #N #NP' appears twice.

4.4 Discontinuous Constituents and Their Representation in an SP-Multiple-Alignment

Regarding the SP-Multiple-Alignment shown in Figure 4, an aspect not mentioned so far is the role of the SP-pattern shown in row 8.

Clues to the role of that SP-pattern lie in the SP-symbols 'PL', 'Npl', and 'Vpl' within the SPpattern in row 8. The first of these SP-symbols, 'PL', shows that the sentence has a 'plural' form. The second of those SP-symbols, 'Npl', marks the subject noun, 'p l u m s' as having the plural form. And the third of those SP-symbols, 'Vpl', marks the main verb as having the plural form.

So in summary, the role of the SP-pattern in row 8, which is the last SP-pattern in the SPgrammar in Figure 5, is to encode the syntactic rule in English and many other natural languages that if the subject of the sentence has a plural form, then the main verb should also have a plural form.

As can be seen in the SP-grammar in Figure 5, the last but one SP-pattern, '!Num !SNG !; Ns Vs', is the one that would be called into play to show the dependency between a singular subject noun and a singular main verb.

These dependencies are called 'discontinuous' because they can jump over any amount of intervening structure.

Arguably, this method for representing discontinuous dependencies in syntax is more elegant than the standard computer science method using variables, described in, for example, [14, Chapter 12]. The method illustrated in Figure 4 and described in [11, Section 5.4] has the merit of growing seamlessly out of the SP-Multiple-Alignment method for representing and processing linguistic information (amongst other kinds of information), without the need for any *ad hoc* addition to the method.

There is more detail about these kinds of discontinuous dependency in Sections 5.4 and 5.5.

4.5 The Main Strengths of the SP Theory of Intelligence

The main strengths of the SPTI are summarised in [1, Section 6]. This section is a shortened version of that description.

- Aspects of intelligence exhibited by the SPCM.
 - [1, Section 6.1.1] describes aspects of intelligence exhibited by the SPCM, excluding probabilistic reasoning.
 - [1, Section 6.1.2] provides a separate section about probabilistic reasoning via the SPCM.
 - [1, Section 6.1.3] is about the representation and processing of several kinds of intelligencerelated knowledge.
 - [1, Section 6.1.4] is about how the SPCM has clear potential to provide seamless integration of diverse aspects of intelligence and diverse kinds of intelligence-related knowledge, in any combination.

It is suggested that this is because of the provision of a uniform framework for all aspects of intelligence and all kinds of intelligence-related knowledge.

And 'It appears that this kind of seamless integration is *essential* in any artificial system that aspires to AGI.'

- Other aspects of intelligence relating to the SPTI.
 - [1, Section 6.2.1] Describes the clear potential of the SPTI to solve twenty significant problems in artificial intelligence research. This is a summary of [2].
 - [1, Section 6.2.2] Describes how the SPTI may be seen as a *foundation* for the Development of artificial general intelligence
 - [1, Section 6.2.3] contains a note about commonsense reasoning and commonsense knowledge.
- [1, Section 6.3] describes some other potential benefits and applications of the SPTI.

5 Examples of the Versatility of the SPCM

As the title of this section suggests, it contains examples of SP-Multiple-Alignments showing the kinds of things that may be done with the SPCM. These examples show some of the versatility of the SP-Multiple-Alignment concept, but SP-Multiple-Alignments are very much more versatile than these few examples may suggest.

5.1 Recursive Processing and the SP Theory of Intelligence

This subsection shows, with the recognition of a palindrome as an example, how the SPCM may accommodate recursive processing. There are two other examples of recursive processing in Figure ?? in Section ??.

Regarding the recognition of a palindrome, Figure 7 shows the best SP-Multiple-Alignment produced by the SPCM with 'a c b a b a b c a' in New and, in Old, the SP-patterns under the heading 'Old' in Figure 6. The SP-Multiple-Alignment may be seen as a recognition that the pattern in New is indeed a palindrome.

New a c b a b a b c a Old L a #L L b #L L c #L L1 a #L1 L2 a #L2 L1 b #L1 L2 b #L2 L1 c #L1 L2 c #L2 X L #L #X X L1 #L1 X #X L2 #L2 #X

Figure 6: SP-atterns for processing by the SPCM to model the recognition of a palindrome.



Figure 7: The best SP-Multiple-Alignment (in terms of compression) produced by the SPCM with the SP-patterns from Figure 6 in New and Old, as shown in that figure.

5.2 Ambiguities in Language

Natural languages are notoriously ambiguous, not only in their meanings but also in their syntax. An example which includes syntactic ambiguity is the second sentence in 'Time flies like an arrow. Fruit flies like a banana', with uncertain origins.¹

Figure 8 shows how the SPCM can accommodate the ambiguity of that example, given an appropriate grammar. In this example, the two parsings shown have compression values that are similar and higher than the compression scores of other alignments formed for the same sentence.



Figure 8: The two best SP-Multiple-Alignments found by the SPCM with SP-patterns representing grammatical rules including words (in the repository of Old SP-patterbs) and the ambiguous sentence 'fruit flies like a banana' (which is a New SP-pattern). Reproduced from [11, Figure 5.1].

¹Based on [11, Section 5.2.2].

5.3 Robustness in the Face of Errors of Omission, Addition, and Substitution

Figure 9 (a) shows how the SPTI may achieve a 'correct' parsing of the sentence 't w o k i t t e n s p l a y', while Figure 9 (b) shows how the SPCM achieves the same 'correct' parsing, except that the sentence contains: an error of omission when the letter 'w' is missing from the word 't w o'; an error of substitution when the letter 'm' replaces the letter 'n' in the word 'k i t t e n s', and an error of addition when the letter 'x' has been added to the word 'p l a y'.²

In effect, the parsing identifies errors in the sentence and suggests corrections for them: 't o' should be 't w o', 'k i t t e m s' should be 'k i t t e n s', and 'p l a x y' should be 'p l a y'.

Examples like this suggest that—by contrast with the way in which DNNs are liable to make large and unexpected errors in recognition ([2, Section 3] and bullet point 2 in Section ??)—the SPTI and its realisation in the SPCM have robust capabilities for recovering from errors in its data.

²This section is based on [12, Section 4.2.2].



Figure 9: (a) The best SP-Multiple-Alignment created by the SP model with a New SP-pattern (representing a sentence to be parsed) shown in row 0, and a store of Old SP-patterns like those in rows 1 to 8 (representing grammatical structures, including words); and (b) The same as in (a) but with errors of omission, substitution and commission as described in the text, and with same set of Old SP-patterns as before. (a) and (b) are reproduced from Figures 1 and 2 respectively in [15], with permission.

5.4 Syntactic Dependencies in French

As we have seen in Section 4.4, sentences in natural languages may contain syntactic dependencies between one part of a sentence and another.³ As described in that section, there is usually a 'number' dependency between the subject of a sentence and the main verb of the sentence: if the subject has a *singular* form then the main verb must have a singular form and likewise for *plural* forms of the subject of a sentence and the main verb.

A prominent feature of these kinds of dependency is that they are often 'discontinuous' in the sense that the elements of the dependency can be separated, one from the next, by arbitrarily large amounts of intervening structure. For example, the subject and main verb of a sentence must have the same number (singular or plural) regardless of the size of qualifying phrases or subordinate clauses that may come between them.

Another interesting feature of syntactic dependencies, not discussed in Section 4.4, is that one kind of dependency, such as number dependency (*singular/plural*) can overlap other kinds of dependency, such as gender dependency (*masculine/feminine*), as can be seen in the following example.

In the French sentence *Les plumes sont vertes* ('The feathers are green') there are two sets of overlapping syntactic dependencies like this:

Р		Ρ	Р			Ρ	Number dependencies
Les	plume	s	sont	vert	е	s	
	F				F		Gender dependencies

In this example, there is a number dependency, which is plural ('P') in this case, between the subject of the sentence, the main verb and the following adjective: the subject is expressed with a plural determiner (*Les*) and a noun (*plume*) which is marked as plural with the suffix (*s*); the main verb (*sont*) has a plural form and the following adjective (*vert*) is marked as plural by the suffix (*s*).

Cutting right across these number dependencies is the gender dependency, which is feminine (F') in this case, between the feminine noun (*plume*) and the adjective (*vert*) which has a feminine suffix (*e*).

For many years, linguists puzzled about how these kinds of syntactic dependency could be represented succinctly in grammars for natural languages. But then elegant solutions were found in transformational grammar [16] and, later, in systems like definite clause grammars [17], based on Prolog [18].

The solution proposed here is different from any established system and is arguably simpler and more transparent than other systems. It is described and illustrated here with a fragment of an SP-grammar of French, shown in Figure 10, which can generate the example sentence just shown.

Apart from the use of SP-patterns as the medium of expression, this SP-grammar differs from systems like transformational grammar or definite clause grammars because the parts of the SPgrammar which express the forms of 'high level' structures like sentences, noun phrases and verb phrases (represented by the first four SP-patterns in Figure 10) do not contain any reference to number or gender. Instead, the SP-grammar contains the eight SP-patterns shown at the end of Figure 10.

In the SP-Multiple-Alignment shown in Figure 11, and in some that come later, SP-Multiple-Alignments are shown with SP-patterns arranged in columns instead of rows. The choice of one arrangement or the other depends mainly on what fits best on the page. With SP-Multiple-Alignments

³Based on [11, Section 5.4].

S NP #NP VP #VP #S (500) NP D #D N #N #NP (700) VP 0 V #V A #A #VP (300) VP 1 V #V P #P NP #NP #VP (200) P 0 sur #P (50) P 1 sous #P (150) V SNG est #V (250) V PL sont #V (250) D SNG M 0 le #D (90) D SNG M 1 un #D (120) D SNG F 0 la #D (130) D SNG F 1 une #D (110) D PL 0 les #D (125) D PL 1 des #D (125) N NR #NR NS1 #NS1 #N (450) NS1 SNG - #NS1 (250) NS1 PL s #NS1 (200) NR M papier #NR (300) NR F plume #NR (400) A AR #AR AS1 #AS1 AS2 #AS2 #A (300) AS1 F e #AS1 (100) AS1 M - #AS1 (200) AS2 SNG - #AS2 (175) AS2 PL s #AS2 (125) AR 0 noir #AR (100) AR 1 vert #AR (200) NP SNG SNG #NP (450) NP PL PL #NP (250) NP M M #NP (450) NP F F #NP (250) N SNG V SNG A SNG (250) N PL V PL A PL (250) N M V A M (300) NFVAF (400)

Figure 10: A fragment of French SP-grammar with SP-patterns for number dependencies and gender dependencies—the last eight SP-patterns in this SP-grammar.

in rows, New information is always in row 0 and Old information is in the remaining rows, one SPpattern in each row. When SP-Multiple-Alignments are in columns, New information is always in column 0 and Old information is always in the remaining columns, one SP-pattern in each column.

The SP-Multiple-Alignment in the figure shows the best SP-Multiple-Alignment found by the SPCM with our example sentence in New and the SP-grammar from Figure 10 in Old. The main constituents of the sentence are marked in an appropriate manner and dependencies for number and gender are marked by SP-patterns appearing in columns 13, 14 and 15 of the SP-Multiple-Alignment.



Figure 11: The best SP-Multiple-Alignment found by the SPCM with the New SP-pattern in column 0, 'les plume s sont vert e s', representing a sentence, and Old SP-patterns, one in each of columns 1 to 15 representing a grammatical structure or word, drawn from the SP-grammar shown in Figure 10. Reproduced from [11, Figure 5.8].

5.5 Dependencies in the Syntax of English Auxiliary Verbs

This section presents an SP-grammar and examples showing how the syntax of English auxiliary verbs may be described in the SPTI.⁴ Before the SP-grammar and examples are presented, the syntax of this part of English is described and alternative formalisms for describing the syntax are briefly discussed.

In English, the syntax for main verbs and the 'auxiliary' verbs which may accompany them follows two quasi-independent patterns of constraint which interact in an interesting way.

The primary pattern of constraint may be expressed with this sequence of symbols,

М Н В В V,

which should be interpreted in the following way:

- Each letter represents a category for a single word:
 - 'M' stands for 'modal' verbs like 'will', 'can', 'would', etc.
 - 'H' stands for one of the various forms of the verb 'to have'.
 - Each of the two instances of 'B' stands for one of the various forms of the verb 'to be'.
 - 'V' stands for the main verb which can be any verb except a modal verb (unless the modal verb is used by itself).
- The words occur in the order shown but any of the words may be omitted.
- Questions of 'standard' form follow exactly the same pattern as statements except that the first verb, whatever it happens to be ('M', 'H', the first 'B', the second 'B', or 'V'), precedes the subject noun phrase instead of following it.

Here are two examples of the primary pattern with all of the words included:

It will have been being washed M H B B V Will it have been being washed? M H B B V

The secondary constraints are these:

- Apart from the modals, which always have the same form, the first verb in the sequence, whatever it happens to be ('H', the first 'B', the second 'B' or 'V'), always has a 'finite' form (the form it would take if it were used by itself with the subject).
- If an 'M' auxiliary verb is chosen, then whatever follows it ('H', first 'B', second 'B', or 'V') must have an 'infinitive' form (i.e., the 'standard' form of the verb as it occurs in the context 'to ...', but without the word 'to').

⁴This section is based on [11, Section 5.5].

- If an 'H' auxiliary verb is chosen, then whatever follows it (the first 'B', the second 'B', or 'V') must have a past tense form such as 'been', 'seen', 'gone', 'slept', 'wanted', etc. In Chomsky's *Syntactic Structures* [16], these forms were characterised as *en* forms and the same convention has been adopted here.
- If the first of the two 'B' auxiliary verbs is chosen, then whatever follows it (the second 'B' or 'V') must have an *ing* form, e.g., 'singing', 'eating', 'having', 'being', etc.
- If the second of the two 'B' auxiliary verbs is chosen, then whatever follows it (only the main verb is possible now) must have a past tense form (marked with *en* as above).
- The constraints apply to questions in exactly the same way as they do to statements.

Figure 12 shows a selection of examples with the dependencies marked.

		H	en	B2	en
It	will	have	 been	 being	 washed
	 М	 inf	 B1	 ing	 V
Wil: M	l he -	B1 be t 	ing alking V	?	
The	y hav H- fir	V 7e fin 1	 ished en		
Are B2- fin	they	y gone en V	?		
Has	he	B1 been 	in workin 	୍ୟ ଅ?	
H- fin		en	V		

Figure 12: A selection of example sentences in English with markings of dependencies between the verbs. Key: 'M' = modal, 'H' = forms of the verb 'have', 'B1' = first instance of a form of the verb 'be', 'B2' = second instance of a form of the verb 'be', 'V' = main verb, 'fin' = a finite form, 'inf' = an infinitive form, 'en' = a past tense form, 'ing' = a verb ending in 'ing'.

5.5.1 Transformational Grammar and English Auxiliary Verbs

In Figure 12 it can be seen that in many cases but not all, the dependencies which have been described may be regarded as discontinuous because they connect one word in the sequence to the suffix of the following word thus bridging the stem of the following word. Dependencies that are discontinuous can be seen most clearly in questions (e.g., the second, fourth and fifth sentences in Figure 12) where the verb before the subject influences the form of the verb that follows immediately after the subject.

In his book *Syntactic Structures*, [16], Noam Chomsky showed that this kind of regularity in the syntax of English auxiliary verbs could be described using transformational grammar. For each pair of symbols linked by a dependency ('M inf', 'H en', 'B1 ing', 'B2 en') the two symbols could be shown together in the 'deep structure' of a sentence and then moved into their proper position or modified in form (or both) using 'transformational rules'.

This elegant demonstration argued persuasively in favour of transformational grammar compared with alternatives which were available at that time. However, as noted in Section 5.4, later research has shown that the same kinds of regularities in the syntax of English auxiliary verbs can be described quite well without recourse to transformational rules, using definite clause grammars or other systems which do not use that type of rule [17, 18]. An example showing how English auxiliary verbs may be described using the definite clause grammar formalism may be found in [19, pp. 183-184]).

5.5.2 English Auxiliary Verbs in the SPTI

Figure 13 shows an SP-grammar for English auxiliary verbs which exploits several of the ideas described above. Figures 14, 15 and 16 show the best SP-Multiple-Alignments in terms of information compression for three different sentences parsed by the SPCM model using the SP-grammar in Figure 13. In the following paragraphs, aspects of the SP-grammar and of the examples are described and discussed.

S ST NP #NP X1 #X1 XR #S (3000) S Q X1 #X1 NP #NP XR #S (2000) NP SNG it #NP (4000) NP PL they #NP (1000) X1 0 V M #V #X1 XR XH XB XB XV #S (1000) X1 1 XH FIN #XH #X1 XR XB XB XV #S (900) X1 2 XB1 FIN #XB1 #X1 XR XB XV #S (1900) X1 3 V FIN #V #X1 XR #S (900) XH V H #V #XH XB #S (200) XB XB1 #XB1 XB #S (300) XB XB1 #XB1 XV #S (300) XB1 V B #V #XB1 (500) XV V #V #S (5000) M INF (2000) H EN (2400) B XB ING (2000) B XV EN (700) SNG SNG (2500) PL PL (2500) V M O will #V (2500) V M 1 would #V (1000) V M 2 could #V (500) V H INF have #V (600) V H PL FIN have #V (400) V H SNG FIN has #V (200) V H EN had #V (500) V H FIN had #V (300) V H ING hav ING1 #ING1 #V (400) V B SNG FIN 0 is #V (500) V B SNG FIN 1 was #V (400) V B INF be #V (400) V B EN be EN1 #EN1 #V (600) V B ING be ING1 #ING1 #V (700) V B PL FIN 0 are #V (300) V B PL FIN 1 were #V (500) V FIN wrote #V (166) V INF 0 write #V (254) V INF 1 chew #V (138) V INF 2 walk #V (318) V INF 3 wash #V (99) V ING O chew ING1 #ING1 #V (623) V ING 1 walk ING1 #ING1 #V (58) V ING 2 wash ING1 #ING1 #V (102) V EN 0 made #V (155) V EN 1 brok EN1 #EN1 #V (254) V EN 2 tak EN1 #EN1 #V (326) V EN 3 lash ED #ED #V (160) V EN 4 clasp ED #ED #V (635) V EN 5 wash ED #ED #V (23) ING1 ing #ING1 (1883) EN1 en #EN1 (1180) ED ed #ED (818)

Figure 13: An SP-grammar for the syntax of English auxiliary verbs.

5.5.3 The Primary Constraints

The first line in the SP-grammar is a sentence pattern for a statement (marked with the symbol 'ST') and the second line is a sentence pattern for a question (marked with the symbol 'Q'). Apart from these markers, the only difference between the two patterns is that, in the statement pattern, the symbols 'X1 #X1' follow the noun phrase symbols ('NP #NP'), whereas in the question pattern they precede the noun phrase symbols. As can be seen in the examples in Figures 14, 15 and 16, the pair of symbols, 'X1 #X1', has the effect of selecting the first verb in the sequence of auxiliary verbs and ensuring its correct position with respect to the noun phrase. In Figure 14 it follows the noun phrase, while in Figures 15 and 16 it precedes the noun phrase.

Each of the next four patterns in the SP-grammar have the form 'X1 ... #X1 XR ... #S'. The symbols 'X1' and '#X1' align with the same pair of symbols in the sentence pattern. The symbols 'XR ... #S' encode the remainder of the sequence of verbs.

The first 'X1' pattern encodes verb sequences which start with a modal verb ('M'), the second one is for verb sequences beginning with a finite form of the verb 'have' ('H'), the third is for sequences beginning with either of the two 'B' verbs in the primary sequence (see below), and the last 'X1' pattern is for sentences which contain a main verb without any auxiliaries.

In the first of the 'X1' patterns, the subsequence 'XR \dots #S' encodes the remainder of the sequence of auxiliary verbs using the symbols 'XH XB XB XV'. In a similar way, the subsequence 'XR \dots #S' within each of the other 'X1' patterns encodes the verbs which follow the first verb in the sequence.

Notice that the pattern 'X1 2 XB1 FIN #XB1 #X1 XR XB XV #S' can encode sentences which start with the first 'B' verb and also contains the second 'B' verb. And it also serves for any sentence which starts with the first or the second 'B' verb with the omission of the other 'B' verb. In the latter two cases, the 'slot' between the symbols 'XB' and 'XV' is left vacant. Figure 14 illustrates the case where the verb sequence starts with the first 'B' verb with the second 'B' verb. Figure 16 illustrates the case where the verb sequence starts with the second 'B' verb (and the first 'B' verb has been omitted).



Figure 14: The best SP-Multiple-Alignment found by the SPCM with 'it is wash ed' in New and the SP-grammar from Figure 13 in Old. Reproduced from [11, Figure 5.12].



Figure 15: The best SP-Multiple-Alignment found by the SPCM with 'will it have be en brok en' in New and the SP-grammar from Figure 13 in Old. Reproduced from [11, Figure 5.13].



Figure 16: The best SP-Multiple-Alignment found by the SPCM with 'are they walk ing' in New and the SP-grammar from Figure 13 in Old. Reproduced from [11, Figure 5.14].

5.5.4 The Secondary Constraints

The secondary constraints (Section 5.5) are represented using the patterns 'M INF', 'H EN', 'B XB ING' and 'B XV EN'. Singular and plural dependencies are marked in a similar way using the patterns 'SNG SNG' and 'PL PL'.

Examples appear in all three SP-Multiple-Alignments in Figures 14, 15 and 16. In every case except one (column 4 in Figure 14), the patterns representing secondary constraints appear in the later columns of the SP-Multiple-Alignment (towards the right). These examples show how dependencies bridging arbitrarily large amounts of structure, and dependencies that overlap each other, can be represented with simplicity and transparency in the medium of SP-Multiple-Alignments.

Notice, for example, how dependencies between the first and second verb in a sequence of auxiliary verbs are expressed in the same way regardless of whether the two verbs lie side by side (e.g., the statement in Figure 14) or whether they are separated from each other by the subject noun-phrase (e.g., the question in Figure 15 and in Figure 16). Notice, again, how the overlapping dependencies in Figure 15 and their independence from each other are expressed with simplicity and clarity in the SPTI.

5.6 The Integration of Syntax With Semantics

In keeping with the remarks about the integration of diverse kinds of knowledge in Sections ??, 4.1, and ??, it has been anticipated that the SPTI would not only support the representation of syntactic and non-syntactic ('semantic') kinds of knowledge but that it would facilitate their integration.⁵

A preliminary example of how this might be done is shown in Figure 18. This is the best SP-Multiple-Alignment produced by the SPCM with 'john kissed mary' as the New pattern and a grammar in Old that contains patterns representing syntax, 'semantics' and the connections between them. The scare quotes are intended to indicate that the representations of semantic structures in this example are, at best, crude. That point made, the quote marks for 'semantics' or 'meanings' will be dropped in the remainder of this section.

In the SP-Multiple-Alignment, the sentence appears in column 0. In the remaining columns, Old patterns with the main roles are as follows:

- The pattern in column 2 represents the overall syntactic structure of the sentence: 'S AR NP #NP #AR AN V1 #V1 #AN OBJ NP #NP #OBJ #S'. Notice that this pattern differs from comparable patterns shown in previous examples because each constituent within the pattern is marked with its semantic role. Thus the first noun phrase ('NP #NP') is enclosed by the pair of symbols 'AR ... #AR' (representing the 'actor' role), the verb ('V1 #V1') is marked as an 'action' with the symbols 'AN ... #AN', and the second noun phrase is marked as an 'object' ('OBJ ... #OBJ').
- In column 7, the pattern 'A X an #an #X Y obj #obj #Y Z ar #ar #Z #A' may be seen as a generalised description of the association between an 'action' ('an #an'), the 'object' of the action ('obj #obj') and the 'actor' or performer of the action ('ar #ar'). Notice that the order in which these concepts are specified is different from the order of the corresponding markers in the syntax pattern. Notice also that these three slots are also marked, in order, as 'X ... #X', 'Y ... #Y' and 'Z ... #Z'. The reason for this marking will be seen shortly.

⁵This section is based on [11, Section 5.7].



Figure 17: The best SP-Multiple-Alignment created by the SPCM with the sentence 'john kissed mary' in New and a grammar in Old that represents natural language syntax, semantics and their integration. Reproduced from [11, Figure 5.18].

- In column 6, the pattern 'SM S #S A #A #SM' provides a link between the pattern in column 2 representing the syntactic structure of the sentence ('S ... #S') and the action-object-actor pattern in column 7 ('A ... #A').
- In columns 9, 12 and 13 are three more patterns that link the syntax with the semantics. In column 9, the pattern 'OBJ 2 #OBJ Y 1 #Y' connects 'NP 2 mary #NP' in the object position of the syntax with 'MARY' in the 'obj #obj' slot of the semantic structure. Here, 'MARY' is intended to represent some kind of conceptual structure that is the meaning of the word 'mary'. More precisely, it is intended to represent a 'code' for that structure (see Section 5.7.1, below). In a similar way, the pattern in column 12 connects 'kissed' with 'KISSED' in the 'an #an' semantic slot; and the pattern in column 13 connects 'john' with 'JOHN' in the 'ar #ar' semantic slot.

The key idea in this example is that the SPTI allows information to be carried from the syntactic part of the knowledge structure to the semantic part and it allows the ordering of information to change from one part to the other. There seems no reason to suppose that this basic capability could not also be applied to examples in which the syntax and the semantics are more elaborate and more realistic.

5.7 The Integration of Syntax With Semantics

In keeping with the remarks about the integration of diverse kinds of knowledge in Sections ??, 4.1, and ??, it has been anticipated that the SPTI would not only support the representation of syntactic and non-syntactic ('semantic') kinds of knowledge but that it would facilitate their integration.⁶

A preliminary example of how this might be done is shown in Figure 18. This is the best SP-Multiple-Alignment produced by the SPCM with 'john kissed mary' as the New pattern and a grammar in Old that contains patterns representing syntax, 'semantics' and the connections between them. The scare quotes are intended to indicate that the representations of semantic structures in this example are, at best, crude. That point made, the quote marks for 'semantics' or 'meanings' will be dropped in the remainder of this section.

In the SP-Multiple-Alignment, the sentence appears in column 0. In the remaining columns, Old patterns with the main roles are as follows:

- The pattern in column 2 represents the overall syntactic structure of the sentence: 'S AR NP #NP #AR AN V1 #V1 #AN OBJ NP #NP #OBJ #S'. Notice that this pattern differs from comparable patterns shown in previous examples because each constituent within the pattern is marked with its semantic role. Thus the first noun phrase ('NP #NP') is enclosed by the pair of symbols 'AR ... #AR' (representing the 'actor' role), the verb ('V1 #V1') is marked as an 'action' with the symbols 'AN ... #AN', and the second noun phrase is marked as an 'object' ('OBJ ... #OBJ').
- In column 7, the pattern 'A X an #an #X Y obj #obj #Y Z ar #ar #Z #A' may be seen as a generalised description of the association between an 'action' ('an #an'), the 'object' of the action ('obj #obj') and the 'actor' or performer of the action ('ar #ar'). Notice that the order in which these concepts are specified is different from the order of the corresponding markers in the syntax pattern. Notice also that these three slots are also marked, in order, as 'X ... #X', 'Y ... #Y' and 'Z ... #Z'. The reason for this marking will be seen shortly.

⁶This section is based on [11, Section 5.7].



Figure 18: The best SP-Multiple-Alignment created by the SPCM with the sentence 'john kissed mary' in New and a grammar in Old that represents natural language syntax, semantics and their integration. Reproduced from [11, Figure 5.18].

- In column 6, the pattern 'SM S #S A #A #SM' provides a link between the pattern in column 2 representing the syntactic structure of the sentence ('S ... #S') and the action-object-actor pattern in column 7 ('A ... #A').
- In columns 9, 12 and 13 are three more patterns that link the syntax with the semantics. In column 9, the pattern 'OBJ 2 #OBJ Y 1 #Y' connects 'NP 2 mary #NP' in the object position of the syntax with 'MARY' in the 'obj #obj' slot of the semantic structure. Here, 'MARY' is intended to represent some kind of conceptual structure that is the meaning of the word 'mary'. More precisely, it is intended to represent a 'code' for that structure (see Section 5.7.1, below). In a similar way, the pattern in column 12 connects 'kissed' with 'KISSED' in the 'an #an' semantic slot; and the pattern in column 13 connects 'john' with 'JOHN' in the 'ar #ar' semantic slot.

The key idea in this example is that the SPTI allows information to be carried from the syntactic part of the knowledge structure to the semantic part and it allows the ordering of information to change from one part to the other. There seems no reason to suppose that this basic capability could not also be applied to examples in which the syntax and the semantics are more elaborate and more realistic.

5.7.1 Codes, Meanings and the Production of Language From Meanings

Section ??, describes how the SPTI can be used to produce a sentence, given a short code for that sentence supplied as New. That example shows in general terms how the system may be used for language production as well as language analysis but it seems unlikely that there would be many applications where there would be a requirement for the production of sentences purely in terms of their syntax and encodings of that syntax. In practice, it is more likely that one would wish to create sentences on the basis of intended *meanings*.

One possibility is that meanings might serve as codes for syntax and be used for language production in the way described in Section ??. In support of this view, we seem to remember what people have said in terms of meanings that were expressed rather than the words that were used to express them. And we can often reconstruct the words that people have used from the meanings that we remember—although there may be an element of lossy compression here because the reconstruction is not always accurate.

Figure 19 shows how, via the building of an SP-Multiple-Alignment, a sentence may be derived from the pattern 'KISSED MARY JOHN', representing an 'internal' code for the meaning to be expressed. The SP-Multiple-Alignment is the best SP-Multiple-Alignment created by the SPCM with that pattern in New and the same grammar in Old as was used for the example in Figure 18. In the top part of the SP-Multiple-Alignment, the words 'john', 'kissed', and 'mary' appear, in that order. If we strip out the 'service' symbols in the SP-Multiple-Alignment, we have the sentence corresponding to the semantic representation in column 0.



Figure 19: The best SP-Multiple-Alignment created by the SPCM with 'KISSED MARY JOHN' in New and the same grammar in Old as was used for the example shown in Figure 18. Reproduced from [11, Figure 5.19].

5.8 Recognition and Retrieval

In this section, Figure 20 provides an example of an SP-Multiple-Alignment created by the SPCM via a process of recognition, which may also be seen as a process of information retrieval.⁷

In this example, the caption to the figure tells us that this is the best SP-Multiple-Alignment found by the SPCM with the features of an unknown plant in column 0 and with SP-patterns drawn from a repository of Old SP-patterns like those shown in columns 1 to 6. This example shows how the SPCM may identify an unknown plant from its features. The answer of course is that the plant with the features shown in the SP-Multiple-Alignment is of the species 'acris' (Meadow Buttercup) (column 1), in the genus 'Ranunculus' (column 6), which is in the family 'Ranunculaceae' (column 5), and so on.

This process of recognition may also be regarded as a process of retrieval because it retrieves information that was not amongst the features of the unknown plant. We learn, for example, that each Meadow Buttercup photosynthesises (column 2), that the sepals are 'not reflexed' (column 1), that each flower has 5 petals (column 6), and so on.

Of particular interest in this example is that the SP-Multiple-Alignment illustrates how classinclusion relations—for example, genus (column 1), family (column 6), order (column 5), and so on—and part-whole relations—for example, a shoot is compose of a stem, leaves, and flowers—may be combined in one SP-Multiple-Alignment.

⁷This section is based on [12, Section 10.1].



Figure 20: The best SP-Multiple-Alignment created by the SPCM, with a set of New SP-patterns (in column 0) that describe some features of an unknown plant, and a set of Old SP-patterns, including those shown in columns 1 to 6, that describe different categories of plant, with their parts and sub-parts, and other attributes. Reproduced from [12, Figure 16].

5.9 Medical Diagnosis

<patient> John_Smith </patient>
<face> flushed </face>
<appetite> poor </appetite>
<breathing> rapid </breathing>
<muscles> aching </muscles>
<chills> yes </chills>
<fatigue> yes </fatigue>
<lymph_nodes> normal </lymph_nodes>
<malaise> no </malaise>
<nose> runny </nose>
<temperature> 38-39 </temperature>
<throat> sore </throat>

Figure 21: The set of New SP-patterns supplied to the SPCM for the example discussed in the text. These patterns represent the patient 'John Smith' and his symptoms.



Figure 22: The best alignment found by THE SPCM with the set of patterns from Figure 21 in New (describing the symptoms of the patient 'John Smith') and a set of patterns in Old describing a range of different diseases and named clusters of symptoms, together with the 'framework' pattern shown in column 1. Reproduced from [20, Figure 6].

5.10 Nonmonotonic Reasoning and Reasoning with Default Values

Conventional deductive reasoning is *monotonic* because deductions made on the strength of current knowledge cannot be invalidated by new knowledge: the conclusion that 'Socrates is mortal', deduced from 'All humans are mortal' and 'Socrates is human', remains true for all time, regardless of anything we learn later.⁸ By contrast, the inference that 'Tweety can probably fly' from the propositions that 'Most birds fly' and 'Tweety is a bird' is *nonmonotonic* because it may be changed if, for example, we learn that Tweety is a penguin or an ostrich.

The elements of nonmonotonic reasoning are illustrated in the following SP-Multiple-Alignments.



Figure 23: The first of the three best SP-Multiple-Alignments formed by the SPCM with 'bird Tweety' in New and SP-patterns in Old as described in the text. The relative probability of this SP-Multiple-Alignment is calculated as 0.66. Reproduced from [12, Figure 17].

In Figure 23, a bird called 'Tweety' (columns 0 and 1) is identified as a bird (column 2) and, as such, it is assumed that it can (probably) fly (column 3). This inference is probabilistic because, as described below, there are two other SP-Multiple-Alignments that can be formed from the information that Tweety is a bird (columns 0 and 1). The SPTI calculates the relative probability of this SP-Multiple-Alignment as 0.66 (calculated from an imaginary frequency of occurrence assigned to each of the Old SP-patterns).

In Figure 24, a bird called 'Tweety' (columns 0 and 1) is identified as a bird (column 2), and as an ostrich (column 3). In this case, we know that Tweety, as an ostrich, would not be able to fly (column 3), but because this SP-Multiple-Alignment is only one of three alternative SP-Multiple-Alignments created from the same New information, the result is less than certain. The relative probability of this SP-Multiple-Alignment is calculated as 0.22.

⁸This section is based on [12, Section 10.1].



Figure 24: The second of the three best SP-Multiple-Alignments formed by the SPCM with 'bird Tweety' in New and patterns in Old as described in the text. The relative probability of this SP-Multiple-Alignment is calculated as 0.22. Reproduced from [12, Figure 18].



Figure 25: The last of the three best SP-Multiple-Alignments formed by the SPCM with 'bird Tweety' in New and SP-patterns in Old as described in the text. The relative probability of this SP-Multiple-Alignment is 0.12.

Figure 25, is much the same as Figure 24 except that, in this case, Tweety is a penguin and, as such, he (or she) would not be able to fly. The relative probability of this SP-Multiple-Alignment is calculated as 0.12.



Figure 26: The best SP-Multiple-Alignment formed by the SPCM with 'penguin Tweety' in New and SP-patterns in Old as described in the text. The relative probability of this SP-Multiple-Alignment is 1.0.

Figure 26 is different from the previous three SP-Multiple-Alignments because, in this case, column 0 tells us that Tweety is a penguin, not a bird. Now there is a sharp change in the probability calculated by the SPTI: the relative probability calculated by the SPCM is 1.0 because there are no alternatives SMAS created from that New information in column 0. The same would apply if column 0 was 'ostrich Tweety'.

These examples illustrate nonmonotonic reasoning as outlined at the beginning of this section because the inferences that are made about Tweety and his (or her) ability to fly can change, depending on the information about Tweety that is supplied. This is much more in keeping with way people normally reason than is classical logic and its procrustean rules that prevents inferences from changing as we learn more about Tweety.

5.11 Problem Solving

As noted in Section ??, Barlow foresaw that '... the operations needed to find a less redundant code have a rather fascinating similarity to the task of answering an intelligence test, ...' [6, p. 210]. In support of his prescient observation, this section shows how IC at the core of the SPCM may solve a modified version of the kind of puzzle that is popular in intelligence tests.

Figure 27 shows an example of this type of puzzle. The task is to complete the relationship 'A is to B as C is to ?' using one of the geometric patterns 'D', 'E', 'F' or 'G' in the position marked with '?' in the figure. For this example, the 'correct' answer is clearly 'E'. Quote marks have been used for the word 'correct' because, in some problems of this type, there may be two or more alternative answers where there is uncertainty about which answer is the right one.



Figure 27: A geometric analogy problem.

Normally, these tests use simple geometric patterns like those shown in Figure 27 but because the SPCM has not yet been developed to process two-dimensional SP-patterns, the geometric patterns are described textually as, for example, in 'small square inside large ellipse', 'small square inside large circle', and so on, as shown in Figure 28.

Computer-based methods for solving this kind of problem have existed for some time (e.g., Evans's [21] well-known heuristic algorithm). In more recent work [22, 23], AIT principles have been applied to good effect. The proposal here is that, within the general framework of Ockham's razor, this kind of problem may be understood in terms of the SP concepts.

Given that the diagrammatic form of the problem has been translated into textual patterns as described above, this kind of problem can be cast as a problem of partial matching, well within the scope of the SPCM.

Figure 29 shows the best SP-Multiple-Alignment created by the SPCM with New information in

```
C1 small square inside large ellipse ;
   D small square inside large circle #C1
C2 small square inside large ellipse ;
   E large square above small ellipse #C2
C3 small square inside large ellipse ;
   F small circle left-of large square #C3
C4 small square inside large ellipse ;
   G small ellipse above large rectangle #C4.
```

Figure 28: Textual patterns corresponding to the combination of 'C' on the bottom left of Figure 27 with one of 'D', 'E', 'F', or 'G' down the right side of the figure. These serve as Old SP-patterns as described in the text.

column 0 corresponding to the geometric patterns 'A' and 'B' in Figure 27, and the Old SP-patterns shown in Figure 28 corresponding to the geometric patterns 'D', 'E', 'F' and 'G' in Figure 27.

0 1 C2 small ---- small circle square inside --- inside large ---- large triangle ellipse ; ----- ; Е large ---- large circle square above ---- above small ---- small triangle ellipse #C2 0 1

Figure 29: The best SP-Multiple-Alignment found by the SPCM for the SP-patterns in New and Old as described in the text.

As can be seen from the figure, the best SP-Multiple-Alignment found by the SPCM shows, in column 2, the combination of textual patterns corresponding to a combination of geometric patterns 'C' and 'E' in Figure 27, and of course this is the 'correct' answer as noted above.

As can be seen from the figure, finding the best SP-Multiple-Alignment from these New and Old SP-patterns depends on the ability of the SPCM to find good partial matches between SP-patterns.

6 How the SPTI With Other Observations Suggest an Entirely Novel Perspective on the Foundations of Mathematics, Logic, and Computing

In view of evidence for the importance of IC in human intelligence [1, Section 2], and in view of the fact that mathematics is the product of human brains and has been designed as an aid to human thinking, it should not be surprising to find that IC is central in the structures and workings of mathematics [1, Section 8]. This idea is supported by evidence that well-known techniques for compression of information may be seen in the structure and workings of mathematics.

Similar arguments may be applied to the foundations of logic and computing [24, Section 7].

7 A Proposed New Mathematics

The idea that both intelligence and mathematics may be understood as IC has led to the proposal that mathematics may be integrated with a mature version of the SPTI to create a *New Mathematics* with many potential benefits and applications ([1, Section 8]).

7.1 The Potential of the SPTI as a Theory of Probability

The SPTI has potential as a theory of probability [1, Section 10]:

- Inference and probabilities via generalisation [1, Section 10.1].
- Inferences and probabilities via partial matching in the sp-multiple-alignment concept [1, Section 10.2].
- Exploiting the asymmetry between information compression and concepts of probability [1, Section 10.3]
- Towards a new science of probability [1, Section 10.4].

7.2 The Potential of the SPTI as a Theory of Computing

The SPTI has potential as a theory of computing [1, Section 11].

- Concepts of computing and The Post Canonical System [1, Section 11.1].
- Potential benefits of the SPTI as a model of computing [1, Section 11.2].

8 Some Other Features of the SPTI and Associated Ideas

This section provides pointers to parts of the book [1] which have been omitted from this short account.

8.1 Deriving an Economical Code from an SPMA and Recreating a New SP-Pattern From Its Code

How an economical code for a New SP-pattern may be derived from an SPMA for that New SP-pattern, and how the New SP-pattern may be recreated from such a code is described in [1, Section 5]. There is also relevant discussion in [1, Section 6.3.3].

8.2 SP-Neural

SP-Neural is a preliminary version of the SPTI in terms of neurons and their inter-connections and inter-communications, described quite fully in Section 7 of [1], and more fully in [25].

References

- [1] J. G. Wolff, Intelligence Via Compression of Information, 2024, in preparation.
- [2] —, "Twenty significant problems in AI research, with potential solutions via the SP Theory of Intelligence and its realisation in the SP Computer Model," *Foundations*, vol. 2, pp. 1045– 1079, 2022.
- [3] —, "Information compression as a unifying principle in human learning, perception, and cognition," *Complexity*, vol. 2019, p. 38, February 2019, article ID 1879746. viXra:1707.0161v3, hal-01624595 v2.
- [4] F. Attneave, "Some informational aspects of visual perception," *Psychological Review*, vol. 61, pp. 183–193, 1954.
- [5] H. B. Barlow, "Sensory mechanisms, the reduction of redundancy, and intelligence," in *The Mechanisation of Thought Processes*, HMSO, Ed. London: Her Majesty's Stationery Office, 1959, pp. 535–559.
- [6] —, "Trigger features, adaptation and economy of impulses," in Information Processes in the Nervous System, K. N. Leibovic, Ed. New York: Springer, 1969, pp. 209–230.
- [7] R. J. Solomonoff, "A formal theory of inductive inference. Parts I and II," Information and Control, vol. 7, pp. 1–22 and 224–254, 1964.
- [8] —, "The discovery of algorithmic probability," Journal of Computer and System Sciences, vol. 55, no. 1, pp. 73–88, 1997.
- [9] J. G. Wolff, "Application of the SP Theory of Intelligence to the understanding of natural vision and the development of computer vision," *SpringerPlus*, vol. 3, no. 1, pp. 552–570.
- [10] —, "Autonomous robots and the SP Theory of Intelligence," *IEEE Access*, vol. 2, pp. 1629–1651, 2014, arXiv:1409.8027 [cs.AI], bit.ly/18DxU5K.
- [11] —, Unifying Computing and Cognition: the SP Theory and Its Applications. Menai Bridge: CognitionResearch.org, 2006, distributors include Amazon.com and Amazon.co.uk. The print version is produced via print-on-demand from "INGRAM Lightning Source" and, via that technology, is unlikely to go out of print.

- [12] —, "The SP Theory of Intelligence: an overview," Information, vol. 4, no. 3, pp. 283–341, 2013, arXiv:1306.3888 [cs.AI], bit.ly/1NOMJ6l.
- [13] —, "The SP Theory of Intelligence, and its realisation in the SP Computer Model as a foundation for the development of artificial general intelligence," *Analytics*, vol. 2, no. 1, pp. 163–192, 20023, web: www.mdpi.com/2813-2203/2/1/10.
- [14] M. Bramer, Logic Programming With Prolog, second, Kindle ed. London: Springer-Verlag, 2013.
- [15] J. G. Wolff, "Towards an intelligent database system founded on the SP Theory of Computing and Cognition," *Data & Knowledge Engineering*, vol. 60, pp. 596–624, 2007, arXiv:cs/0311031 [cs.DB], bit.ly/1CUldR6.
- [16] N. Chomsky, Syntactic Structures. The Hague: Mouton, 1957.
- [17] F. C. N. Pereira and D. H. D. Warren, "Definite Clause Grammars for language analysis a survey of the formalism and a comparison with augmented transition networks," *Artificial Intelligence*, vol. 13, pp. 231–278, 1980.
- [18] G. Gazdar and C. Mellish, Natural Language Processing in Prolog. Wokingham: Addison-Wesley, 1989.
- [19] J. G. Wolff, "Cognitive development as optimization," in *Computational Models of Learning*, L. Bolc, Ed. Heidelberg: Springer-Verlag, 1987, pp. 161–205.
- [20] —, "Medical diagnosis as pattern recognition in a framework of information compression by multiple alignment, unification and search," *Decision Support Systems*, vol. 42, pp. 608–625, 2006, arXiv:1409.8053 [cs.AI], bit.ly/1F366o7.
- [21] T. G. Evans, "A program for the solution of a class of geometric-analogy intelligence-test questions," in *Semantic Information Processing*, M. L. Minsky, Ed. Cambridge Mass.: MIT Press, 1968, pp. 271–353.
- [22] T. Belloti and A. Gammerman, "Experiments in solving analogy problems using Minimal Length Encoding," in Proceedings of Stream 1: "Computational Learning and Probabilistic Reasoning", Applied Decision Technologies, Brunel University, April 1995, 1996.
- [23] A. J. Gammerman, "The representation and manipulation of the algorithmic probability measure for problem solving," Annals of Mathematics and Artificial Intelligence, vol. 4, pp. 281– 300, 1991.
- [24] J. G. Wolff, "Mathematics as information compression via the matching and unification of patterns," *Complexity*, vol. 2019, p. 25, 2019, article ID 6427493, Archives: vixra.org/abs/1912.0100 and hal.archives-ouvertes.fr/hal-02395680. This paper is reproduced in *New Ideas Concerning Science and Technology*, Vol. 13, 12 June 2021, Pages 132–169, doi.org/10.9734/bpi/nicst/v13/8449D.
- [25] —, "Information compression, multiple alignment, and the representation and processing of knowledge in the brain," *Frontiers in Psychology*, vol. 7, no. 1584, 2016, arXiv:1604.05535 [cs.AI], bit.ly/2esmYyt.