Proposal for the creation of a research facility for the development of the SP machine*

J Gerard Wolff†   Vasile Palade‡

July 5, 2016

Abstract

This is a proposal to create a research facility for the development of a high-parallel version of the *SP machine*, based on the *SP theory of intelligence*. We envisage that the new version of the SP machine will be an open-source software virtual machine, derived from the existing *SP computer model*, and hosted on an existing high-performance computer. It is intended as a means for researchers everywhere to explore what can be done with the system and to create new versions of it. The SP system is a unique attempt to simplify and integrate observations and concepts across artificial intelligence, mainstream computing, mathematics, and human perception and cognition, with information compression via the matching and unification of patterns as a unifying theme. Many potential benefits and applications flow from this simplification and integration. These include potential with problems associated with big data; potential to facilitate the development of autonomous robots; and potential in other areas including: unsupervised learning, natural language processing, several kinds of reasoning, fuzzy pattern recognition and recognition at multiple levels of abstraction, computer vision, best-match and semantic forms of information retrieval, software engineering, medical diagnosis, simplification of computing systems, and the seamless integration of diverse kinds of knowledge and diverse aspects of intelligence. Additional motivations for creating the proposed facility include

---

*An electronic copy of this document may be downloaded via bit.ly/1zZjjIs.

†Dr Gerry Wolff, BA (Cantab), PhD (Wales), CEng, MIEEE, 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.

‡Dr Vasile Palade, MEng (Tech. Univ. of Bucharest), PhD (University of Galaţi), IEEE Senior Member; Reader in Pervasive Computing, Department of Computing, Coventry University, UK; vasile.palade@coventry.ac.uk; +44 (0) 2477 659190; +44 (0) 7912 043982; *Skype*: v_palade; *Web*: bit.ly/1xNM5xr, bit.ly/1Buyk9E.

the potential of the SP system to help solve problems in defence, security, and the detection and prevention of crime; the potential impact of the development in terms of economic, social, environmental, and academic criteria, and in terms of publicity; and the potential for international influence in research. The main elements of the proposed facility are described, including support for the development of *SP-neural*, a neural version of the SP machine. The facility should be permanent in the sense that it should be available for the foreseeable future, and it should be designed to facilitate its use by researchers anywhere in the world.

1 Introduction

"There is nothing more practical than a good theory", Kurt Lewin [7, p. 169].

This is a proposal to create a high-parallel version of the *SP machine*, which we will refer to as "SPM", and a research facility for its further development, which we will refer to as "SPF". It is envisaged that SPM will be created as an open-source software virtual machine, derived from the existing *SP computer model* (SP71), and hosted on an existing high-performance computer. SPF is intended to be a means for researchers everywhere to explore what can be done with the SP system and to create new versions of it.

It would be helpful if any reader were to let the authors know (via jgw@cognitionresearch.org) if he or she may be interested in using such a facility or believes the facility would be useful.

1.1 The SP theory and the SP machine

The basis of the proposed development is the *SP theory of intelligence.* This is a unique attempt to simplify and integrate observations and concepts across artificial intelligence, mainstream computing, mathematics, and human perception and cognition, with information compression via *multiple alignment* as a unifying theme.[1]

Many potential benefits and applications flow from this simplification and integration, as detailed in Section 2 and Appendix C. Several of those benefits and applications may be realised on relatively short timescales, as indicated in the introduction to Appendix C.

---

[1]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".

The SP theory was not dreamed up overnight. It is the product of about 20 years of development and testing.

A central idea in the SP theory is the previously-mentioned concept of *multiple alignment*, borrowed and adapted from that concept in bioinformatics, and outlined in Appendix A.4.2. This concept, as it has been developed in the SP programme of research, has the potential to be as significant in computing and cognition as the double helix has proved to be in biological sciences.

Expressing the theory in a computer model has helped to reduce vagueness in the theory, it has provided a means of testing candidate ideas—many of which have been rejected as a result of this testing—and it is a means of demonstrating what can be done with the system.

There is an outline of the SP theory and the SP computer model in Appendix A, with pointers to where fuller information may be found. Distinctive features and apparent advantages of the SP system are summarised in Appendix B. Some of the potential benefits and applications of the SP system are summarised in Section 2 and Appendix C, with pointers to more comprehensive sources of information.

The wide scope of the SP system, and the wide range of its potential benefits and applications, means that there are many more avenues to be explored than could be tackled by any one research group. Hence the need for SPF. It is intended to facilitate large-scale collaboration, very much in the spirit of:

> "The creation of this new era of [cognitive] computing is a monumental endeavor ... no company can take on this challenge alone. So we look to our clients, university researchers, government policy makers, industry partners, and entrepreneurs—indeed the entire tech industry—to take this journey with us." [6, Preface].

and

> "... fund activities that support integration and collaboration within the research community; for instance, CPPs [Collaborative Computational Projects], consortia, networks, international interactions."[2]

1.2   Presentation

In what follows, we'll first describe some motivations for creating SPF (Section 2). Then we'll describe the facility itself and how it may be used (Section 3). There is an estimate of costs in Section 4, and concluding remarks in Section 5.

---

[2]From p. 9 in *EPSRC E-Infrastructure Roadmap*, August 2014, bit.ly/1d8Eh3I.

2 Motivations

In terms of policies of the UK Government and the UK Research Councils, SPM and SPF tick several boxes, and there are other reasons for creating SPF and developing SPM. These motivations are described in subsections below.

2.1 The power and potential of multiple alignment

Probably the most distinctive feature of the SP system is the powerful concept of multiple alignment, as it has been adapted from bioinformatics (Appendix B). In broad terms, this concept provides a key to:

- *Simplification of concepts and systems.* Diverse concepts in computing and cognition may be assimilated to the relatively simple concept of multiple alignment ([19], [26, Section2 and 4]), and that concept can mean substantial simplifications of computing systems ([19, Section 4.4.2], [26, Section 5]).

- *Versatility in intelligence.* The SP system has strengths in unsupervised learning, natural language processing, pattern recognition, information retrieval, several kinds of reasoning, planning, problem solving, information compression, and more [19, 21].

- *Adaptability in intelligence.* The SP system's strengths in unsupervised learning [19, Chapter 9] may provide a foundation for other kinds of learning [23, Section V]. Via unsupervised learning, and perhaps via other kinds of learning, the system has potential to promote human-like adaptability in intelligence.

- *Seamless integration of structures and functions.* The SP system, with multiple alignment centre stage, has potential to promote seamless integration of diverse kinds of knowledge and diverse aspects of intelligence [26, Section 7], an integration that appears to be essential for effective reasoning and problem solving in intelligent systems.

More specifically, there are many potential benefits and applications that flow from the multiple alignment concept, described in outline in Appendix C and in more detail in [19, 26, 24, 23, 20, 18]. As noted in Section 2.6, the potential value, worldwide, of those benefits and application, has been estimated to be at least \$190 billion, *every year*.

The estimated cost of setting up SPF is £634,700 and recurrent costs are estimated to be £173,800 pa (Section 4).

Taking account of these estimated costs, and the potential benefits and applications, it appears that the proposed facility would be a good investment.

2.2   UK industrial strategy

The development of SPF would chime well with the goal of "supporting emerging technologies", one of five main themes in the UK's industrial strategy.[3]

More specifically, the SPF fits well with the Government's "information economy strategy",[4] and its focus on "eight 'great technologies' which will propel the UK to future growth":[5]

The SP system has considerable potential to contribute to development of two of those "great technologies"—*big data* (discussed in the next subsection) and *robotics and autonomous systems* (discussed in Section 2.4). The SP system is also relevant to development in a third area—*energy*—discussed briefly in Section 2.5).

In addition, the very broad scope of the SP system (Section 2.1), and the provision of a facility for researchers anywhere in the world (Section 3.5), is likely, via synergies of various kinds, encourage "better science", "new kinds of science", and "interdisciplinary projects", in keeping with the Government's *A Strategic Vision for UK e-Infrastructure*.[6]

2.3   Big data

Outlined here are some significant challenges associated with big data and how the SP system may help to overcome them.

2.3.1   Challenges

There is a tendency amongst some people to write or speak as if big data was an unalloyed blessing and that all the necessary technologies are available to exploit it. But it is clear from the book *Smart Machines* by John Kelly and Steve Hamm (both of IBM) [6], and also from *Frontiers in Massive Data Analysis* from the US National Research Council [9], that much of the potential value of big data cannot be realised currently because of significant unsolved problems in that area.

It is also clear from *Smart Machines* that significant problems associated with big data cannot be solved by tinkering with existing designs for computers. Radical

---

[3]See *Industrial strategy: government and industry in partnership*, *GOV.UK*, 2013-08-06, bit.ly/1kYdtBS.

[4]See *Information Economy Strategy* (PDF), June 2013, bit.ly/1d5tn8h.

[5]See £600 *million investment in the eight great technologies*, Department for Business, Innovation & Skills, 2013-01-24, bit.ly/QFwhwi; and "The 'eight great technologies' which will propel the UK to future growth receive a funding boost", speech given by the Rt Hon David Willetts MP, former Minister for Universities and Science, *GOV.UK*, 2013-01-23, bit.ly/1wKCqII.

[6]Report from the Department for Business, Innovation and Skills, January 2012, bit.ly/1FrAQzW, p. 21.

rethinking of the architecture of computers will be needed, probably drawing on what we can learn about the workings of the human brain.

And achieving this transition to "cognitive computing" will require large-scale collaboration, as indicated in Section 1.1.

2.3.2 Potential solutions

SPM and SPF are relevant to the issues just described because: the SP system has been developed in the spirit of cognitive computing, drawing heavily on research on human perception and cognition, and neuroscience; it offers radical solutions to problems posed by big data; and the existence of SPF would encourage and facilitate large-scale collaboration amongst researchers all around the world.

The paper *Big data and the SP theory of intelligence* [24] describes how the SP system may help to solve nine of the problems associated with big data, as outlined in Appendix C.13.1. Three of those problems, and their potential solutions, are highlighted here:

- *The problem of variety in big data.* There is a pressing need in computing to tame the great variety of formalisms and formats for the representation of knowledge, each with their own mode of processing. The multiple alignment framework in the SP system has clear potential as a *universal framework for the representation and processing of diverse kinds of knowledge* (UFK) [24, Section III].

- *Energy consumption.* The SP system, in conjunction with the concept of "data-centric computing" [6, Chapter 5], has potential to "make computers many orders of magnitude more energy efficient" [6, p. 88] via two key principles [24, Section IX]: 1) Taking advantage of statistical information that the system gathers as a by-product of how it works; 2) Cutting out much searching by making direct connections between "neural symbols" in "SP-neural" (Appendix A.7).

- *Transmission of information.* The SP system has potential to yield dramatic economies in the transmission of information, partly by making big data smaller [24, Section VII], but perhaps more importantly via analysis/synthesis [11, Chapter 18], based on the judicious separation of *encoding* and *grammar* [24, Section VIII].

As noted in Appendix C.13.1, it appears that, considering the nine proposed solutions collectively, and in several cases individually, there are no alternatives that can rival the potential of what is described in [24].

With regard to the Government's information economy strategy,[7] "The UK

---

[7] *Information Economy Strategy* (PDF), June 2013, bit.ly/1d5tn8h.

now has the opportunity to take a lead in the global efforts to deal with the volume, velocity and variety of data created each day. This will require continued infrastructure investment ..." (p. 7). SPF has potential to be a major plank in that infrastructure, and at modest cost (Section 4).

2.4 Robotics and autonomous systems

As mentioned above, *robotics and autonomous systems* is another of the Government's "eight great technologies" where the SP system may make a contribution—perhaps via the UK Robotics and Autonomous Systems Network (UK-RAS Network)[8] or otherwise—and the potential is considerable.

In the paper *Autonomous robots and the SP theory of intelligence* [23], robots are "autonomous" if they do not depend on external intelligence or power supplies, and are mobile. It is also assumed that they are designed to exhibit as much human-like intelligence as possible. The paper describes how the SP system may contribute in three main areas:

- *Computational efficiency, the use of energy, and the size and weight of computers.* "To field a conventional computer with [human-like] cognitive capacity would require gigawatts of electricity and a machine the size of a football field." [6, p. 75]. If a robot is to be autonomous as described above, it needs a 'brain' that is efficient enough to do all the necessary processing without external assistance, does not require an industrial-scale power station to meet its energy demands, and is small enough and light enough to be carried around—things that are difficult or impossible to achieve with current technologies.

  Section III of [23] describes how these things may be achieved in a revised and updated version of arguments in [24, Section IX]. The SP system may help: by reducing the size of data to be processed; by exploiting statistical information that the system gathers as an integral part of how it works; and via a revised version of Donald Hebb's [5] concept of a "cell assembly".

- *Towards human-like versatility in intelligence.* If a robot is to operate successfully in an environment where people cannot help, or where such opportunities are limited, it needs as much as possible of the versatility in intelligence that people may otherwise provide.

  The SP system demonstrates versatility via its strengths in areas such as unsupervised learning, natural language processing, fuzzy pattern recognition

---

[8]See, for example, "Widespread backing for UK robotics network", The Engineer, 2015-06-24, bit.ly/1BR9qE2.

and recognition at multiple levels of abstraction, best-match and semantic forms of information retrieval, several kinds of reasoning, planning, problem solving, information compression, and more [23, Section IV].

But the SP system is not simply a kludge of different AI functions. Owing to its focus on simplification and integration of concepts in computing and cognition (Appendix A), it promises to reduce or eliminate unnecessary complexity and to avoid awkward incompatibilities between poorly-integrated subsystems, as indicated in Appendix C.1.

- *Towards human-like adaptability in intelligence.* The SP system's strengths in unsupervised learning and other aspects of intelligence may help to achieve human-like adaptability in intelligence via: one-trial learning; the learning of natural language; learning to see; building 3D models of objects and of a robot's surroundings; learning regularities in the workings of a robot and in the robot's environment; exploration and play; learning major skills; and learning via demonstration [23, Section V].

This new approach to the development of autonomous robots has several distinctive features (Appendix B), especially the concept of multiple alignment. It appears that this approach has considerable potential to promote computational and energy efficiency in a robot's brain, and to achieve human-like versatility and adaptability in intelligence.

2.5 Energy

The SP system may also contribute indirectly to another of the areas identified as one of the "eight great technologies"—*energy*—via increases in efficiency in the workings of computers and in the transmission of data, as outlined in Sections 2.3 and 2.4, and described more fully in [24, Sections VIII and IX] and also in [23, Section III].

2.6 Other potential benefits and applications of the SP system

In addition to big data (Section 2.3), robotics and autonomous systems (Section 2.4), and energy (Section 2.5), there are several other potential benefits and applications of the SP system, outlined in Appendix C. As noted there, several of these potential benefits and applications may be realised on relatively short timescales with existing high-performance computers or even ordinary computers.

Peer-reviewed papers about potential benefits and applications that have not already been mentioned include [22] (application of the SP theory to the understanding of natural vision and the development of computer vision), [18] (application of the SP system to medical diagnosis), and [20] (the SP system as an intelligent database), and [26] (several other potential benefits and applications including unsupervised learning, natural language processing, pattern recognition, computer vision, software engineering, information compression, the semantic web, bioinfomatics, data fusion, simplification of computing systems, and the seamless integration of structures and functions in diverse kinds of knowledge).

In view of the wide scope of the SP system, and evidence of its potential, it seems reasonable to estimate that it could add at least 5% to the value of IT investments, worldwide. Since these are about \$3.8 trillion annually,[9] the value of the SP concepts, *every year*, would be at least \$190 billion! [26, Section 8].

2.7 Defence, security, and the detection and prevention of crime

With regard to "Security in a changing world", part of the "vision" [13, p. 13] of the UK's Science & Technology Facilities Council (STFC), several of the strengths of the SP system that have been mentioned are relevant to defence, security, and the detection and prevention of crime. These include:

- Big data (Section 2.3), and data fusion [26, Section 6.10.5].

- Robotics and autonomous systems (Section 2.4).

- Natural language processing [19, Chapter 5], [26, Section 6.2].

- Pattern recognition (see Appendix C.5 and [19, Chapter 6]).

- The SP system may function as an intelligent database [20].

- Computer vision [22].

- Several kinds of reasoning, including one-step 'deductive' reasoning, abductive reasoning, reasoning with probabilistic networks and trees, reasoning with if-then rules, nonmonotonic reasoning, 'explaining away', and causal diagnosis [19, Chapter 7].

- Planning and problem solving [19, Chapter 8].

---

[9]See, for example, "Gartner: Big data will help drive IT spending to \$3.8 trillion in 2014", *InfoWorld*, 2013-01-03, bit.ly/Z00SBr.

- Unsupervised learning [19, Chapter 9], [26, Section 6.1].

- The potential of the system to simplify computing system [26, Sections 4 and 5].

- The potential of the system to promote seamless integration of diverse forms of knowledge and diverse aspects of intelligence [26, Section 7].

This last point is particularly relevant to forensic work and security work, because of the need to remove artificial barriers between different aspects of intelligence, different kinds of knowledge, and different kinds of processing, so that all kinds of intelligence and knowledge can be brought to bear in solving crimes and in the prevention of acts of terrorism. Similar things may be said about autonomous robots that aspire to human-like versatility and adaptability in intelligence.

2.8 Contributing to the work of the planned Alan Turing Institute for Data Science and the planned Cognitive Computing Research Centre

"The Alan Turing Institute for Data Science will benefit from a £42 million government investment over 5 years that will strengthen the UK's aims to be a world leader in the analysis and application of big data. It will also ensure that the UK is at the forefront of data-science in a rapidly moving, globally competitive area, enabling first-class research in an environment that brings together theory and practical application."[10]

The SP system and SPF have potential to contribute to the work of the planned *Alan Turing Institute for Data Science*, and also the planned *Cognitive Computing Research Centre*,[11] by facilitating both the analysis and application of big data (Section 2.3).

2.9 Impact

"The Higher Education Funding Council for England (HEFCE), Research Councils UK (RCUK) and Universities UK (UUK) have a shared commitment to support and promote a dynamic and internationally

---

[10]See "Plans for world class research centre in the UK", *GOV.UK*, 2014-03-18, bit.ly/1j6CZYB.

[11]See "Autumn Statement 2014", *H M Treasury*, December 2014, p. 50, PDF, bit.ly/15OO6QS.

competitive research and innovation base that makes an increased and sustainable contribution, both nationally and globally, to economic growth, wellbeing, and the expansion and dissemination of knowledge.", "Impact Policies", *Research Councils UK*, bit.ly/1znJpts.

On five main fronts, the potential impact of SPF is large:

- *Economic.* As noted in Section 2.6, a conservative estimate of the potential value of the SP concepts is at least $190 billion, every year.

- *Social.* As already noted, there are many potential benefits and applications of the SP system, some of them outlined in Sections 2.3 to 2.6, and in Appendix C.

- *Environmental.* The SP concepts have potential to deliver more for less— with consequent environmental benefits—via a combination of: 1) gains in economic and social impact (as just noted); 2) reductions in energy consumption and in the size and weight of computers (Sections 2.3.2 and 2.4).[12]

- *Academic.* We expect the establishment of SPF, with associated publicity, to encourage and facilitate contributions by research groups and individual researchers all around the world. Given the wide scope of the SP system and the wide range of its potential benefits and applications, the academic impact of SPF is potentially very large. Research that is done with SPF may be seen as a 'dividend' from the investment. From the perspective of the UK, research that is done by research groups and individual researchers outside the UK may be seen as a bonus, additional to what would be achieved if the facility were restricted to UK researchers. In addition, there are likely to be benefits arising from synergies and large-scale collaboration (Section 2.2).

  Initially, we aim to raise awareness of the facility amongst researchers and to encourage relevant research projects. But later, we expect that research with the facility will develop its own momentum.

- *Publicity.* As a condition for using SPF, researchers would be asked to ensure that every relevant publication in an academic journal, collection of academic papers, or conference proceedings, contains an acknowledgement of the facility and of the body or bodies that have provided the necessary funds. This in itself would be publicity. But, in addition, there is potential

---

[12]There would also be a need for measures to counteract the effects of "Jevon's paradox" (*Wikipedia*, bit.ly/1dUJF9C, retrieved 2014-12-22)—how gains in efficiency may, without appropriate controls, lead to an overall increase in consumption.

for publicity in articles in newspapers and magazines, in programmes on radio or TV, and on the internet, in blogs, mailing lists, Facebook pages, and the like.

2.10 International influence

"Our priorities for maximising international impact are: to seek increased influence for the UK in research, especially in Europe; in the long term to work towards the siting of a major international research facility in the UK.", [13, p. 21].

Given the wide scope of the SP system, the international dimension of SPF, and its potential impact (Section 2.9), it has potential to be "a major international research facility" and to increase the international influence of STFC or other hosting organisation.

3 The proposed facility

As mentioned in the Introduction, it is intended that SPF will facilitate the development of a high-parallel version of the SP machine, realised as an open-source software virtual machine, and hosted on an existing high-performance computer. How things may develop is shown schematically in Figure 1.

Aspects of the proposal are described in the following subsections.

3.1 The proposed facility and public policies

Some aspects of SPF may seem unfamiliar and may, superficially, appear to conflict with funding policies. These remarks apply, in particular, to the proposal that SPF should be hosted on its own dedicated machine (Section 3.3), that it should be permanent (Section 3.4), that it should be available to researchers anywhere in the world (Section 3.5), and that, with few if any exceptions, there should be no charges for using the facility (Section 3.6).

But it appears that, at high levels, it is recognised that funding for e-infrastructure should be responsive and flexible, that it should recognise the diverse needs of the research community, and that there is a need for long-term perspectives. In support of such policies:

- "In the roadmap, EPSRC aims to ... understand the requirements of the EPS [engineering and physical sciences] research community that make use of e-infrastructure, ensuring there are no gaps or duplication." (From p. 3
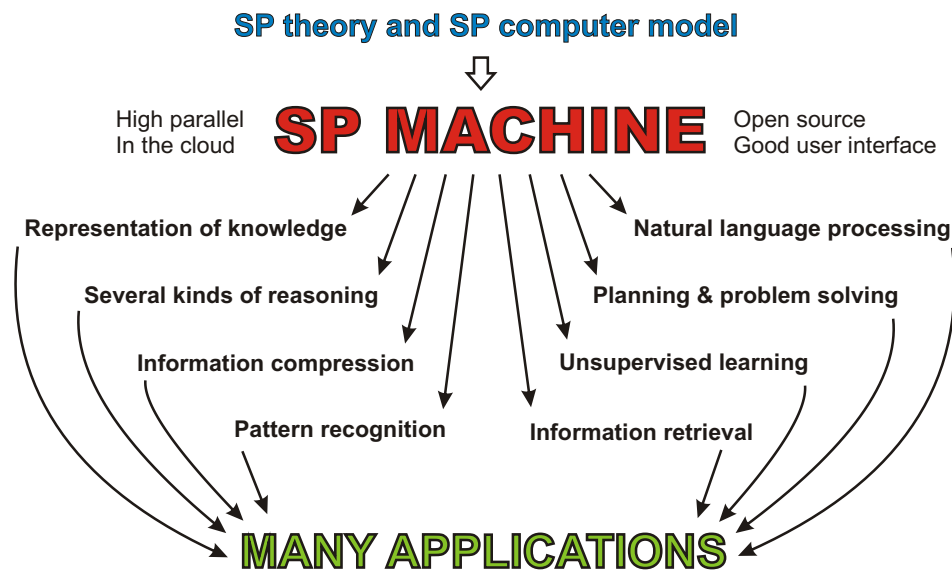
Figure 1: Schematic representation of the development and application of the SP machine.

of "EPSRC e-infrastructure roadmap", Engineering and Physical Sciences Research Council, August 2014, bit.ly/1HvM34x).

- "The funding agencies should recognise the need for flexibility of funding and put in place mechanisms to enable rapid responses to hardware and software developments and to new research opportunities and requirements." (Recommendation #3 from p. 17 in "Strategy for the UK Research Computing Ecosystem", The e-Science Institute at The University of Edinburgh, October 2011, bit.ly/1KslhrQ).

3.2   Elements of SPF

It is envisaged that the main features of SPF will be these:

- *The SP machine, version 1.* SPF will contain the first high-parallel version of the SP machine, which may be referred to as the "SP machine, version 1" or "SPM1".[13] The name "SPM" would normally mean SPM1 but, depending on the context, it may be used to refer to any other version of the SP machine or to all versions, collectively.

---

[13]Although the SP computer model may be seen as a version of the SP machine (Appendix A.2), we think it is probably best if the first SPF version is called "SPM1".

- *Founded on the SP71 computer model.* The SP71 computer model (Appendix A.2), the basis for SPM, will be ported on to the host machine. This should be relatively straightforward to do since the model is written in C++, with good comments, and since a slightly earlier version, SP70, is described quite fully, with pseudocode, in [19, Sections 3.9, 3.10, and 9.2].

- *High levels of parallel processing.* As far as possible, processes in the SP71 code will be modified to take advantage of parallel processing. This may be done in three main parts of the system, in procedures for: the matching and unification of patterns (as outlined in Appendix A.4.1); the building of multiple alignments (Appendix A.4.2); and the unsupervised learning of grammars (Appendix A.4.3).

- *Good user interface.* SPF will be accessed via a well-designed website, with user-friendly interfaces for each part of the system, using multimedia where appropriate. It would be best if users did not have to download and install any software. Any such downloading of software can probably be avoided by developing the user interface using the facilities of HTML5. In accordance with Google's new search policies, announced in April 2015,[14] the website should be "mobile friendly", meaning that it should be easy to use and legible on mobile phones.

- *Registration of users.* There will be procedures for the registration of each user, assigning disk space, and so on.[15] Here, a "user" will be a research group or an individual researcher. To avoid undue complications, at least initially, each research group will be treated as if it was an individual researcher, with no attempt to differentiate individual researchers within each group, or their work. Each research group would need to make its own arrangements for the division of labour within the group.

- *Creating new versions of the SP machine.* Any user may create one or more new versions of SPM:

  - Each such version will have its own storage space with documentation, source files, executable files, data files, and so on.
  - There will an hierarchical scheme of the usual kind for the identifiers of different versions, with identifier(s) for the creator(s) of each version.

---

[14]See, for example, "Google search changes will promote mobile-friendly sites", *BBC News*, 2015-04-20, bbc.in/1Q9AORA.

[15]As noted in Section 2.9, each user would be asked to ensure that, in all academic publications arising from research with SPF, there would be an acknowledgement of the facility and of the body or bodies that have provided the necessary funds.

- The documentation should make clear the origins of each version in one or more previous versions, which may include SPM1 or other versions by the same user or other users.

- Any version of SPM will be in one of two states: "under development" or "completed". Once a version has been "completed", further modifications will not be allowed. The executable code of any completed version may be run by any other user (next).

- *Running the SP machine.* Any version of SPM that is under development may be run by the user that is developing it. Any completed version of SPM may be run by any user, which would of course include the user that created it. For any user to run any given version of SPM, facilities will be needed for:

  - Assigning an area of storage to hold data and results.

  - Setting parameters for the given version of SPM.

  - Creating and editing SP patterns.

  - Uploading SP patterns that have been created elsewhere.

  - Online debugging of software (for SPM versions that are under development). It is likely that a suitable debugger is already established on the host machine.

  - Viewing results, including multiple alignments and grammars created by the program. Since some results—multiple alignments in particular—may be larger than can be seen comfortably on a typical computer screen, there should be facilities for zooming and scrolling.

  - Downloading SP patterns, included those newly-created by the given version of SPM.

- *SP-neural, version 1.* Alongside SPM1 will be a first, tentative, version of the SP-neural machine, which may be referred to as "SPNRL1". The main elements of the SP-neural facility will be:

  - As with SPM, each user may create one or more versions of SPNRL.

  - SPNRL1 will have a means of translating a set of SP patterns (or one or more multiple alignments) into an inter-connected set of pattern assemblies. With the object-oriented facilities of C++, there will be a class for neurons and a class for pattern assemblies.

15

- In this preliminary model, each SP symbol will be represented by a single neuron although it is likely that in living brains, neural "symbols" are realised with clusters of inter-connected neurons.

- Connections between neurons will be represented with C++ pointers. As SPNRL is developed, it is likely that there will be excitatory connections and also inhibitory connections, in accordance with what is known about neural tissue. It is also likely that there will be lateral connections between neighbouring neurons within each pattern assembly.

- There will be a means of displaying, in graphical form, a set of inter-connected neurons and pattern assemblies, much as in Figure 5.

- For any given set of inter-connected neurons and pattern assemblies, there will be a means of downloading it as a set of SP patterns.

- As SPNRL is developed:

  * There will be a means of sending excitatory and inhibitory signals through inter-connected pattern assemblies and displaying the results in terms of the excitatory levels of individual neurons.

  * There will a means of creating new pattern assemblies and learning neural 'grammars', in accordance with learning processes in any given version of SPM.

  * Parallel processing will be applied to mimic the high levels of parallelism that exist in the workings of real neural tissue.

3.3 The facility probably requires a dedicated machine

For the following main reasons, the proposed facility probably needs to be set up on its own dedicated machine:

- Most high-parallel machines are geared mainly to the running of compute-intensive programs and make few if any concessions to what is required for the development of software—unrestricted access to run programs at any time, and the use of interactive debugging software.

- The need for permanence of research facilities (Section 3.4) is not widely recognised. It is true that all hardware needs to be replaced periodically, but existing high-parallel facilities are quite likely to have replacement policies that do not meet the needs of the proposed research—giving researchers confidence that the research machine, or a machine that will run precisely the same software, will be available for the foreseeable future.

- Most high-parallel machines have restrictions on who is eligible to use them. They are not, by default, accessible to researchers anywhere in the world (Section 3.5).

- The procedures for gaining access to existing high-parallel machines can be unnecessarily complicated. The procedures need to be as simple as possible and, for most uses of the machine, there should be no charge (Section 3.6).

In the early stages, the dedicated machine need not be very large or expensive. It may be set up initially as a cluster of processors and then expanded later. Of course, a key requirement would be that the machine can be expanded in that way. Alternatively, it should be possible to add new machines, with the proviso that two or more machines can work together seamlessly as if they were one.

Other key requirements would be the provision of a C++ compiler with an interactive debugger.

3.4   The facility should be permanent

For reasons that follow, SPF should be permanent in the sense that it should be available for the foreseeable future:

- *Bridging the 'Valley of Death'.* In accordance with "It is our historic failure to back [R&D, technology and engineering] which lies behind the familiar problems of the so-called 'valley of death' between scientific discoveries and commercial applications",[16] the facility should support and encourage research that has sufficient breadth and depth to bridge that 'valley of death'. Short-termism is just as damaging for research as it is for business.

- *Long-term research programmes*:

  - *Researchers need confidence that SPF will not be withdrawn.* We envisage the proposed facility providing the basis for research that is likely to include long-term research programmes. Any research group that wishes to embark on that kind of research programme needs to have confidence that the research facility will not be withdrawn. *Without assurance for researchers that the facility is permanent, there is a high risk that the entire project will fail.*

  - *BIS: software as a sustained infrastructure in the long term.* "Current investments in software development should be reviewed with a view to

---

[16]Speech by the Rt Hon David Willetts MP, "The 'eight great technologies' which will propel the UK to future growth receive a funding boost", 2013-01-24, bit.ly/1wKCqII.

developing models of support for software as a sustained infrastructure in the long term, as opposed to being supported by significant one off investments."[17]

- *e-Science Institute: the need for long-term funding for ambitious software development projects.* "There needs to be long-term funding for ambitious software development projects .... The current approach is hardware centric and short-term ..., but does not recognise the critical and long-term role that software and people play in the [research computing] enterprise."[18]

- *Time to develop research programmes.* It is likely to take time for a research group to learn about the SP system, to decide to mount a research project, to prepare a research proposal, for the proposal to be assessed, to provide PCs and other facilities, and to recruit staff. If the proposed research facility is only available for a few years (bearing in mind the time needed to create the facility), much of that time may be taken up with the processes described. Then it is likely that there would be little or no time left to do the actual research.

- *Consistency across facilities.* In the same way that large and expensive facilities—large telescopes, atom smashers, supercomputers, and the like— are maintained for many years, and replaced as necessary, the same should apply to less expensive or smaller-scale facilities. In both cases, research benefits from the long-term perspective that stability makes possible.

- *Archival status.* As outlined in Section 3.2, the facility will be designed so that research groups and individual researchers may create their own versions of the SP machine. Each such version needs to have archival status:

  - *The need for long-term access to and execution of software.* Other researchers, now and in the future, need to be able to access the source code and documentation for any version of the SP machine and to run the executable code. Naturally, the host machine for SPF may need to be replaced from time to time. In any such case, SPF, including all associated projects and data, may be ported on to the replacement machine, in accordance with standard practice in industry and commerce.

---

[17]From p. 8 in "Report of the e-Infrastructure Advisory Group", Department for Business, Innovation and Skills, June 2011, bit.ly/1LMXOSD.

[18]Recommendations #4 and #5, p. 17, in "Strategy for the UK Research Computing Ecosystem", The e-Science Institute at The University of Edinburgh, October 2011, bit.ly/1KslhrQ.

– *RCUK: software as a vital research infrastructure.* The need for long-term preservation of software and associated computing systems is implied by: "Software developed for experimental facilities and instrumentation, modelling and simulation and data-analysis is a critical and valuable resource. Software and algorithm development represents major investments by skilled researchers, and the large suite of codes and algorithms used in research should be regarded as a vital research infrastructure, requiring support and maintenance along the innovation chain, and throughout its life cycle. The reproducibility of research is at the very heart of the scientific method."[19].

– *Long-term preservation of software.* This requirement accords with the trend for academic journals to allow the author or authors of any published paper to provide supplementary material, which may include data, software, and things that may be regarded as forms of software, such as MP4 files.[20] It is increasingly recognised that many bodies of software should have the same archival status as peer-reviewed papers in academic journals or conference proceedings, and bodies of data on which those papers are based.

– *Preserving the execution environment.* Since software will often not run, or will not run correctly, unless it is running on the type of computer (with its operating system and related software) for which it was written, there is also a need for long-term preservation of each type of computer. There is more detail here:

  ∗ *Vint Cerf: digital "dark age" and "digital vellum".* The case has been argued by internet pioneer Vint Cerf, as reported in, for example: "Google's Vint Cerf warns of 'digital Dark Age'"[21]. Vint Cert also makes the case in a lecture that he gave on 2015-02-11, "Digital vellum and the expansion of the internet into the solar system" (see bit.ly/1aqDeKw);

  ∗ *Olive Executable Archive.* How software and associated computing systems may be preserved is demonstrated in the "Olive Executable Archive", www.olivearchive.org. This project has been highlighted by Vint Cerf in his articles and lectures.

---

[19]From p. 8 of "E-infrastructure roadmap", RCUK e-Infrastructure Group, 2014, bit.ly/1Jx59EE.

[20]The journal *Royal Society Open Science* says "Supplementary material can be used for supporting data sets, supporting movies, figures and tables, and any other supporting material." ("Instructions for Authors", bit.ly/1wD8ezN); and *IEEE Access* allows authors to include "data collections and multimedia materials" ("IEEE Access—Information for Authors", PDF, bit.ly/1wD5H8x).

[21]BBC News, 2015-02-13, bbc.in/1D3pemp

* *New gold standard established for open and reproducible research.* The case has also been argued by computer scientists at Cambridge University who "have set a new gold standard for openness and reproducibility in research by sharing the more than 200GB of data and 20,000 lines of code behind their latest results—an unprecedented degree of openness in a peer-reviewed publication."[22]

* *Maintaining access to digitised information.* The need to maintain access to digitised information in the face of changes in the hardware or software environment that it needs is discussed in "Preserving the digital record of computing history".[23]

As noted above, replacement of SPF's host machine should not be a problem. The system may be ported onto any new machine, as or when that becomes necessary.

- *Avoiding waste and gaining full value from SPF.* Given that the bulk of the cost of SPF is in its initial development and setting up, there is little scope for saving money by terminating the facility within a few years. Indeed, any such move is likely to mean a waste of resources and a failure to gain full value from the facility:

  - *The damaging effect of short-termism.* As noted above, it is likely that the entire project will fail unless researchers have confidence that the facility is permanent.

  - *Loss of valuable research.* Premature termination of SPF would mean losses of valuable research that may otherwise accrue, much of it provided at no cost to the UK by researchers elsewhere (Sections 2.9 and 3.6).

  - *SPF will not wear out and its value will increase.* Unlike ordinary research facilities such as space telescopes or atom smashers, SPF will not wear out. And its value for researchers will increase progressively as new versions of SPM are created and made available for the whole research community, including all new insights and improvements.

In case a commitment to permanence of SPF seems unduly expensive, it should be born in mind: 1) That the proposed facility will be extraordinarily cheap compared with things like the Hubble space telescope or the Large Hadron Colliger;

---

[22] "New gold standard established for open and reproducible research", Cambridge University press release, 2015-05-04, bit.ly/1IgFkKG.

[23] Article by David Anderson in *Communications of the ACM*, vol. 58, no. 7, July 2015, pp. 29–31.

and 2) The cost of maintaining the system for the foreseeable future is likely to be very much less than the value of research that is undertaken with the system.

Of course, if people stop using the facility, it would make sense to save on development and maintenance costs. But the software and its execution environment should be preserved in accordance with Vint Cerf's arguments outlined above. And it would be useful if people could still access the software and run it.

3.5 The facility should be available to researchers anywhere in the world

As previously noted, the wide scope of the SP system means that there are far more avenues to be explored than any one research group could tackle on its own. Accordingly, we are are aiming to create a facility for research groups and individual researchers anywhere in the world. It should not be restricted to UK scientists:

- This accords with what John Kelly and Steve Hamm say about the need for large-scale collaboration (quoted in Section 1.1) and also with the need for integration of e-infrastructures "internationally, across other national e-infrastructures, to deliver end-to-end services in the global environment of collaborative research"[24] and the need to facilitate "... research collaboration between industry and academia."[25]

- It also accords with the principles that:

  - The EPSRC aims to "Understand the whole UK e-infrastructure landscape, view it holistically and *consider it within an international context* (emphasis added)."[26]

  - "The UKs Research and Innovation e-infrastructure needs to be led and driven to deliver a UK wide vision for research e-infrastructure, *embedded in the international context essential to todays research challenges.*" (emphasis added).[27]

- Opening up the research in this way is likely to yield synergies that might not arise if there were more restrictions (Section 2.2)—with corresponding benefits for UK researchers.

---

[24]p. 2 in *E-infrastructure roadmap*, Research Councils UK, bit.ly/1Jx59EE.

[25]*ibid.*, p. 5.

[26]The EPSRC "E-Infrastructure Roadmap", retrieved 2015-07-01, bit.ly/1R3V767.

[27]"Report of the e-Infrastructure Advisory Group", Department of Business, Innovation and Skills, June 2011, p. 3, bit.ly/1LMXOSD.

- The international perspective will help to achieve full value from the facility (Section 2.1).

- And it will help to increase the impact of the facility (Section 2.9) and its international influence (Section 2.10).

It may seem perverse for UK research funding to be providing a research facility for researchers all around the world. But instead of viewing it as an unreasonably generous form of charity, it would be better to view it as a way of getting valuable research done for nothing.

There are also security reasons for opening up the research to a large community of researchers: "It is well known among cryptography and computer security experts that closed-source software and algorithms are less secure than their free and open-source counterpart."[28]

The reason that science has been and still is a highly international enterprise is because of the valuable synergies that arise from the sharing of ideas and results. The cost of the proposed facility is likely to be much smaller than the value of opening up the research to researchers anywhere in the world—meaning a net gain in value for the UK.

It is not without interest that this approach to the conduct of research—providing a computational facility to enable researchers anywhere to investigate a given area—is now being promoted by at least the following organisations:

- *The Human Brain Project.* Starting 30 March 2016, the scientific community worldwide can begin exploring the initial versions of the six ICT Platforms of the Human Brain Project.[29]

- *OpenAI.* The OpenAI research organisation is now promoting the "OpenAI Gym" (bit.ly/1Y3cVhc), a platform for artificial intelligence research, including "a site and API allowing people to meaningfully compare performance of their trained agents".[30]

- *Google.* Google has announced:

  - That "TensorFlow", its open-source platform for machine learning, will be accessible by anyone with an internet connection.[31]

---

[28]Quoted in "How to create a malevolent artificial intelligence". *MIT Technology Review*, 2016-05-19, bit.ly/1OZnSjS.

[29]See "Platforms overview" of the Human Brain Project, bit.ly/1Uj4LxJ.

[30]See "Elon Musks artificial intelligence group opens a 'gym' to train AI", *Popular Science*, 2016-04-27), bit.ly/1VH6fYm.

[31]See "How Google aims to dominate AI", *Popular Science*, 2016-04-29) bit.ly/1klAiF2.

– That some of its software for the processing of natural language will be be made freely available to researchers and others.[32]

- *IBM.* IBM has announced that it is making its 5-qubit quantum processor available to the public via its cloud computing facility.[33]

- *VisionLabs.* A global open-source computer vision project has been set up by VisionLabs.[34]

3.6 The facility should be easy to access and easy to use

In keeping with the main points in the preceding subsection, it is important to minimise obstacles to the use of the system for preliminary investigations, for teaching students, for individual research projects, and for long-term research programmes. An additional reason is that, for any research group or individual researcher that is not already familiar with the SP concepts, a significant amount of time and effort is likely to be required to get up to speed, so it is important to minimise other obstacles to progress.

For these reasons:

- *Ease of access.* The process of registering with the system and gaining access to it should be as simple and straightforward as possible.

- *No restrictions on use.* In the development of software, it is essential that researchers should have entirely unrestricted access to the host machine so that programs under development may be run very frequently, often with the use of an interactive debugger. Anything less than this is likely to be more costly in terms of researchers' time than any possible saving in terms of computer time. Restrictions on use may put the entire project at risk of failure. In short, there needs to be a recognition that the host machine is to be used for the creation and development of software, not merely the running of established programs.

- *User friendly.* For all aspects and parts of the system, there should be a user-friendly interface.

---

[32]See "Google's artificial-intelligence tool is offered for free", *The Wall Street Journal*, 2016-05-12, on.wsj.com/1qgYAlJ.

[33]See "IBM makes quantum computing available in the cloud", *Computerworld*, 2016-05-04, bit.ly/1TtAlZd.

[34]See "Developer teams up with Facebook and Google to make 'machines see'", *TechWorm*, , 2016-06-05, bit.ly/1ZK1XOv.

- *Free at the point of use.* With few if any exceptions, the system should be free for users. In addition to minimising obstacles for researchers, reasons include:

  - *Research dividends.* In keeping with remarks about research 'dividend' in Section 2.9, the overall value of research done with SPF is likely to outweigh the cost of providing and maintaining the facility, and by a considerable margin.

  - *Research bonuses.* As noted in Section 2.9, research that is done by research groups and individual researchers outside the UK may be seen as a bonus, additional to what would be achieved if the facility were restricted to UK researchers.

  - *Research synergies.* Also, as noted in Section 2.2, there are likely to be benefits from synergies and large-scale collaboration.

  - The 'Valley of Death'. In accordance with remarks in Section 3.4 about the 'valley of death', it is important to avoid any premature requirement that the research should, commercially, "stand on its own two feet" or "become self-sustaining", in much the same way that we do not expect young children to go out and earn a living.

- *Modern development environment.* To facilitate the process of development, a modern development environment would be helpful, such as Microsoft's Visual C++, part of Microsoft's Visual Studio.

With regard to costs, these may be met as follows:

- *Development and maintenance costs.* These costs may be met from a variety of sources, perhaps including funds for the support of researchers based in the UK.

- *Running costs.* Although it is likely that mature versions of SPM will be used for computationally-demanding tasks, we believe that much research with SPF can be done with small examples that make small computational demands on the host computer, and that the associated costs would be relatively small. It seems likely that the cost of administering a system of charges for running costs would outweigh the actual costs. In that case, it would be best, probably, not to attempt to make any charges for the use of the system. A possible exception is where a user wishes to run the system with computational tasks that are likely to consume large computational resources.

3.7 Open source

As already mentioned, SPM should be open source. More specifically, it should be possible for anyone—including people who are not registered users of SPF—to view the source code and documentation of any completed version of SPM, and for any registered user to run any completed version.

In keeping with these points, all versions of SPM should be "free software" conforming to the GNU General Public License, described on www.gnu.org/licenses/gpl.html. In brief, this means that every user should have:

- The freedom to use the software for any purpose;

- The freedom to change the software to suit his or her needs;

- The freedom to share the software with friends and neighbours; and

- The freedom to share the changes that any user makes.

This policy conforms to and supports the long-established principle that science works best when ideas, observations and experimental results are freely available to everyone, without restrictions. This kind of openness and transparency is essential if we are to make the difficult transition to cognitive computing (Section 2.3.1).

4 Costs

Costs associated with SPF are estimated in the following subsections. A spreadsheet that summarises all the estimated costs, with totals, is shown in Section 4.4.

4.1 Development costs

We estimate that a 'starter' cluster machine may be purchased for about £100,000.

We estimate that three members of staff with the right skills who are already familiar with the host machine would be able to develop SPF within 2 years. We suggest that it would be appropriate to employ one relatively senior person, perhaps at a salary of about £70,000, one more junior person, perhaps at a salary of about £50,000, and a third person, possibly a student, at a salary of about £30,000. We would also need the services of a technician at a salary of about £30,000.

We have assumed that the cost of employing someone is twice the cost of his or her salary, and of course salaries would be paid for two years.

We have allowed £15,000 for PCs, printers, software, and related facilities for the three staff.

In view of the importance of the website for SPF and the specialised skills required for good website design—including the skills needed to ensure that the website is "mobile friendly"—we believe it would be prudent to employ a web design company to establish the framework for the website and advise on how it should be developed. We estimate that £2000 should be allowed for this work.

We have allowed 10% for contingencies.

4.2 Future costs

Our estimate of development costs includes the cost of a relatively small 'starter' cluster machine. If use of the facility grows as expected, it is likely that the starter machine will either need to be expanded or replaced with a bigger machine. It is possible that that may need to be repeated later. This is a contingency that needs to be recognised as part of the potential cost of the project.

4.3 Recurrent costs

The annual cost of running SPF are estimated in the following subsections and summarised in Section 4.4.

4.3.1 Maintenance costs

We believe that when SPF is up and running, it will need support: to make sure that it runs smoothly and to deal with any snags that arise; to answer queries by users; to fix bugs in the system; and to refine the system in the light of feedback from users.

We believe that these things can be done by one person with a salary of about £50,000 pa, with an assistant at about £30,000 pa. We would also need the services of a technician at a salary of about £30,000 pa.

4.3.2 Running costs

As mentioned in Section 3.6, we believe that much of the research with SPF would make small demands on the host computer and, probably, that the associated costs would be smaller than the cost of administering a system of charges. Nevertheless, such costs do need to be covered. We estimate that they would be £10,000 pa.

4.3.3 Awareness-raising costs

In accordance with "impact" policies of the UK Government and UK Research Councils (Section 2.9), it is important that SPF and what it can do should be well known around the world, amongst academic and industrial researchers, amongst research administrators and policy-makers, and amongst members of the general public. This may be achieved via the following main routes, each one with associated costs:

- Via papers in academic journals and conference proceedings. Most of these would be prepared by research groups and individual researchers that are using SPF, and they would pay the associated costs. But, in addition, there would be academic papers about SPF and SPM, prepared by the two authors of this proposal and staff recruited to develop SPF. The main costs here would be the costs of attending conferences and charges for open-access publication.

- The proposers, and perhaps development staff as well, may travel to give talks about SPF and the SP concepts.

- Perhaps via conferences (online or traditional) dedicated to research using SPF. The organisation may be undertaken by a specialist conference organiser, and there would be associated costs.

- Via articles in popular-science magazines, both those for the general reader and those with a more technical or specialist orientation. Here, it would be useful if articles could be prepared by writers with relevant skills (who would need to be paid), although technical input would be provided by the authors of this proposal and research/development staff. Articles may be translated from English into other languages.

- Via the mainstream media (newspapers, radio and TV), internet blogs, and online videos. Again, it would be useful if articles, videos and the like could be prepared by people with relevant skills, with guidance from people with relevant technical knowledge.

We estimate that a budget of £50,000 pa would be needed to cover these costs during the two-year development phase, reducing to £20,000 pa for a further three years. Although we believe it is important that SPF should be a permanent facility, as described in Section 3.4, we shall assume for present purposes that awareness-raising costs will be spread over a period of 20 years. In that case, the annual cost will be $£((50,000 \times 2) + (20,000 \times 3))/20 = £8,000$.

4.3.4 Contingencies

As with the non-recurrent development costs, we have allowed 10% for contingencies.

4.4 Summary of costs

In this section, Table 1 summarises our estimates of non-recurrent costs for SPF ($£546,700$ in total) and recurrent costs ($£151,800$ pa in total).

It addition to these costs, there is a clear possibility that, with growing use, the host machine will either need to be expanded or replaced with a bigger machine, and this may happen more than once (Section 4.2).

5 Conclusion

The wide scope of the SP theory and the wide range of its potential benefits and applications means that there are many more avenues to be explored than could be tackled by any one research group. By providing the means for researchers worldwide to participate and collaborate, the research facility that is proposed in this document would facilitate the further development of the SP system and the realisation of its potential.

In view of that potential, and the relatively small cost of the proposed facility, the creation of that facility would be a good investment.

Appendices

A Outline of the SP theory and SP computer model

This Appendix is intended to provide readers with sufficient information about the SP theory to make the rest of the proposal intelligible. The outline starts with a short informal account of the SP system, followed by subsections that describe the main elements of the system in a little more detail.

A.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* information via its senses, and compress some or all of it to create *Old* information, as illustrated schematically in Figure 2.

Category	£	£
Non-recurrent (development)		
Development salaries		
Senior	70,000	
Middle	50,000	
Junior	30,000	
Technician	30,000	
Total (salaries, one year)	180,000	
Total (salaries + overheads, one year)	360,000	
Total (salaries + overheads, two years)		720,000
High-parallel machine (small)		100,000
PCs etc		15,000
Website design		2,000
Awareness raising		8,000
Sub-total (non-recurrent)	845,000	
Contingencies (10%)		84,500
Total (non-recurrent)		929,500
Recurrent (per year)		
Maintenance salaries		
Middle	50,000	
Junior	30,000	
Technician	30,000	
Total (salaries)	110,000	
Total (salaries + overheads)		220,000
Running costs		10,000
Awareness raising		8,000
Sub-total (recurrent)	238,000	
Contingencies (10%)		23,800
Total (recurrent)		261,800

Table 1: A spreadsheet summarising our estimates of non-recurrent and recurrent costs for SPF.
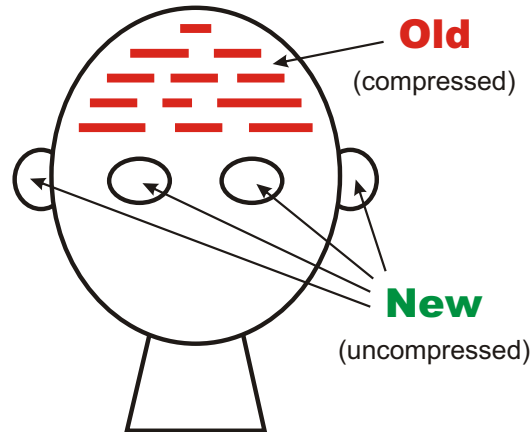
Figure 2: Schematic representation of the SP system from an 'input' perspective.

A central idea in the SP system is a concept of *multiple alignment*, borrowed and adapted from that concept in bioinformatics. This is outlined in Appendix A.4.2 and illustrated in Figure 4.

In the early stages of learning, most New information will be simply stored directly as Old information, with the addition of 'identification' symbols at each end. For example, a New pattern like '`w i n d o w c l e a n e r`' may be stored as an Old pattern something like '`A w i n d o w c l e a n e r #A`', where '`A`' and '`#A`' are 'identification' symbols and all the lower-case characters are 'contents' symbols.

Later, if a New pattern like '`w i n d o w`' were to be received, the system would create a multiple alignment between that pattern and the previously-stored '`A w i n d o w c l e a n e r #A`', as shown in Figure 3. Then:

- From '`w i n d o w`', the system would create an Old pattern like '`X w i n d o w #B`', much as before.

- From the unmatched contents symbols in '`A w i n d o w c l e a n e r #A`' in the multiple alignment, the system would create an Old patterns like '`C c l e a n e r #C`'.

- And, in addition, the system would create a pattern like '`A B #B C #C #A`', where the pair of symbols 'textttB #B' may be seen as a 'reference' to '`B w i n d o w #B`', and 'textttC #C' may be seen as a reference to the pattern '`C c l e a n e r #C`'.

Many of the Old patterns that are created in this way may be seen as 'good', but the system also creates many that people are likely to regard as 'bad'. To

```
0   w i n d o w           0
    | | | | | |
1 A w i n d o w c l e a n e r #A 1
```

Figure 3: A multiple alignment between 'B w i n d o w #B' and 'A w i n d o w c
l e a n e r #A', as described in the text.

get rid of the bad patterns, SP system applies a process for evaluating sets of
patterns (termed *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 (Appendix
A.4).

The building of multiple alignments provides the means by which New informa-
tion may be encoded economically in terms of Old information. It also provides
much of the versatility of the system in such functions as the parsing and pro-
duction of natural language, pattern recognition, computer vision, information
retrieval, several kinds of reasoning, planning, problem solving, and information
compression. And it has a central role in unsupervised learning—how the system
discovers or infers good grammars.

A.2   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*.

As noted in the Introduction, the use of a computer model as a vehicle for the
theory has helped to reduce vagueness in the theory, it has provided a means of
testing candidate ideas, many of which have been rejected, and it is a means of
demonstrating what can be done with the system.

An outline of how the SP computer model works may be found in [19, Section
3.9], with more detail, including pseudocode, in [19, Sections 3.10 and 9.2].[35] 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.

---

[35]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 [19, Sections 3.9.1 and 3.10], of
a subset of the SP70 model called SP61.

As previously noted, we envisage that the SP computer model will be the basis for the creation of the high-parallel, open-source version of the SP machine. How things may develop is shown schematically in Figure 1.

A.3 Patterns and symbols

Knowledge in the SP system is represented with arrays of atomic symbols in one or two dimensions called *patterns*. The SP71 model works with 1D patterns but it is envisaged that the system will be generalised to work with 2D patterns [21, Section 3.3].

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.

In themselves, SP patterns are not particularly expressive. But within the multiple alignment framework (Appendix A.4.2), they support the representation and processing of a wide variety of kinds of knowledge (Appendix A.5). A goal of the SP research programme is to establish one system for the representation and processing of *all* kinds of knowledge (see also [24, Section III]). Evidence to date suggests that this may be achieved with SP patterns in the multiple alignment framework.

Any collection of SP patterns is termed a *grammar*. Although that term is most closely associated with linguistics, it is used in the SP research programme for a collection of SP patterns describing any kind of knowledge.

A.4 Information compression

In the SP theory, the emphasis on information compression derives from earlier research on grammatical inference [16] and the principle of *minimum length encoding* (MLE) [12, 15, 10]).

At an abstract level, information compression means the detection and reduction of *redundancy* in information. In more concrete terms, redundancy means *recurrent patterns*, *regularities*, *structures*, and *associations*, including *causal associations*. Thus information compression provides a means of discovering such things as words in natural language [16], objects [23, Section V-E], and associations (see, for example, [24, Section III-A.1]), in accordance with the DONSVIC principle [21, Section 5.2].[36]

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

---

[36]*DONSVIC* = "The discovery of natural structures via information compression".

(see, for example, [24, Section X-B]) but any such discard is reversible.

In the SP system, information compression is achieved via the matching and unification of patterns. 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.

A.4.1 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. At the heart of the SP71 model is a method for finding good full and partial matches between sequences with advantages compared with classical methods [19, Appendix A].[37]

A.4.2 Information compression via the building of multiple alignments

That 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 4. This concept is similar to multiple alignment in bioinformatics but with important differences [19, Section 3.4]. It is a powerful and distinctive feature of the SP system.

---

[37]The main advantages are [19, 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 can be controlled by parameters.

```
0                 1                 2                 3                 4                 5                 6

                  <species>
                  acris
                  <genus> -------------------------------------------------------------------- <genus>
                  Ranunculus ----------------------------------------------------------------- Ranunculus
                                                                                      <family> --------- <family>
                                                                                      Ranunculaceae ---- Ranunculaceae
                                                                    <order> ------ <order>
                                                                    Ranunculales - Ranunculales
                                                  <class> ------ <class>
                                                  Angiospermae - Angiospermae
                                    <phylum> --------- <phylum>
                                    Plants ----------- Plants
                                    <feeding>
has_chlorophyll ------------------ has_chlorophyll
                                    photosynthesises
                                    <feeding>
                                    <structure> ------ <structure>
                                                       <shoot>
<stem> ---------- <stem> -------------------------- <stem>
hairy ---------- hairy
</stem> -------- </stem> ------------------------- </stem>
                  <leaves> ------------------------ <leaves>
                  compound
                  palmately_cut
                  </leaves> ---------------------- </leaves>
                                                   <flowers> ------------------ <flowers>
                                                                                <arrangement>
                                                                                regular
                                                                                all_parts_free
                                                                                </arrangement>
                  <sepals> --------------------------------------------------- <sepals>
                  not_reflexed
                  </sepals> -------------------------------------------------- </sepals>
<petals> ------- <petals> --------------------------------------------------- <petals> --------- <petals>
                                                                                <number> --------- <number>
                                                                                                   five
                                                                                </number> ------- </number>
                  <colour> --------------------------------------------------- <colour>
yellow --------- yellow
                  </colour> -------------------------------------------------- </colour>
</petals> ------ </petals> --------------------------------------------------- </petals> ------- </petals>
                                                                                <hermaphrodite>
<stamens> ------------------------------------------------------------------- <stamens>
numerous -------------------------------------------------------------------- numerous
</stamens> ------------------------------------------------------------------ </stamens>
                                                                                <pistil>
                                                                                ovary
                                                                                style
                                                                                stigma
                                                                                </pistil>
                                                                                </hermaphrodite>
                                                   </flowers> ----------------- </flowers>
                                                   </shoot>
                                                   <root>
                                                   </root>
                                    </structure> ----- </structure>
<habitat> ------- <habitat> ------ <habitat>
meadows --------- meadows
</habitat> ------ </habitat> ----- </habitat>
                  <common_name> -- <common_name>
                  Meadow
                  Buttercup
                  </common_name> - </common_name>
                                    <food_value> ------------------------------ <food_value>
                                                                                poisonous
                                    </food_value> ----------------------------- </food_value>
                                    </phylum> -------- </phylum>
                                                  </class> ----- </class>
                                                             </order> ----- </order>
                                                                                      </family> -------- </family>
                  </genus> ------------------------------------------------------------------- </genus>
                  </species>

0                 1                 2                 3                 4                 5                 6
```

Figure 4: The best multiple alignment created by the SP computer model, with a set of New patterns (in column 0) that describe some features of an unknown plant, and a set of Old patterns, including those shown in columns 1 to 6, that describe different categories of plant, with their parts and sub-parts, and other attributes.

This example shows the best multiple alignment created by the SP computer model when a set of New patterns (in column 0)[38] is processed in conjunction with a set of pre-existing Old patterns like those shown in columns 1 to 6. Here, the multiple alignment is 'best' because it is the one that achieves the most economical description of the New patterns in terms of the Old patterns. The way in which that description or 'encoding' is derived from a multiple alignment is explained in [19, Section 3.5] and [21, Section 4.1]. Like all other kinds of knowledge, encodings derived from multiple alignments are recorded using SP patterns (Appendix A.3).

This multiple alignment may be interpreted as the result of a process of recognition (Appendix C.5). The New patterns represent the features of some unknown plant and the Old patterns in columns 1 to 6 show how the plant has been identified at several levels of abstraction: species 'Meadow Buttercup' (column 1), genus *Ranunculus* (column 6), family *Ranunculaceae* (column 5), and so on.

A.4.3 Information compression via the unsupervised learning of grammars

As outlined in [19, Section 3.9.2] and [21, Section 5.1], and described more fully in [19, Chapter 9], the SP system may, without assistance from a "teacher" or anything equivalent, derive one or more plausible grammars from a body of New patterns, with minimum length encoding as a guiding principle. In that process, multiple alignment has a central role as a source of SP patterns for possible inclusion in any grammar [23, Section V-B1].

A.4.4 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 in all three cases men-

---

[38]Specifically, the New patterns in this example are '`has_chlorophyll`' (a pattern with one symbol), '`<stem> hairy </stem>`', '`<petals> yellow </petals>`', '`<stamens> numerous </stamens>`', and '`<habitat> meadows </habitat>`'. The patterns in a set like that may be presented to the system and processed in any order.

tioned above. This means searching for solutions in stages, with a pruning of the search tree at every stage, guided by measures of success [19, Appendix A; Sections 3.9 and 3.10; Chapter 9]. 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.

A.5 Multiple alignment and the representation and processing of diverse kinds of knowledge

The expressive power of SP patterns within the multiple alignment framework derives in large part from the way that symbols in one pattern may serve as links to one or more other patterns or parts thereof. One of several examples in Figure 4 is how the pair of symbols '`<family>` ... `</family>`' in column 6 serves to identify the pattern '`<family>` ... `Ranunculales` ... `<hermaphrodite>` ... `poisonous` ... `</family>`' in column 5.

In the figure, these kinds of linkages between patterns mean that the unknown plant (with characteristics shown in column 0) may be recognised at several different levels within a hierarchy of classes: genus, family, order, class, and so on. Although it is not shown in this example, the system also supports cross classification.

In the figure, the parts and sub-parts of the plant are shown in such structures as '`<shoot>`' (column 3), '`<flowers>`' (column 5), '`<petals>`' (column 6), and so on.

As in conventional systems for object-oriented design, the system provides for inheritance of attributes (Appendix C.8). But unlike such systems, there is smooth integration of class hierarchies and part-whole hierarchies, without awkward inconsistencies [20, Section 4.2.1].

More generally, SP patterns within the multiple alignment framework provide for the representation and processing of a wide variety of kinds of knowledge including: the syntax and semantics of natural language; class hierarchies and part-whole hierarchies (as just described); networks and trees; entity-relationship structures; relational knowledge; rules and several kinds of reasoning; patterns and pattern recognition; images; structures in three dimensions; and procedural knowledge. There is a summary in [24, Section III-B], and more detail in Appendix C.

A.6 Information compression, prediction, and probabilities

Owing to the close connection between information compression and concepts of prediction and probability [8], the SP system is fundamentally probabilistic.[39] Each SP pattern has an associated frequency of occurrence and probabilities may be calculated for each multiple alignment and for any inference that may be drawn from any given multiple alignment.

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 [19, Section 2.3.7].

A.7 SP-neural

Part of the SP theory is the idea, described most fully in [19, 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.

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 Donald Hebb's [5] 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 [19, Section 11.4.1]).

It is pertinent to mention that unsupervised learning in the SP theory ([19, Chapter 9], [21, Section 5]) is quite different from "Hebbian learning" as described by Hebb [5] and widely adopted in the kinds of artificial neural networks that are popular in computer science.[40] By contrast with Hebbian learning, the SP system, like a person, may learn from a single exposure to some situation or event. And, by contrast with Hebbian learning, it takes time to learn a language in the SP

---

[39]There are reasons to believe that the same is true of "computing" at its most fundamental level, and also mathematics. For example, Gregory Chaitin has written "I have recently been able to take a further step along the path laid out by Gödel and Turing. By translating a particular computer program into an algebraic equation of a type that was familiar even to the ancient Greeks, I have shown that there is randomness in the branch of pure mathematics known as number theory. My work indicates that—to borrow Einsteins metaphor—God sometimes plays dice with whole numbers." [4, p. 80].

[40]See, for example, "Hebbian theory", *Wikipedia*, bit.ly/1sW6ATt, retrieved 2014-12-19.

system because of the complexity of the search space, not because of any kind of gradual strengthening or "weighting" of links between neurons [19, Section 11.4.4].

Figure 5 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.[41] 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.

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 [19, 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.

B Distinctive features and apparent advantages of the SP system

Information compression and concepts of probability are themes in other research, including research on Bayesian inference, Kolmogorov complexity, deep learning, artificial neural networks, minimum length encoding, unified theories of cognition, natural language processing and more. The main features that distinguish the SP system from these other areas of research, and apparent advantages compared with these other approaches, are described quite fully in "The SP theory of intelligence: its distinctive features and advantages" [27]. The main points are summarised here:

- *Simplification and integration of observations and concepts.* As mentioned above, the SP theory is a unique attempt to simplify and integrate observations and concepts across artificial intelligence, mainstream computing, mathematics, and human perception and cognition. The canvass is much broader than it is, for example, in "unified theories of cognition". It has quite a lot to say, for example, about the nature of mathematics [17], [19, Chapter 10], [25].

- *Simplification and integration of computing systems.* The provision of one simple format for knowledge, and one framework—multiple alignment—for

---

[41]Not shown in the figure are lateral connections within each pattern assembly and inhibitory connections elsewhere, as outlined in [19, Sections 11.3.3 and 11.3.4].

To higher-level pattern assemblies

< *NP* < D > < N > >

< *D 0* t h i s >

< *D 1* t h a t >

< *N 0* g i r l >

< *N 1* b o y >

From lower-level pattern assemblies
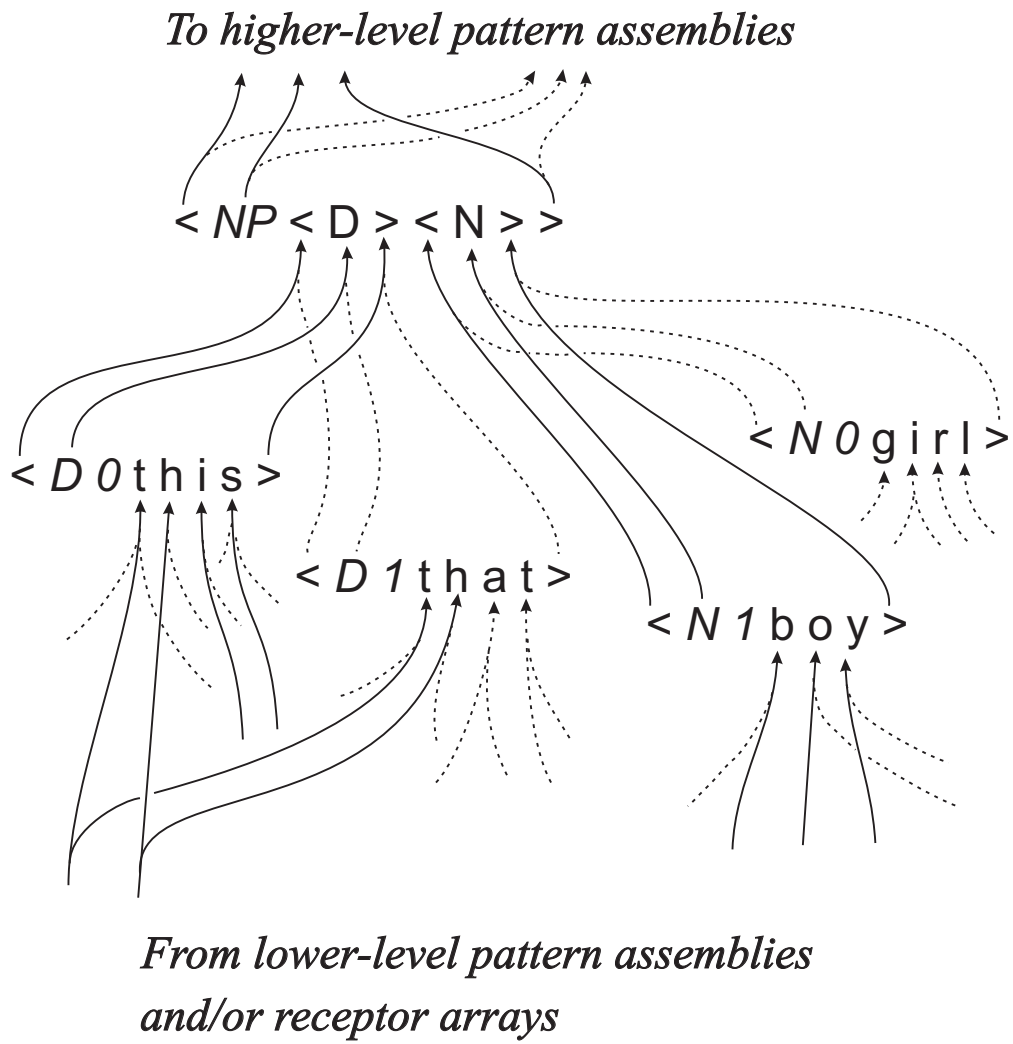*and/or receptor arrays*

Figure 5: 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.

the representation and processing of knowledge, promote an overall simplification of computing systems, including both hardware and software [26, Section 5]. They also promote seamless integration of diverse structures and functions [26, Section 7].

- *Descriptive and explanatory range.* It appears that the descriptive and explanatory range of the SP system is much wider than any of the alternatives mentioned above. In terms of demonstrable capabilities as well as abstract concepts, it has strengths in areas that include the representation of diverse forms of knowledge (including class hierarchies, part-whole hierarchies, and their seamless integration), unsupervised learning, natural language processing, fuzzy pattern recognition and recognition at multiple levels of abstraction, best-match and semantic forms of information retrieval, several kinds of reasoning (one-step 'deductive reasoning', abductive reasoning, probabilistic networks and trees, reasoning with 'rules', nonmonotonic reasoning, explaining away, causal reasoning, and reasoning that is not supported by evidence), planning, problem solving, information compression, and aspects of neuroscience and of human perception and cognition.

- *Many potential benefits and applications.* In the spirit of the quote at the start of the Introduction ("There is nothing more practical than a good theory"), the SP theory, with its broad base of support, has many potential benefits and applications, as outlined in Appendix C.

- *Strong theoretical underpinnings.* Compared with, for example, artificial neural networks and deep learning, the SP system appears to have much stronger and more coherent theoretical underpinnings. The multiple alignment framework provides a simple but powerful realisation of the concept on which the SP theory is founded: information compression via the matching and unification of patterns.

- *The SP theory is a theory of computing.* Most other research is founded on the idea that computing may be understood in terms of the Universal Turing Machine or equivalent models such as Lamda Calculus or Post's Canonical System. By contrast, *the SP theory is itself a theory of computing* [19, Chapter 4]. *What is distinctive about the SP theory as a theory of computing is that it provides much of the human-like intelligence that is missing from earlier models.*[42]

_____

[42]Although Alan Turing saw that computers might become intelligent [14], the Universal Turing Machine, in itself, does not tell us how! The SP theory goes some way towards plugging the gap, with potential to do more.

- *Information compression via the matching and unification of patterns.* In trying to cut through complexities, the SP research programme focuses on a simple, 'primitive' idea: that information compression may be understood as a search for patterns that match each other, with the merging or 'unification' of patterns that are the same. The potential advantage of this approach, focussing on the simple concept of matching and unifying patterns, is that it can help us avoid old tramlines, and open doors to new ways of thinking.

- *Multiple alignment.* More specifically, information compression via the matching and unification of patterns provides the basis for a concept of *multiple alignment*, borrowed and adapted from that concept in bioinformatics. Developing this idea as a framework for the simplification and integration of concepts across a broad canvass has been a major undertaking. *Multiple alignment is a distinctive and powerful idea in the SP research programme.*

- *Transparency in the representation and processing of knowledge.* By contrast with sub-symbolic approaches to artificial intelligence (artificial neural networks, deep learning, and related approaches), and notwithstanding objections to symbolic AI,[43] knowledge in the SP system is transparent and open to inspection, and likewise for the processing of knowledge.

- *The DONSVIC principle.* A related point is that unsupervised learning in the SP system is geared to the realisation of the "DONSVIC" principle— *The Discovery of Natural Structures Via Information Compression* [21, Section 5.2]. By contrast with sub-symbolic approaches to artificial intelligence, structures created by learning should, normally, be comprehensible by people.

- *Perception and cognition.* The SP theory draws extensively on research on human and animal perception and cognition, and neuroscience. In particular, an important part of its inspiration is research on the learning of natural language (see www.cognitionresearch.org/lang_learn.html).

- *SP-neural.* The SP theory includes proposals—SP-neural—for how abstract concepts in the theory may be realised in terms of neurons and neural processes. The SP-neural proposals (Appendix A.7) are significantly different from artificial neural networks as commonly conceived in computer science, and arguably more plausible in terms of neuroscience.

---

[43]See, for example, "Hubert Dreyfus's views on artificial intelligence", *Wikipedia*, bit.ly/1hGHVm8, retrieved 2014-08-19.

C Some potential benefits and applications of the SP system

With the SP system, there is a large range of potential benefits and applications, some of which are outlined in the following subsections.

Some potential applications 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 (Appendix C.6), and applications in such areas as medical diagnosis (Appendix C.13.3), pattern recognition (Appendix C.5), information compression (Appendix C.13.5), highly-economical transmission of information (Appendix C.13.1, [24, Section VIII]), bioinformatics (Appendix C.13.4), and natural language processing (Appendix C.4.3).

More radical solutions, that may take longer to develop, include a radically new architecture for computers (Appendix C.2), and developing the full potential of the system in computer vision (Appendix C.7), and natural language processing (Appendix C.4).

C.1 Simplification and integration of concepts

Although some people may argue otherwise, the world of computing suffers from a deep malaise: its fragmentation into a myriad of concepts, shown schematically in Figure 6, a myriad of different formalisms and formats for representing knowledge, each with its own mode of processing [24, Section III], and the extraordinary complexity of many computing systems, especially software.

The quest for simplification and integration in the SP theory accords with Occam's Razor, one of the most widely-accepted principles in science. In terms of that principle, seeking to combine conceptual *simplicity* with descriptive and explanatory *power*:

- A relatively simple conceptual framework—multiple alignment—provides an account of a wide range of concepts and phenomena [19, 21].

- There is potential for very substantial simplification of computing systems, taking hardware and software together [26, Section 5], and for seamless integration of diverse structures and functions [26, Sections 2 and 7].

- Like any theory that simplifies and integrates a good range of observations and concepts, the SP theory promises deeper insights and better solutions to problems than may otherwise be achieved [26, Section 6], [23, Section IV-A3].

Figure 6: Schematic representation of the fragmentation of computer science.

C.2   A radically new architecture for computers

The SP system has the potential to provide the basis for a radically new architecture for computers, either via the abstract model (with a central role for multiple alignment) or, more likely, via SP-neural. Potential benefits of this new architecture would be human-like versatility and adaptability in computing, substantial simplification of software, and dramatic reductions in the energy demands of computers, and in their size and weight [24, Section IX].

C.3   Unsupervised learning

The SP theory originates in part from an earlier programme of research on grammatical inference and the unsupervised learning of natural language, with minimum length encoding as a central principle [16], [23, Section V-A3]. However, meeting the goals of the SP research programme has meant a radical reorganisation of computer models, with the development of multiple alignment as a framework for the simplification and integration of diverse structures and functions [23, Section V-A4]. The new model demonstrates capabilities in grammatical inference ([19, Chapter 9], Appendix A.4.3) and appears to have considerable potential for further development with linguistic, visual, and other kinds of knowledge.

C.4 Natural language processing

In addition to the learning of linguistic knowledge (Appendix C.3), the SP system has strengths in the parsing of natural language, the production of natural language, and the integration of syntactic and semantic knowledge, as outlined in this section. These aspects of the system are described more fully in [21, Section 8] and in [19, Chapter 5].

C.4.1 Parsing of natural language

Figure 7 shows how, via multiple alignment, a sentence (in row 0) may be parsed in terms of grammatical structures including words (rows 1 to 8).[44] It also shows, in row 8, how the system may mark the syntactic dependency between the plural subject of the sentence ('Np') and the plural main verb ('Vp') (see also [19, Sections 5.4 and 5.5], [21, Section 8.1]).

```
0                       t h e         a p p l e   s         a r e       s w e e t      0
                        | | |         | | | | |   |         | | |       | | | | |
1                       | | |       N Nr 6 a p p l e #N |         | | |       | | | | |      1
                        | | |         | |     | |           | | |       | | | | |
2                       | | |   N Np N Nr             #N s #N         | | |       | | | | |      2
                        | | |     | |                   |           | | |       | | | | |
3               D 17 t h e #D | |                 |         | | |       | | | | |      3
                  |       | | |                 |         | | |       | | | | |
4           NP 0a D       #D N |               #N #NP       | | |       | | | | |      4
            |           |               |       | | |       | | | | |
5           |           |                 | V Vp 11 a r e #V   | | | | |      5
            |           |                 | | |         |   | | | | |
6 S Num    ; NP          |             #NP V |         #V A  | | | | | #A #S 6
    |       |           |                 |           |   | | | | | |
7   |       |           |                 |         A 21 s w e e t #A   7
    |       |           |                 |
8   Num PL ;                   Np                 Vp                  8
```

Figure 7: The best multiple alignment created by the SP model with a store of Old patterns like those in rows 1 to 8 (representing grammatical structures, including words) and a New pattern (representing a sentence to be parsed) shown in row 0.

To create a multiple alignment like the one in the figure, the system needs a grammar of Old patterns, like those shown, one per row, in rows 1 to 8 of the figure. In this example, the patterns represent linguistic structures including words.

Although SP patterns are remarkably simple, it appears that, within the multiple alignment framework, they have at least the expressive power of a context-sensitive grammar [19, Sections 5.4 and 5.5]. As previously noted (Appendix A.3), there is reason to believe that, within the multiple alignment framework, all kinds of knowledge may be represented by SP patterns.

---

[44]Compared with the multiple alignment shown in Figure 4, this multiple alignment is rotated through 90°, replacing columns with rows. The choice between these two styles, which are equivalent, depends largely on what fits best on the page.

C.4.2   Production of natural language

A neat feature of the SP system is that one set of mechanisms may achieve both the analysis or parsing of natural language (Appendix C.4.1) and the generation or production of sentences. This is explained in [19, Section 3.8] and [21, Section 4.5].

C.4.3   The integration of syntax and semantics

The use of one simple format for all kinds of knowledge is likely to facilitate the seamless integration of syntax and semantics. Preliminary examples of how this may be done are shown in [19, Section 5.7], both for the derivation of meanings from surface forms [19, Figure 5.18] and for the production of surface forms from meanings [19, Figure 5.19].

It appears that, on relatively short timescales, there is potential via the SP system to create natural language interfaces for such things as internet TVs, DVD recorders and the like, and thus help to overcome much of the difficulty that many people now experience in controlling these things.

C.4.4   Parallel streams of information

Up to now, most work on natural language within the SP research programme has made the simplifying assumption that language may be represented with a sequence of symbols, as in ordinary text. But with some aspects of natural language such as formants in speech, and the relationship between syntax and semantics, there seem to be parallel streams of information. The way in which such parallelism may be represented and processed with 2D patterns in the SP system is described in [23, Section IV-B4 and Appendix C].

C.5   Pattern recognition

As described quite fully in [19, Chapter 6] and more briefly in [21, Section 9], the SP system has strengths in several aspects of pattern recognition:

- It can recognise patterns at multiple levels of abstraction, with the integration of class-inclusion relations and part-whole relations, as shown in the example in Figure 4.

- It can model "family resemblance" or polythetic categories, meaning that recognition does not depend on the presence absence of any particular feature or combination of features.

- Recognition is robust in the face of errors of omission, commission or substitution in the New pattern or patterns.

- For any given identification, or any related inference, the SP system may calculate associated probabilities.

- As a by-product of how recognition is achieved via the building of multiple alignments, the system provides a model for the way in which context may influence recognition.

C.6 Information storage and retrieval, with intelligence

The SP theory provides a versatile model for database systems, with the ability to accommodate object-oriented structures, as well as relational 'tuples', and network and tree models of data [20]. It lends itself most directly to information retrieval in the manner of query-by-example but it appears to have potential to support the use of natural language or query languages such as SQL.

Unlike some ordinary database systems:

- The storage and retrieval of information is integrated with other aspects of intelligence such as pattern recognition, reasoning, planning, problem solving, and learning—as outlined elsewhere in this document.

- The SP system provides a simple but effective means of combining class hierarchies with part-whole hierarchies, with inheritance of attributes (Appendix A.5).

- It provides for cross-classification with multiple inheritance.

- There is flexibility and versatility in the representation of knowledge arising from the fact that the system does not distinguish 'parts' and 'attributes' [20, Section 4.2.1].

- Likewise, the absence of a distinction between 'class' and 'object' facilitates the representation of knowledge and eliminates the need for a 'metaclass' [20, Section 4.2.2].

- SP patterns provide a simpler and more direct means of representing entity-relationship models than do relational tuples [20, Section 4.2.3].

C.7   Vision

With generalisation of the SP system to accommodate 2D patterns, it has potential to model several aspects of natural vision and to facilitate the development of human-like abilities in artificial vision [22]. In these connections, the main strengths and potential of the SP system are:

- Low level perceptual features such as edges or corners may be identified via the multiple alignment framework by the extraction of redundancy in uniform areas in the manner of the run-length encoding technique for information compression.

- The system may be applied in the recognition of objects and in scene analysis, with the same strengths as in pattern recognition (Appendix C.5).

- There is potential for the learning of visual entities and classes of entity and the piecing together of coherent concepts from fragments [22, Section 5].

- There is potential, via multiple alignment, for the creation of 3D models of objects and of surroundings [22, Section 6].

- The SP theory provides an account of how we may see things that are not objectively present in an image, how we may recognise something despite variations in the size of its retinal image, and how raster graphics and vector graphics may be unified.

- And the SP theory has things to say about the phenomena of lightness constancy and colour constancy, ambiguities in visual perception, and the integration of vision with other senses and other aspects of intelligence.

C.8   Reasoning

As described in quite fully in [19, Chapters 7 and 10, Section 6.4] and more selectively in [21, Section 10], the SP system lends itself to several kinds of reasoning:

- One-step 'deductive' reasoning.

- Abductive reasoning.

- Reasoning with probabilistic decision networks and decision trees.

- Reasoning with 'rules'.

- Nonmonotonic reasoning and reasoning with default values.

- Reasoning in Bayesian networks, including "explaining away".

- Causal diagnosis.

- Reasoning which is not supported by evidence.

- Inheritance of attributes in an object-oriented class hierarchy or heterarchy.

There is also potential for spatial reasoning [23, Section IV-F1] and what-if reasoning [23, Section IV-F2].

These several kinds of reasoning may work together seamlessly without awkward incompatibilities, and likewise for how they may integrate seamlessly with such AI functions as unsupervised learning, pattern recognition, and so on [26, Sections 2, 4, and 7].

For any given inference reached via any of these kinds of reasoning, the SP system may calculate associated probabilities (Appendix A.6).

Although the system is fundamentally probabilistic, it may imitate the effect of logic and other 'exact' forms of reasoning [19, Section 10.4.5].

C.9   Planning and problem solving

With data about flights between different cities, represented using SP patterns, the SP computer model may find a route between any two cities (if such a route exists) and, if there are alternative routes, it may find them as well [19, Section 8.2].

Provided they are translated into textual form, the SP computer model can solve geometric analogy problems of the kind found in puzzle books and some IQ tests [19, Section 8.3], [21, Section 12].

C.10   Software engineering

Although it may not seem obvious at first sight, the multiple alignment framework can model several devices used in ordinary procedural programming, including: *procedure*, *function*, or *subroutine*; *variable*, *value* and *type*; *function with parameters*; *conditional statement*; and the means of repeating operations such as *repeat ... until* or *do ... while* [26, Section 6.6.1]. In accordance with good practice in software engineering, the SP system facilitates the integration of 'programs' with 'data'. And as previously noted (Appendix A.5), the SP system supports object-oriented concepts such as class hierarchies with inheritance of attributes.

In [26, Section 6.6.3], it is suggested that, since SP patterns at the 'top' level are independent of each other, they may serve to model processes that may run in parallel. Now it appears that a better option is to model parallel processes as

parallel streams of information, represented in 2D SP patterns as described in [23, Appendix C]. The advantage of this latter scheme is that it provides the means of showing when two or more events occur at the same time and, more generally, the relative timings of events.

Within the SP system, these structures and mechanisms may serve in the representation and processing of sequential and parallel procedures from the real world such as those required for cooking a meal, organising a party, going shopping, and so on.

Potential benefits in software engineering include the elimination of compiling or interpretation, automatic programming, benefits in verification and validation, and helping to overcome the problem of technical debt [26, Section 6.6].

C.11 Mathematics

Aspects of mathematics may be understood in terms of some basic techniques for information compression: *chunking-with-codes*, *schema-plus-correction*, and *run-length coding* [25], and some features of mathematics may be modelled in the SP system [19, Chapter 10].

On the strength of this evidence and some other considerations, there is reason to believe that all of mathematics may be understood in terms of information compression. There appear to be considerable implications for mathematics, and also for science—because of the importance of mathematics as a language of science.

C.12 Human perception and cognition, and neuroscience

Part of the inspiration for the SP theory has been earlier research on the role of information compression in the workings of brains and nervous systems [1, 2, 3], and a programme of research on the learning of a first language or languages [16]. Thus there is reason to believe that the SP theory may help to illuminate human perception and cognition.

Since the converse is true—that insights into the nature of human perception and cognition may help to inform the development of the SP theory—the two things are often considered together, as can be seen in [19, 22] and other writings about the SP system.

As outlined in Appendix A.7, the SP theory includes the proposal—called *SP-neural*—that abstract concepts in the theory may be realised in terms of neurons and their interconnections [19, Chapter 11].

C.13　Other potential benefits and applications

The versatility of the SP system may be seen not only in the areas outlined above but in potential benefits and applications summarised in the following subsections.

C.13.1　Big data

The SP system may help to solve nine problems associated with big data [24]. In brief, these problems and their potential solutions are:

- *Overcoming the problem of variety in big data.* Harmonising diverse kinds of knowledge, diverse formats for knowledge, and their diverse modes of processing, via a universal framework for the representation and processing of knowledge (UFK).

- *Learning and discovery.* The unsupervised learning or discovery of 'natural' structures in data.

- *Interpretation of data.* The system has strengths in areas such as pattern recognition, information retrieval, parsing and production of natural language, translation from one representation to another, several kinds of reasoning, planning and problem solving.

- *Velocity: analysis of streaming data.* The SP system lends itself to an incremental style, assimilating information as it is received, much as people do.

- *Volume: making big data smaller.* Reducing the size of big data via lossless compression can yield several benefits.

- *Transmission of data.* There is potential for substantial economies in the transmission of data by judicious separation of 'encoding' and 'grammar'.

- *Energy, speed, and bulk.* There is potential for big cuts in the use of energy in computing, for greater speed of processing with a given computational resource, and for corresponding reductions in the size and weight of computers.

- *Veracity: managing errors and uncertainties in data.* The SP system can identify possible errors or uncertainties in data, suggest possible corrections or interpolations, and calculate associated probabilities.

- *Visualisation.* Knowledge structures created by the system, and inferential processes in the system, are all transparent and open to inspection. They lend themselves to display with static and moving images.

Considering these proposed solutions collectively, and in several cases individually, it appears that there are no alternatives that can rival the potential of what is described in [24].

C.13.2 Autonomous robots

The SP system may help in the design of autonomous robots, meaning robots that do not depend on external intelligence or power supplies, are mobile, and are designed to exhibit as much human-like intelligence as possible [23]. The three main areas where it may make a contribution—problems with the efficiency and bulkiness of computers; developing human-like versatility in robots; and developing human-like adaptability in robots—are summarised in Section 2.4.

C.13.3 Medical diagnosis

The way in which the SP system may be applied in medical diagnosis is described in [18]. The expected benefits of the SP system in that area of application include:

- A format for representing diseases that is simple and intuitive.

- An ability to cope with errors and uncertainties in diagnostic information.

- The simplicity of storing statistical information as frequencies of occurrence of diseases.

- The system provides a method for evaluating alternative diagnostic hypotheses that yields true probabilities.

- It is a framework that should facilitate the unsupervised learning of medical knowledge and the integration of medical diagnosis with other AI applications.

The main emphasis in [18] is on medical diagnosis as pattern recognition. But the SP system may also be applied to causal diagnosis [19, Section 7.9], [21, Section 10.5] so that it may be possible, for example, to reason that "The patient's fatigue may be caused by anemia which may be caused by a shortage of iron in the diet".

C.13.4 Bioinformatics

Because of the central importance of multiple alignment in the SP system, and because of the importance of that concept in bioinformatics, there is clear potential for the SP system to find applications in that area [26, Section 6.10.2]. Multiple alignment as it has been developed in the SP system has potential advantages compared with multiple alignment as it has been developed for bioinformatics.

Another potential advantage of the SP system is its capabilities and potential in unsupervised learning (Appendices A.4.3 and C.3, and sources referenced there):

- They have potential to discover recurrent patterns of arbitrary size within DNA sequences, amino acid sequences, and the like—including sequences that are discontinuous, jumping over intervening structures.

- They have potential to discover disjunctive classes of entity, with the context in which they are embedded.

- With the right kind of data, there is potential to discover structures and associations within and between DNA, proteins, enzymes, and associated biochemical processes.

- Likewise, there is potential to discover associations between diseases and their symptoms on the one hand, and biochemical structures and processes on the other.

C.13.5 Information compression

Since information compression is central in the workings of the SP system (Appendix A.4), there is reason to believe that an industrial-strength version of the system will be useful in that area of application.

References

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

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

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

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

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

[6] J. E. Kelly and S. Hamm. *Smart machines: IBM's Watson and the era of cognitive computing.* Columbia University Press, New York, Kindle edition, 2013.

[7] K. Lewin. *Field theory in social science: Selected theoretical papers by Kurt Lewin.* Tavistock, London, 1952.

[8] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications.* Springer, New York, 2009.

[9] National Research Council. *Frontiers in Massive Data Analysis.* The National Academies Press, Washington DC, 2013. ISBN-13: 978-0-309-28778-4. Online edition: bit.ly/14A0eyo.

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

[11] K. Sayood. *Introduction to data compression.* Morgan Kaufmann, Amsterdam, 2012.

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

[13] STFC. A new vision for new times. Published by the Science and Technology Funding Council of the UK, 2010. See bit.ly/1wCGIlN.

[14] A. M. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950.

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

[16] 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.

[17] J. G. Wolff. Computing, cognition and information compression. *AI Communications*, 6(2):107–127, 1993. See bit.ly/XL359b.

[18] 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].

[19] 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.

[20] 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].

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

[22] 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.

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

[24] 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.

[25] 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.

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

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