Branching Varieties of (Evolved) Consciousness

DRAFT -- WILL BE UPDATED

Notes relating to talk on consciousness at the Models of Consciousness conference at the Mathematical Institute, Oxford, 9-12 Sept, 2019.
https://models-of-consciousness.org/

Aaron Sloman http://www.cs.bham.ac.uk/~axs
School of Computer Science, The University of Birmingham UK.

Background
(Last updated: 16 Jan 2020)

This document is
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/consciousness-varieties.html
(also available as PDF).
Criticisms welcome: this is an early draft and may have gaps, errors and infelicities that will be remedied later.

Talks at the Oxford Consciousness conference were recorded and the videos made available here: https://www.youtube.com/channel/UCWgIDgfzRDp-PmQvMsYiNlg/videos.

My talk was on Tuesday morning 10th Sept. The video can be seen here:
https://www.youtube.com/watch?v=0DTYh37U8uE

Part of the conference talk (starting around 5:15) summarised the theory, developed with Jackie Chappell, that genomes can be "Meta-Configured" Chappell/Sloman(2007), an idea that is still under development. A more detailed explanation of the Meta-Configured Genome (MCG) hypothesis, including a short video, with additional discussion of implications for cognitive science and philosophy of mind is available here (still work in progress):
http://www.cs.bham.ac.uk/research/projects/cogaff/movies/meta-config/

This document provides draft explanatory notes about the background to the biological theory presented in the conference video and the above document. The ideas below seem to be consistent with, but significantly extend (I hope), some of the thoughts in William James (1879) 'Are we automata?' (mentioned in Adrian Kent's talk at the conference https://www.youtube.com/watch?v=Pli4TsEmODI). James asked (in effect) whether mental states and processes can be real and causally efficacious if everything is ultimately fully implemented in purely physical mechanisms. We now have an answer, which William James could not have given, in terms of what we have learnt about ways of implementing enduring virtual machines with causal powers that can persist across changes of physical implementation, as is constantly happening now to virtual machines running on the internet.

For a more detailed explanation of powers of virtual machinery see
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html. (This topic is resumed later.)

I plan to reorganise these notes later. A related unfinished document is concerned with varying roles during evolution and individual development of DNA and its products, indirectly supporting the hypothesis that there is far more sub-neural chemistry-based computation in brains than neural-net theorists take account of.
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/dna-uses.html

Conjecture: Chemical mechanisms may be able to explain forms of consciousness (e.g. mathematical consciousness of possibility, necessity and impossibility) that have nothing to do with statistics-based probabilities and cannot be explained by probabilistic neural nets, however "deep" they are!

I have a large and growing collection of examples available online. A somewhat disorganised subset originally planned for a lecture in Bristol University can be found here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html (also pdf)

HIGH LEVEL OVERVIEW
(Draft notes, added 14 Jan 2020)

William James was right to characterise consciousness as something that is (a) efficacious and (b) "has been slowly evolved in the animal series", since the earliest life forms, and resembles in this all organs that have a use. Moreover he points out that the most perfected parts of the brain are those whose action are least determinate, i.e. the least predictable on the basis of physical stimuli.

This difference is to be expected as control mechanisms make use of increasing amounts of information, simultaneously attempting to achieve or maintain increasingly varied goal states, or avoid increasingly complex and varied dangers, while managing increasingly complex and varied risks and opportunities.

The more complex states and processes involved in conscious activity (internal or externally visible), have many different forms, functions and mechanisms, and cannot be defined by some simple slogan (e.g. "what it's like to be/feel...") -- to which I add that types, functions and mechanisms of consciousness can also change dramatically during individual development (e.g. from new-born infant, or pre-born foetus, to professor of mathematics), producing considerable diversity of functions and therefore also diversity of requirements for explanatory mechanisms.

Requirements for a more complex elaboration of James' ideas:

- Deflating the criticism that his proposal admits "a dualism which, in the same breath that it allows to mind an independent status of being, banishes it to a limbo of causal inertness, from whence no intrusion or interruption on its part need ever be feared". Since the 1950s we have learnt about, designed and built, increasingly complex and causally powerful virtual machines that are fully implemented in physical machinery but include networks of causation that are not directly translatable into networks of physical causation, not least because we know now that sometimes replacing physical components with new ones can provide better support for the old, still running, virtual machinery (e.g. better because the physical components are smaller, cheaper, more energy efficient, faster, more reliable, able to cope with a short term power failure, etc.).

For more on this see the distinction between "atomic state functionalism" and "virtual machine functionalism" in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html.

- Taking into account features of sexual reproduction, symbiosis, and structured individual development, implies that evolutionary *branching trees* need to be replaced by more complex *tangled networks*, as in the theory of *symbiogenesis* promoted by Lynn Margulis and others. This potentially has implications for varieties of types of consciousness, some of which may combine different evolutionary heritages.

  A less obvious consequence of an evolutionary view of consciousness is that changes between the very earliest and most recent forms of consciousness descended from them cannot all be continuous.

  There may be some continuous changes between stages in individual development (e.g. height, muscle strength), but there is no continuously developing individual that started as an ancient microscopic organism and is now an elephant or a human being: biological reproduction in millions of species, including humans must be discrete because between any current individual and any of its ancient ancestors there can be only a finite number of intervening ancestors.

  Some of the discrete changes happen during staggered gene expression in individuals, using meta-configured genomes http://www.cs.bham.ac.uk/research/projects/cogaff/movies/meta-config/ and in a multitude of types of chemical construction during individual development and physical functioning. Molecular bonds exist or don't exist: there is not a continuum of intermediate states (a feature that was emphasized by Schrödinger in *What is life?*(1944)). (In contrast, Newtonian physics is incapable of explaining how new stable structures can come into existence.)[This needs more discussion.]

- Adding multi-layered, temporally staggered, gene expression, with later stages using generic structures instantiated on the basis of environmental parameters acquired earlier, implies that *unitary* data-driven mechanisms are biologically inaccurate: instead, different abstractly specified mechanisms are used at different stages of gene-expression, using parameters acquired during earlier stages. E.g. brains of young mammals are not born with a fixed learning engine and a vast empty memory.
- This is *obviously* true for language development in humans, in view of the multi-layered diversity of human languages, with differences at every level of structure, each level based partly on some genetic influence, using delayed gene-expression, and partly on environmental influences, e.g. local linguistic usage. Recognition that this applies to other competences, seems to be rare.
- Similarly, consciousness can have different contents, functions, and mechanisms at different stages of development. Where each stage combines abstract structures influenced by newly expressed genetic information, with information acquired from the environment at earlier stages (some of it used to provide parameters for the later abstractions). [Including cultural influences such as having a mathematical education that includes Euclidean geometry!]
- (That is the "meta-configured genome" thesis summarised here.)
- Some of the mechanisms, in some species (including squirrels, crows, apes, elephants, etc.), provide abilities to detect and make use of *impossibilities* (e.g. a rigid object with maximum

diameter larger than a fixed width gap cannot pass through the gap in all orientations).

In some cases an impossibility is more naturally described as a necessity --where being false is impossible-- e.g. containment is necessarily transitive: It is impossible for A to contain B and B to contain C, at the same time as A does not contain C.

- As Immanuel Kant noted (in CPR 1781), many such mathematical insights cannot be based entirely on empirical generalisation, or on logical deductions from definitions. Other mechanisms are required, though he thought humans may not be clever enough to think of them -- even though they use them! (I think this is related to Turing's comments about the differences between mathematical intuition, and mathematical ingenuity in 1936.)
- (Some of these ideas recur in claims by Roger Penrose: he and I use related examples, in my case influenced originally by Kant, but we derive different conclusions from the examples.)
- The ability to *detect* and *use* spatial/mathematical (including topological) impossibilities and necessities does not necessarily provide abilities to *reflect on*, *reason about*, *communicate* or *teach* what has been learnt. (Compare apes, squirrels, crows, human toddlers, etc.)
- Additional evolutionary/development processes are required for expert teaching, extending cognitive architectures with meta-cognitive mechanisms, including other-directed meta-cognition used by good educators.
- In humans, ill-understood additional mechanisms provide reflective mathematical consciousness (a kind of self-teaching ability), as in ancient geometers: Appolonius, Archimedes, Euclid, Zeno, etc.
- (Unfortunately too many schools no longer teach Euclidean geometry, including finding constructions and proofs, so most graduates, including science, engineering, and philosophy graduates, now fail to understand most of this. Some Eastern European countries seem to be exceptions.)
- At present, I don't believe any currently proposed mechanism in psychology, neuroscience, philosophy or AI is capable of explaining these capabilities for mathematical discovery.
- I suspect (and I think Turing suspected) that the mechanisms required cannot be provided either by current, digital, logic-based forms of computation, or by neural statistics-based learning mechanisms. The search for an alternative may have motivated his interest in chemistry-based morphogenesis (1952). [There is also one unexplained sentence about chemistry in his 1950 paper, quoted below.]
- Could sub-neural molecular computational mechanisms with their combinations of continuous and discrete processes, emphasised by Schrodinger in *What is life?* (1944), provide the missing explanations -- with disastrous implications for currently fashionable AI theories and mechanisms (including both "symbolic" logic-based or rule-based mechanisms, and statistics-based neural mechanisms)?
- Perhaps these questions explain why Turing wrote in his 1950 paper: "In the nervous system chemical phenomena are at least as important as electrical". Unfortunately he gave no explanation. I suspect von Neumann had similar thoughts towards the end of writing his unfinished, posthumously published book, *The Computer and the Brain*. (Silliman memorial lectures. 1958. I don't know whether he had read Turing 1952.)
- Working out these ideas in more detail may require younger, more competent, brains than mine. But perhaps this framework will motivate and help them?

Branching evolved varieties of consciousness

The key idea of this document is that asking the question "What is consciousness?" while

expecting a short answer is as misguided as "What is an animal?", which cannot be answered adequately without discussing forms of life on this planet and how animals are distinguished and can differ in many different ways, both from non-animal organisms (e.g. plants) and from other animals, e.g. fleas compared with elephants.

Part of the answer to "What is an animal?" must include the evolution of varieties of animals of different sub-types. Similar remarks can be made about varieties of plants, varieties of insects, and varieties of much simpler, even single-celled, organisms that need to select between different actions on the basis of information about their own states and/or information about the immediate or remote environment (e.g. based on chemicals originating from remote sources -- including their chemical structure, their density, their rates of flow, and their direction of flow).

> Exercise for the reader:
> How do those processes differ from the influences and behaviours of a ball rolling down a helical slide: https://en.wikipedia.org/wiki/Helter_skelter_(ride)

Some of the features of the question "What is an animal?" are also features of the question "What is consciousness?" for the same reason: consciousness takes many forms produced over millions of years by evolution, and also changes in complexity of functions and mechanisms during development of individuals, partly related to changes in physical size, shape, needs, and capabilities -- as discussed in more detail below, where I'll talk about branching forms of biological consciousness, criticising theories of consciousness that build on simplistic fashionable definitions or slogans (e.g. "What it is like, or feels like to ..." Nagel (1974), thereby ignoring the great variety of types and functions of consciousness in different products of biological evolution and at different stages of development within individual conscious organisms.

I am not making an original point. One of the invited speakers at the Oxford conference, Adrian Kent, mentioned an important article by William James (1879) that I had not previously encountered. William James expressed an understanding of the need for the kind of analysis of consciousness summarised here that is either present or absent in an organism (or machine). In that article, James presents consciousness not as some unique, simply defined, phenomenon, but rather as a variety of products of biological evolution, changing and growing in complexity as a result of evolution -- to which we should add: and also as a result of individual development. A six year old child is capable of much richer and more varied forms of consciousness than a six month old child. There must also be differences between forms of consciousness in different species, partly because of different mechanisms used (e.g. vision *vs* sonar) but also because of the types of information acquired and different ways in which it is used. (A blind person is likely to use sound in richer ways than a sighted person -- perhaps sharing more with bat consciousness.)

In 1879 James could not have thought about varieties of information processing architectures based on combinations of increasingly sophisticated forms and functions of information processing produced by biological evolution, since many of the concepts now available for describing those functions did not exist at that time -- e.g. the notion of the meaning of a sentence or phrase being tree- or graph-structured. (I wonder whether James had read and been inspired by the reflections of Ada Lovelace (1842) on the powers of computation?)

Only since about 1965 (perhaps later) have we been able to think about virtual machines consisting of multiple virtual machines interacting with one another and the environment -- such as the world wide web, or the email system now used every day by millions of people, both running on a

constantly changing interconnected network of physical machines linked to human individuals and organisations. Understanding of the complexities and diversity of mechanisms and functions of virtual machines has grown enormously in recent decades, mostly unnoticed by philosophers, psychologists and neuroscientists. (Ada Lovelace partly anticipated some of the features of virtual machines, however. For more details on powers of virtual machinery see http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html)

Background notes added Oct/Nov 2019:
The conference presentation and additional notes are related to a collection of publications and discussion notes related to my 1962 DPhil thesis defending Immanuel Kant's philosophy of mathematics (which is widely, but mistakenly, believed to have been refuted by Einstein and Eddington), and various publications contributing to AI and its relations to philosophy and psychology, e.g. including explaining (in 1981) why intelligent robots will (inevitably) have emotions http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#36).

This was part of the Cognition and Affect (CogAff) project, begun at Sussex University in the 1980s, and extended at Birmingham University (UK) since 1991 http://www.cs.bham.ac.uk/research/projects/cogaff/.

There are also some partly related Youtube videos on philosophy, AI, evolution, and the nature of mathematical consciousness:
https://www.youtube.com/playlist?list=PLYC-dSIlAaYa6Mk1g6hBGUyqCwrIvyOWB

The Alan Turing centenary, 2012, triggered a new direction for the CogAff project, to include the new Turing-inspired Meta-Morphogenesis project summarised in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html (also PDF).

This significantly extended the CogAff project to include a survey of evolution of forms of information processing since the simplest life forms.

Some of the ideas in my talk at the Models of Consciousness conference overlap with a presentation summarising *The Meta-Configured Genome* (MCG) hypothesis, developed with Jackie Chappell, available with additional comments here http://www.cs.bham.ac.uk/research/projects/cogaff/movies/meta-config/

Varieties of biological consciousness

The key implication, which I did not have time to present adequately at the Oxford conference is that whereas many or most theories of mind assume that minds start up (at or shortly before birth) armed with some powerful form of learning machine which goes on acquiring (and reorganising) information throughout life, the MCG hypothesis proposes instead that biological evolution has provided, in humans and many other species, a collection of increasingly sophisticated and abstract (i.e. partly specified) forms of learning, that evolved at different times and which are encoded in the genome but not all activated at or soon after birth.

Instead, activation is staggered: the newer, more abstract evolved forms of learning are activated at later stages of development and what they do is not fully specified in the genome, but depends on parameters derived from what has been learnt during earlier stages of development (which is why I call these genetic mechanisms and competences "meta-configured").

What is learnt during early stages of development can vary widely between individuals in the same species, depending on geographical, physical, social, linguistic, technological, etc. features of the current environment -- partly a product of accumulated effects of genome expression in ancestors and other predecessors, who unwittingly or deliberately alter the learning environments of individuals born later with the same genome.

Therefore, what is derived when later stages of the genome are activated can vary even more widely. (For example, linguistic structures in a language, e.g. words, phrases, sentences, paragraphs, stories, theories, etc. vary more widely as they become more complex, and are less likely to be used, so that beyond a certain size they are almost all unique -- for instance I am confident that this parenthetical comment has never been produced previously.)

One of the more obvious results of the MCG mechanism is that humans with (roughly) the same genome at birth can end up speaking languages that vary widely in their physical sounds, vocabulary, grammatical structures, semantic powers and communicative functions. Some communities even make use mostly of sign languages (which almost certainly first evolved before spoken languages, yet can have similar expressive power). The same "meta-configured" genetic mechanisms can account for huge variations in tool uses, tools, mathematical achievements, technology, tastes in food, science, culture, art, belief systems, etc.

Ancient mathematical discoveries

Aspects of the genome concerned with perceiving, understanding, creating, transforming and using spatial structures and processes, when combined with meta-cognitive abilities to reflect on one's own reasoning, errors, corrections, etc. may have provided the foundation for the amazing ancient mathematical discoveries in geometry, topology and arithmetic begun thousands of years ago, long before the use of logical formalisms and explicit axiom systems.

(It is not widely appreciated that the number competences underlying arithmetic depend on a grasp of mathematical properties of the one-to-one correspondence relation between sets of items, especially the fact that the relation is both transitive and symmetric, which is what makes it possible to use number names as surrogates for elements of any collection of items. It seems that children do not grasp the properties of these structures before they are five or six years old, and any evidence that they have innate number competences, or number competences that develop much earlier is seriously mistaken: those earlier competences could be much simpler pattern recognition competences, for example.

Many of the deep ancient mathematical discoveries are still in constant use (e.g. Pythagoras' theorem). As Kant noted, those mathematical discoveries are concerned with impossibilities and necessary consequences, which cannot be derived from statistical evidence and probabilistic reasoning by neural nets (however deep the nets are!).

In fact neural nets that derive probabilities from statistical evidence cannot even represent necessity or impossibility, let alone identify cases.

And since the formalisms and mechanisms of logic-based theorem proving had not yet been discovered when ancient geometrical discoveries were first made, and there's no evidence that modern logical mechanisms existed in ancient brains, the ancient mathematics could not have been based on logic. Instead, logic is a special case of a more general class of structure-based forms of reasoning, that we seem to share with other intelligent species (e.g. squirrels, apes,

elephants, crows, and others), though only humans seem to have the ability to notice what they are doing, reflect on it, generalise it, discuss it with conspecifics and teach younger individuals about it, so that later generations go further than their ancestors because their mathematical learning is accelerated. Current AI systems lack these kinds of functionality, and, as far as I know, psychologists and neuroscientists have no idea what brain mechanisms provide them. (E.g. could it be sub-neural chemistry?)

I don't know of any theory of consciousness that takes proper account of the forms of consciousness involved in such ancient mathematical discoveries, but I suggest that any adequate theory of mathematical consciousness must explain the role of Meta-Configured genomes in production of such forms of consciousness.

BRANCHING EVOLVED FORMS OF CONSCIOUSNESS

A consequence of this is that instead of "What is consciousness?" having a single answer it will need a collection of answers of varying complexity describing and explaining the steadily increasing and varying forms of consciousness produced in different organisms -- including some evolving in parallel, with partly overlapping forms of consciousness, e.g. because consciousness of spatial structures and processes is so fundamental to the intelligence of many different species that interact with, move around in, and manipulate their spatial environment, though only humans seem to have developed reflective forms of meta-consciousness that not only allow individuals to debug and improve their own modes of reasoning, but also allow them to think about the strengths and weaknesses of abilities in conspecifics (and others) and in some cases help to accelerate progress in others, both implicitly through provision of playthings, challenges and cooperative tasks, and explicitly by using linguistic and diagrammatic forms of communication that can hugely accelerate rates of progress in others, e.g. through explicit schooling -- though current educational fashions can seriously interfere with those ancient mechanisms.

Another corollary of the above view of evolution is that recent fashionable attempts to define consciousness using phrases similar to "What it's like to be..." "What it feels like to be..." are shallow and seriously misguided, and divert attention from the richness and diversity of what needs to be explained, or modelled.

Instead of such simple slogans, we need a concept that accommodates the constantly branching (and sometimes merging -- e.g. through sexual reproduction) trajectories in evolution of forms of biological information processing forming the kind of tree or network shown in familiar diagrammatic tree-like summaries of the variety of products of biological evolution, but enhanced with further branching structures showing how, in more complex species the forms of consciousness instead of being fixed at birth can develop and change during the life of an individual, forming yet more trees or networks of possibilities growing out of branches in the standard depictions of evolution's products -- e.g. https://en.wikipedia.org/wiki/Tree_of_life_(biology).

Evolved (parametrised) designs or mechanisms can be combined in novel ways through sexual reproduction or symbiotic relationships that merge different (partly overlapping) evolved designs. In such cases, evolutionary trees are extended to become evolutionary networks.

A common subgoal of much of this work is investigation of uses of information, varieties of forms of information, and varieties of information-processing mechanism produced, until recently, by biological evolution and and/or by individual or social development. Those processes have been

enormously enhanced by rapidly growing uses of forms and mechanisms of information processing produced by human science and technology, especially since the 20th century.

The relevant concept of information here is not Shannon's syntactic measure, but the much older concept of useful semantic information content well understood, for example, by the novelist Jane Austen writing over a century before Shannon, as explained here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/austen-info.html (also PDF)

At the Oxford conference there was partial overlap between my presentation and parts of the presentation of Xerxes Arsiwallah https://www.youtube.com/watch?v=0ImBKHuORqg

Note on required biological mechanisms
The evolutionary and developmental processes sketched here and in related documents depend on physical/chemical features of DNA molecules and the mechanisms that make use of DNA during processes of reproduction, development, maintenance, and normal functioning of cells in organisms of many kinds. A grossly oversimplified, but possibly useful, introduction to some of that complexity can be found in:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/dna-uses.html
Evolution's Uses of DNA (part of a discussion of genome replication/gene expression).

SOME REFERENCES TO BACKGROUND MATERIAL
(To be extended)

Jackie Chappell and Aaron Sloman, (2007). Natural and artificial meta-configured altricial information-processing systems. *International Journal of Unconventional Computing,* 3(3), 211-239.
http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#717
Updated version: The Meta-Configured Genome (work in progress)
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-configured-genome.html

William James, 1879, Are we automata?, in *MIND, A Quarterly Review of Psychology and Philosophy*, pp. 1--22,

Ada Lovelace, 1842, Notes upon the Memoir by the Translator of ''Sketch of The Analytical Engine Invented by Charles Babbage'' By L. F. Menabrea, http://www.fourmilab.ch/babbage/sketch.html

T. Nagel (1974), What is it like to be a bat?, *Philosophical Review*, 83, Oct. 1974, pp. 435--50, Duke Univ. Press, http://dx.doi.org/10.2307/2183914

Erwin Schrödinger, 1944, *What is life?,* CUP, Cambridge.
I have a partial transcript with extended discussion here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/schrodinger-life.html

A. Sloman, 1962, *Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth* (DPhil Thesis), Oxford University. (A defence of Immanuel Kant's philosophy of mathematics, before I had learnt about computing and AI. Now online, with some explanatory comments).
http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-1962

A. Sloman, 1978 *The Computer Revolution in Philosophy,* Harvester Press (and Humanities Press), Hassocks, Sussex. (Revised, extended version online.)
http://www.cs.bham.ac.uk/research/cogaff/62-80.html#crp

Alan Turing's work inspired the meta-morphogenesis project, begun during the Turing Centenary. This document is a part of that project:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html (or pdf),

---

Perhaps there will be more content here by March 2020 after a submission deadline for post conference proceedings!