Construction Kits for Biological Evolution
Aaron Sloman

Pre-publication version Dec 2016

Published by Springer, 2017
Invited contribution to: *The Incomputable: Journeys Beyond the Turing Barrier*
Eds. Mariya Soskova and S Barry Cooper
http://www.springer.com/gb/book/9783319436678
Ongoing work on evolved construction kits reported here:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/construction-kits.html

Abstract This is part of the Turing-inspired Meta-Morphogenesis project, which aims to identify transitions in information processing since the earliest proto-organisms, in order to provide new understanding of varieties of biological intelligence, including the mathematical intelligence that produced Euclid's *Elements*. (Explaining evolution of mathematicians is much harder than explaining evolution of consciousness!) Transitions depend on "construction kits", including the initial "Fundamental Construction Kit" (FCK) based on physics, and Derived Construction Kits (DCKs) produced by evolution, development, learning and culture. Some construction kits (e.g. Lego, Meccano, plasticine, sand) are *concrete*: using physical components and relationships. Others (e.g. grammars, proof systems and programming languages) are *abstract*: producing abstract entities, e.g. sentences, proofs, and new abstract construction kits. Mixtures of the two are *hybrid* kits. Some are meta-construction kits: they are able to create, modify or combine construction kits. Construction kits are generative: they explain sets of possible construction processes and possible products, with mathematical properties and limitations that are mathematical consequences of properties of the kit and its environment. Evolution and development both make new construction kits possible. Study of the FCK and DCKs can lead us to new answers to old questions, e.g. about the nature of mathematics, language, mind, science, and life, exposing deep connections between science and metaphysics. Showing how the FCK makes its derivatives, including all the processes and products of natural selection, possible is a challenge for science and philosophy. This is a long-term research programme with a good chance of being progressive in the sense of Lakatos. Later, this may explain how to overcome serious current limitations of AI (artificial intelligence), robotics, neuroscience and psychology.

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

1 Background: What is Science? Beyond Popper and Lakatos

How is it possible for very varied forms of life to evolve from lifeless matter, including a mathematical species able to make the discoveries presented in Euclid's *Elements*?[1] Explaining evolution of mathematical insight is much harder than explaining evolution of consciousness! (Even insects must be conscious of aspects of their surroundings.) An outline answer is based on construction kits that make other things (including new construction kits) possible. The need for science to include theories that explain how something is possible has not been widely acknowledged. Explaining how X is possible (e.g. how humans playing chess can produce a certain board configuration) need not provide a basis for *predicting* when X will be realised, so the theory used cannot be falsified by non-occurrence. Popper [1934] labelled such theories "non-scientific" – at best metaphysics. His falsifiability criterion has been blindly followed by many scientists who ignore the history of science. E.g. the ancient atomic theory of matter was not falsifiable, but was an early example of a deep scientific theory. Later, Popper shifted his ground, e.g. in his [1978], and expressed great admiration for Darwin's theory of Natural Selection, despite its unfalsifiability.

Lakatos [1980] extended Popper's philosophy of science, showing how to evaluate competing scientific research programmes over time, according to their progress. He offered criteria for distinguishing "progressive" from "degenerating" research programmes, on the basis of their patterns of development, e.g. whether they systematically generate questions that lead to new empirical discoveries and new applications. It is not clear to me whether he understood that his distinction could also be applied to theories explaining how something is possible. Chapter 2 of Sloman [1978][2] modified ideas of Popper and Lakatos to accommodate scientific theories about what is *possible*, e.g. types of plant, types of animal, types of reproduction, types of consciousness, types of thinking, types of learning, types of communication, types of molecule, types of chemical interaction, and types of biological information processing. It presented criteria for evaluating theories of what is possible and how it is possible, including theories that straddle science and metaphysics. Insisting on sharp boundaries between science and metaphysics harms both. Each can be pursued with rigour and openness to specific kinds of criticism. A separate paper[3] includes a section entitled "Why allowing non-falsifiable theories doesn't make science soft and mushy", and discusses the general concept of "explaining possibilities", its importance in science, the criteria for evaluating such explanations, and how this notion conflicts with the falsifiability requirement for scientific theories. Further examples are in Sloman [1996a]. The extremely ambitious Turing-inspired Meta-

---

1 http://www.gutenberg.org/ebooks/21076

2 http://www.cs.bham.ac.uk/research/projects/cogaff/crp/#chap2

3 http://www.cs.bham.ac.uk/research/projects/cogaff/misc/explaining-possibility.html

Morphogenesis project, proposed in Sloman [2013a][4], depends on these ideas, and will be a test of their fruitfulness, in a combination of metaphysics and science.

This paper, straddling science and metaphysics, asks: *How is it possible for natural selection, starting on a lifeless planet, to produce billions of enormously varied organisms, in environments of many kinds, including mathematicians able to discover and prove geometrical theorems?* An outline answer is presented in terms of *construction kits*: the Fundamental (physical) Construction Kit (the FCK), and a variety of "concrete", "abstract" and "hybrid" Derived Construction Kits (DCKs), that together are conjectured to explain how evolution is possible, including evolution of mathematicians. The FCK and its relations to DCKs are crudely depicted later in Sect. 2.1.

Inspired by Kant's ideas in [1781], construction kits are also offered as providing Biological/Evolutionary foundations for core parts of mathematics, including parts used by evolution (but not consciously, of course) long before there were human mathematicians.

Note on "Making Possible": "X makes Y possible" as used here does not imply that if X does not exist then Y is impossible, only that *one* route to existence of Y is via X. Other things can also make Y possible, e.g., an alternative construction kit. So "makes possible" is a relation of sufficiency, not necessity. The exception is the case where X is the FCK – the *Fundamental Construction Kit* – since all concrete constructions must start from it (in this universe?). If Y is abstract, there need not be something like the FCK from which it must be derived. The space of abstract construction kits may not have a fixed "root". However, the abstract construction kits that can be thought about by physically implemented thinkers may be constrained by a future replacement for the Church-Turing thesis, based on later versions of ideas presented here. Although my questions about explaining possibilities arise in the overlap between philosophy and science [Sloman, 1978, Chap. 2], I am not aware of any philosophical work that explicitly addresses the theses discussed here, though there seem to be examples of potential overlap, e.g. Bennett [2011], Wilson [2015].

2 Fundamental and Derived Construction Kits (FCK, DCKs)

Natural selection alone cannot explain how evolution happens, for it must have options to select from. What sorts of mechanisms can produce options that differ so much in so many ways, allowing evolution to produce microbes, fungi, oaks, elephants, octopuses, crows, new niches, ecosystems, cultures, etc.? Various sorts of construction kit, including evolved/derived construction kits, help to explain the emergence of new options. What explains the possibility of these construction kits? Ultimately, features of fundamental physics, including those emphasised by Schrödinger ([1944]), discussed below. Why did it take so much longer for evolution

[4] Expanded in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html

to produce baboons than bacteria? Not merely because baboons are more complex, but also because evolution had to produce more complex construction kits to make baboon-building possible.

Construction-kits are the “hidden heroes” of evolution. Life as we know it requires construction kits supporting construction of machines with many capabilities, including growing many types of material, many types of mechanism, many types of highly functional bodies, immune systems, digestive systems, repair mechanisms, reproductive machinery, and even mathematicians!

A kit needs more than basic materials. If all the atoms required for making a loaf of bread could somehow be put into a container, no loaf could emerge. Not even the best bread-making machine, with paddle and heater, could produce bread from atoms, since that would require atoms pre-assembled into the right amounts of flour, sugar, yeast, water, etc. Only different, separate, histories can produce the molecules and multi-molecule components, e.g. grains of yeast or flour. Likewise, no fish, reptile, bird, or mammal could be created simply by bringing together enough atoms of all the required sorts; and no machine, not even an intelligent human designer, could assemble a functioning airliner, computer, or skyscraper directly from the required atoms. Why not, and what are the alternatives? We first state the problem of constructing very complex working machines in very general terms and indicate some of the variety of strategies produced by evolution, followed later by conjectured features of a very complex, but still incomplete, explanatory story.

2.1 Combinatorics of Construction Processes

Reliable construction of a living entity requires: appropriate types of matter, machines that manipulate matter, physical assembly processes, stores of energy for use during construction, and usually *information* e.g. about which components to assemble at each stage, how to assemble them, and how to decide in what order to do so. This requires, at every stage, at least: (i) components available for the remaining stages, (ii) mechanisms capable of assembling the components, (iii) mechanisms able to decide what should happen next.

If there are N types of basic component and a task requires an object of type O composed of K basic components, the size of a blind exhaustive search for a sequence of types of basic component to assemble an O is up to N^K sequences, a number that rapidly grows astronomically large as K increases. If, instead of starting from the N types of *basic* component, construction uses M types of *pre-assembled* component, each containing P basic components, then an O will require only K/P pre-assembled parts. The search space for a route to O is reduced in size to $M^{(K/P)}$.

Compare assembling an essay of length 10,000 characters (a) by systematically trying elements of a set of about 30 possible characters (including punctuation and spaces) with (b) choosing from a set of 1000 useful words and phrases, of average length 50 characters. In the first case each choice has 30 options but 10,000 choices are required. In the second case there are 1000 options per choice, but far fewer stages: 200 instead of 10,000 stages. So the size of the (exhaustive) search space is reduced from 30^{10000}, a number with 14,773 digits, to about 1000^{200}, a number with only 602 digits: a very much smaller number. Therefore trying only good pre-built substructures at each stage of a construction process can produce a huge reduction to the search space for solutions of a given size, though some solutions may be missed.

So, learning from experience by storing useful subsequences can achieve dramatic reductions, analogous to a house designer moving from thinking about how to assemble atoms, to thinking about assembling molecules, then bricks, planks, tiles, then pre-manufactured house sections. The reduced search space contains fewer samples from the original possibilities, but the original space has a much larger proportion of useless options. As sizes of pre-designed components increase, so does the variety of pre-designed options to choose from at each step, though far, far, fewer search steps are required for a working solution: a very much shorter evolutionary process. The cost may be exclusion of some design options.

This indicates intuitively, but very crudely, how using increasingly large, already tested useful part-solutions can enormously reduce the search for viable solutions. The technique is familiar to many programmers, in the use of "memo-functions" ("memoization") to reduce computation time, e.g. computing fibonacci numbers. The family of computational search techniques known as "Genetic Programming"[5] makes use of related ideas. The use of "crossover" in evolution (and in Genetic Algorithms), allows parts of each parent's design specification to be used in new combinations.

In biological evolution, instead of previous *solutions* being stored for future re-use, *information about how to build components of previous solutions* is stored in genomes. Evolution, the Great Blind Mathematician, discovered memoization long before we did. A closely related strategy is to record fragments that cannot be useful in certain types of problem, in order to prevent wasteful attempts to use such fragments. Expert mathematicians learn from experience which options are useless (e.g. dividing by zero). This could be described as "negative-memoization". Are innate aversions examples of evolution doing something like that?

Without prior information about useful components and combinations of pre-built components, random assembly processes can be used. If mechanisms are available for recording larger structures that have been found to be useful or useless, the search space for new designs can be shrunk. By doing the searching and experimentation using *information* about how to build things rather than directly recombining the built physical structures themselves, evolution reduces the problem of *recording* what has been learnt.

---

[5] https://en.wikipedia.org/wiki/Genetic_programming

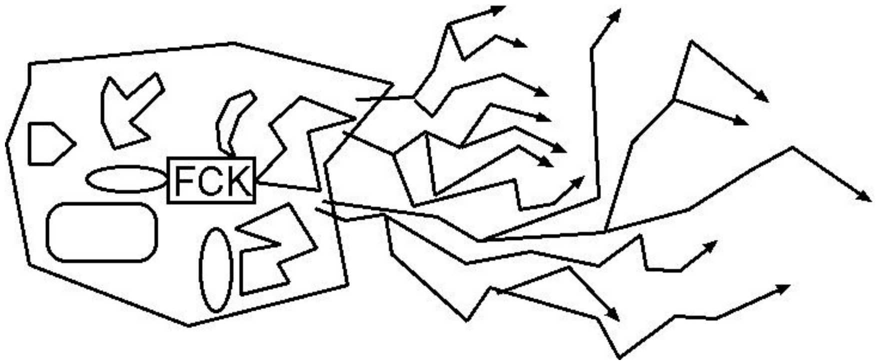

Fig. 1 This is a crude representation of the Fundamental Construction Kit (FCK) (on *left*) and (on *right*) a collection of trajectories from the FCK through the space of possible trajectories to increasingly complex mechanisms.

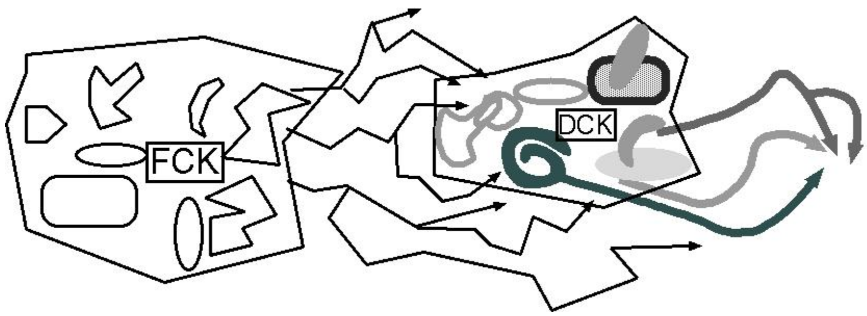

Fig. 2 Further transitions: a fundamental construction kit (FCK) on *left* gives rise to new evolved "derived" construction kits, such as the DCK on the *right*, from which new trajectories can begin, rapidly producing new more complex designs, e.g. organisms with new morphologies and new information-processing mechanisms. The *shapes* and *colours* (crudely) indicate qualitative differences between components of old and new construction kits, and related trajectories. A DCK trajectory uses larger components and is therefore much shorter than the equivalent FCK trajectory.

The Fundamental Construction Kit (FCK) provided by the physical universe made possible all the forms of life that have so far evolved on earth, and also possible, but still unrealised, forms of life, in possible types of physical environment.

Figure 2 shows how a common initial construction kit can generate many possible trajectories, in which components of the kit are assembled to produce new instances (living or non-living). The space of possible trajectories for combining basic constituents is enormous, but routes can be shortened and search spaces shrunk by building derived construction kits (DCKs), that are able to assemble larger structures in fewer steps,[6] as indicated in Fig. 2.

The history of technology, science and engineering includes many transitions in which new construction kits are derived from old ones. That includes the science and technology of digital computation, where new advances used an enormous variety of discoveries and inventions, including punched cards (used in Jacquard looms),

[6] Assembly mechanisms are part of the organism, as illustrated in a video of grass growing itself from seed https://www.youtube.com/watch?v=JbiQtfr6AYk. In mammals with a placenta, more of the assembly process is shared between mother and offspring.

many types of electronic device, many types of programming language, many types of external interface, (not available on Turing machines!), many types of operating system, many types of network connection, and many types of virtual machine, in an enormous variety of applications. Particular inventions were generalised, using mathematical abstractions, to patterns that could be reused in new contexts. New applications frequently led to production of new more powerful tools.

Natural selection did all this on an even larger scale, with far more variety, probably discovering many obscure problems and solutions still unknown to us. (An educational moral: teaching only what has been found most useful can discard future routes to possible major new advances – like depleting a gene pool.)

Biological construction kits derived from the FCK can combine to form new Derived Construction Kits (DCKs), some specified in genomes, and (very much later) some discovered or designed by individuals (e.g. during epigenesis Sect. 2.3), or by groups, for example new languages. Compared with derivation from the FCK, the rough calculations above show how DCKs can enormously speed up searching for new complex entities with new properties and behaviours. See Fig. 2.

DCKs that evolve in different species in different locations may have overlapping functionality, based on different mechanisms: a form of *convergent evolution*. E.g., mechanisms enabling elephants to learn to use trunks, eyes, and brains to manipulate food may share features with those enabling primates to learn to use hands, eyes, and brains to manipulate food. In both cases, competences evolve in response to structurally similar affordances in the environment. This extends Gibson's ideas in [1979] to include affordances for a species, or collection of species.[7]

2.2 Construction Kit Ontologies

A construction kit (and its products) can exist without being described. However, scientists need to use various forms of language in order to describe the entities they observe or postulate in explanations, and to formulate new questions to be answered. So a physicist studying the FCK will need one or more construction kits for defining concepts, formulating questions, formulating theories and conjectures, constructing models, etc. Part of the process of science is extending construction kits for theory formation. Something similar must be done by natural selection: extending useful genetic information structures that store specifications for useful components.

This relates to claims that have been made about requirements for control systems and for scientific theories. For example, if a system is to be capable of distinguishing between N different situations and responding differently to them, it must be capable of being in at least N different states (recognition+control states). This is a variant of Ashby's "Law of Requisite Variety" [1956].

[7] Implications for evolution of vision and language are discussed in http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk111

Many thinkers have discussed representational requirements for scientific theories, or for specifications of designs. Chomsky [1965] identified requirements for theories of language, which he labelled *observational adequacy* (covering the variety of observed uses of a particular language), *descriptive adequacy* (covering the intuitively understood principles that account for the scope of a particular language) and *explanatory adequacy* (providing a basis for explaining how any language can be acquired on the basis of data available to the learner). These labels were vaguely echoed by McCarthy and Hayes [1969], who described a form of representation as *metaphysically* adequate if it can express anything that can be the case, *epistemologically* adequate if it can express anything that can be known by humans and future robots, and *heuristically* adequate if it supports efficient modes of reasoning and problem-solving. (I have simplified all these proposals.)

Requirements can also be specified for powers of biological construction kits. The fundamental construction kit (FCK) must have the power to make any form of life that ever existed or will exist possible, using huge search spaces if necessary. DCKs may meet different requirements, e.g. each supporting fewer types of life form, but enabling those life forms to be "discovered" in a shorter time by natural selection, and replicated (relatively) rapidly. Early DCKs may support the simplest organisms that reproduce by making copies of themselves perhaps as Ganti [2003] described.

At later stages of evolution, DCKs are needed that can construct organisms that change their properties during development and change their control mechanisms appropriately as they grow Thompson [1917]. This requires the ability to produce individuals whose features are *parametrised*, with parameters that change over time. More sophisticated DCKs must be able to produce species with epigenetic mechanisms that modify their knowledge and their behaviours not merely as required to accommodate their own growth but also to cope with changing physical environments, new predators, new prey and new shared knowledge. A special case of this is having genetic mechanisms able to support development of a wide enough range of linguistic competences to match any type of human language, developed in any social or geographical context. However, the phenomenon is far more general than language development, as discussed in the next section.

2.3 Construction Kits Built During Development (Epigenesis)

Some new construction kits are products of evolution of a species and are initially shared among only a few members of the species (barring genetic abnormalities), alongside cross-species construction kits shared between species, such as those used in mechanisms of reproduction and growth in related species. Evolution also discovered the benefits of "meta-construction-kits": mechanisms that allow members of a species to build new construction kits during their own development.

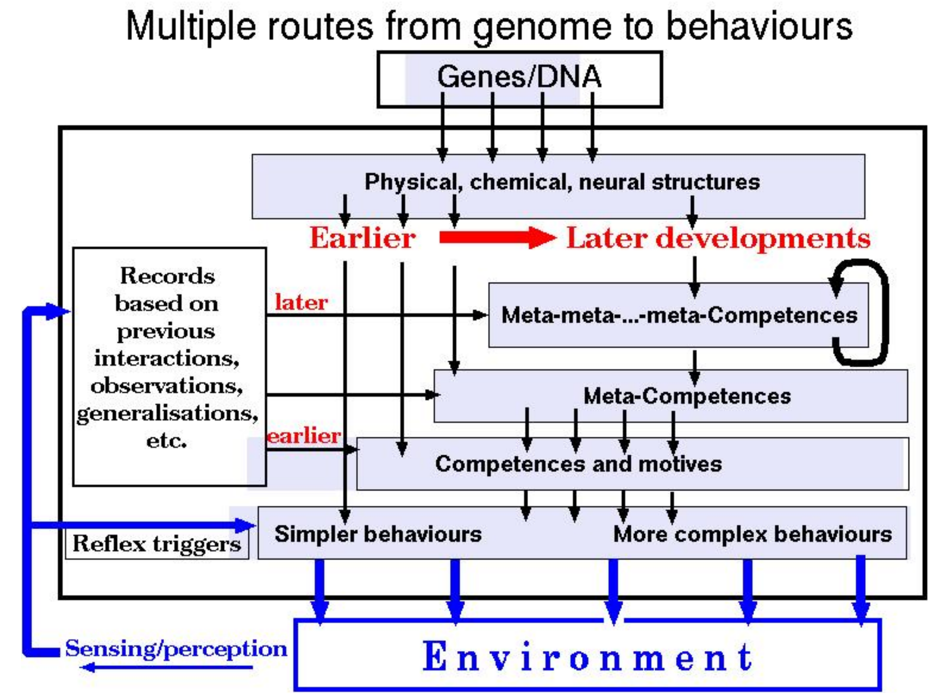

Fig. 3 A construction kit gives rise to very different individuals if the genome interacts with the environment in increasingly complex ways during development. Precocial species use only the downward routes on the *left*, producing preconfigured competences. Competences of altricial species, using staggered development, may be far more varied. Results of using earlier competences interact with the genome, producing meta-configured competences on the *right*.

Examples include mechanisms for learning that are initially generic mechanisms shared across individuals, and developed by individuals on the basis of their own previously encountered learning experiences, which may be different in different environments for members of the same species. Human language learning is a striking example: things learnt at earlier stages make new things learnable that might not be learnable by an individual transferred from a different environment part way through learning a different language. This contrast between genetically specified and individually built capabilities for learning and development was labelled a difference between “pre-configured” and “meta-configured” competences in Chappell and Sloman [2007], summarised in Fig. 3.

The meta-configured competences are partly specified in the genome, but that specification is combined with information abstracted from individual experiences. Mathematical development and language development in humans both seem to be special cases of growth of meta-configured competences. Karmiloff-Smith presented closely related ideas in [1992].

Construction kits used for assembly of new organisms that start as seeds or eggs make possible many different processes in which components are assembled in parallel, using abilities of the different sub-processes to constrain one another. Nobody knows the full variety of ways in which parallel construction processes can exercise mutual control in developing organisms. One implication of Fig. 3 is that there are not always simple correlations between genes and organism features.

Explaining the many ways in which a genome can orchestrate parallel processes of growth, development, formation of connections, etc. is a huge challenge. A framework allowing abstract specifications in a genome to interact with details of the environment in instantiating complex designs is illustrated schematically in Fig. 3. An example might be Popper's proposal in [1976] that newly evolved desires of individual organisms (e.g. desires to reach fruit in taller trees) could indirectly and gradually, across generations, influence selection of physical characteristics (e.g. longer necks, abilities to jump higher) that improve success-rates of actions triggered by those desires. Various kinds of creativity, including mathematical creativity, might result from such transitions. This generalises Waddington's [1957], "epigenetic landscape" metaphor, by allowing individual members of a species to partially construct and repeatedly modify their own epigenetic landscapes instead of merely following paths in a landscape that is common to the species. Mechanisms that increase developmental variability may also make new developmental defects possible (e.g. autism?).

2.4 The Variety of Biological Construction Kits

As products of physical construction kits become more complex, with more ways of contributing to needs of organisms, and directly or indirectly to reproductive fitness, they require increasingly sophisticated control mechanisms. New sorts of control often use new types of information. Processing that information may require new mechanisms. That may require new construction kits for building new types of information-processing mechanism. The simplest organisms use only a few types of (mainly chemical) sensor, providing information about internal physical and chemical states and the immediate external physical environment. They have very few behavioural options. They acquire, use and replace fragments of information, using the same forms of information throughout their life, to control deployment of a fixed repertoire of capabilities.

More complex organisms acquire information about enduring spatial locations in extended terrain, including static and changing routes between static and changing resources and dangers. They need to construct and use far more complex (internal or external) information stores about their environment, and, in some cases, "meta-semantic" information about information-processing, in themselves and in others, e.g. conspecifics, predators and prey.

What forms can such information take? Many controlled systems have states that can be represented by a fixed set of physical measures, often referred to as "variables", representing states of sensors, output signals, and internal states of various sorts. Relationships between such state-components are often represented mathematically by equations, including differential equations, and constraints (e.g. inequalities) specifying restricted, possibly time-varying, ranges of values for the variables, or magnitude relations between the variables. A system with N variables (including derivatives) has a state of a fixed dimension, N. The only way to record new information in such a system is in static or dynamic values for numeric variables — changing "state vectors" — and possibly alterations in the equations. A typical example is W.T.Powers [1973], inspired by Wiener [1961] and Ashby [1952]. There are many well understood special cases, such as simple forms of homeostatic control using negative feedback. Neural net based controllers often use large numbers of variables clustered into strongly interacting sub-groups, groups of groups, etc.

For many structures and processes, a set of numerical values and rates of change linked by equations (including differential equations) expressing their changing relationships is an adequate form of representation, but not for all, as implied by the discussion of types of *adequacy* in Sect. 2.2. That's why chemists use *structural* formulae, e.g. diagrams showing different sorts of bonds between atoms and collections of diagrams showing how bonds change in chemical reactions. Linguists, programmers, computer scientists, architects, structural engineers, map makers and users, mathematicians studying geometry and topology, composers, and many others, work in domains where structural diagrams, logical expressions, grammars, programming languages, plan formalisms, and other *non-numerical* notations express information about structures and processes that is not usefully expressed in terms of collections of numbers and equations linking numbers.[8]

Of course, any information that can be expressed in 2-D written or printed notation, such as grammatical rules, parse trees, logical proofs, and computer programs, can also be converted into a large array of numbers by taking a photograph and digitising it. Although such processes are useful for storing or transmitting documents, they add so much irrelevant numerical detail that the original functions are obstructed, such as checking whether an inference is valid, manipulating a grammatical structure by transforming an active sentence to a passive one, determining whether two sentences have the same grammatical subject, removing a bug from a program, or checking whether a geometric construction proves a theorem — unless the original non-numerical structures are extracted, often at high cost.

Similarly, collections of numerical values will not always adequately represent information that is biologically useful for animal decision making, problem solving, motive formation, learning, etc. Moreover, biological sensors are poor at acquiring or representing very precise information, and neural states often lack reliability and stability. (Such flaws can be partly compensated for by using many neurons per numerical value and averaging.) More importantly, the biological functions, e.g. of visual systems, may have little use for absolute measures if their functions are based on *relational* information, such as that A is closer to B than to C, A is biting B,

[8] Examples include: https://en.wikipedia.org/wiki/Parse_tree, https://en.wikipedia.org/wiki/Flowchart, https://en.wikipedia.org/wiki/Structural_formula, https://en.wikipedia.org/wiki/Euclidean_geometry, https://en.wikipedia.org/wiki/Entity-relationship_model. https://en.wikipedia.org/wiki/Programming_language

A is keeping B and C apart, A can fit through the gap between B and C, the joint between A and B is non-rigid, A cannot enter B unless it is reoriented, and many more. As Schrödinger [1944] pointed out, topological structures of molecules can reliably encode a wide variety of types of genetic information, and may also turn out to be useful for recording other forms of structural information. Do brains employ them? Chomsky [1965] pointed out that using inappropriate structures in models can divert attention from important biological phenomena that need to be explained–see Sect. 2.2, above. Max Clowes, who introduced me to AI in 1969, made similar points about research in vision around that time.[9] So a subtask for this project is to identify types of non-numerical, e.g. relational, information content that are of biological importance, and the means by which such information can be stored, transmitted, manipulated, and used, and to explain how the mechanisms performing those tasks can be built from the FCK, using appropriate DCKs.

2.5 Increasingly Varied Mathematical Structures

Electronic computers made many new forms of control possible, including use of logic, linguistic formalisms, planning, learning, problem solving, vision, theorem proving, teaching, map-making, automated circuit design, program verification, and many more. The world wide web is an extreme case of a control system made up of millions of constantly changing simpler control systems, interacting in parallel with each other and with millions of display devices, sensors, mechanical controllers, humans, and many other things. The types of control mechanism in computer-based systems now extend far beyond the numerical sorts familiar to control engineers.[10]

Organisms also need multiple control systems, not all numerical. A partially constructed percept, thought, question, plan or terrain description has parts and relationships, to which new components and relationships can be added and others removed as construction proceeds and errors are corrected. So the structures change – unlike a fixed-size collection of variables assigned changing values. Non-numerical types of mathematics are needed for describing or explaining such systems, including topology, geometry, graph theory, set theory, logic, formal grammars, and theory of computation. A full understanding of mechanisms and processes of evolution and development may need new branches of mathematics, including mathematics of non-numerical structural processes, such as chemical change, or changing "grammars" for internal records of complex structured information. The importance of non-numerical information structures has been understood by many mathematicians, logicians, linguists, computer scientists and engineers, but many scientists still focus only on numerical structures and processes. They sometimes seek to remedy failures by using

[9] http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#61

[10] Often misleadingly labelled "non-linear" — like calling apples, apes and avalanches non-bananas! http://en.wikipedia.org/wiki/Control_theory http://en.wikipedia.org/wiki/Nonlinear_control

statistical methods, which can be spectacularly successful in restricted contexts, as shown by recent AI successes, whose limitations I have commented on elsewhere.[11]

The FCK need not be able to produce all biological structures and processes *directly*, in situations without life, but it must be rich enough to support successive generations of increasingly powerful DCKs that together suffice to generate all possible biological organisms evolved so far, and their behavioural and information-processing abilities. Moreover, the FCK, or DCKs derived from it, must include abilities to acquire, manipulate, store, and use information structures in DCKs that can build increasingly complex machines that encode information, including non-numerical information. Since the 1950s we have also increasingly discovered the need for new *virtual* machines as well as *physical* machines Sloman [2010, 2013b].

Large-scale physical processes usually involve a great deal of variability and unpredictability (e.g. weather patterns), and sub-microscopic indeterminacy is a key feature of quantum physics; yet, as Schrödinger [1944] pointed out, life depends on very complex objects built from very large numbers of small-scale structures (molecules) that can preserve their *precise* chemical structure despite continual thermal buffetting and other disturbances. Unlike non-living natural structures, important molecules involved in reproduction and other biological functions are copied repeatedly, predictably transformed with great precision, and used to create very large numbers of new molecules required for life, with great, but not absolute, precision. This is *non-statistical* structure preservation, which would have been incomprehensible without quantum mechanics, as explained by Schrödinger. That feature of the FCK resembles "structure-constraining" properties of construction kits such as Meccano, TinkerToy and Lego[12] that support structures with more or less complex, discretely varied topologies, or kits built from digital electronic components, that also provide extremely reliable preservation and transformations of precise structures, in contrast with sand, water, mud, treacle, plasticine, and similar materials. Fortunate children learn how structure-based kits differ from more or less amorphous construction kits that produce relatively flexible or plastic structures with non-rigid behaviours — as do many large-scale natural phenomena, such as snow-drifts, oceans, and weather systems.

Schrödinger [1944] stressed that quantum mechanisms can explain the structural stability of individual molecules, and how a set of atoms in different arrangements can form discrete stable structures with very different properties (e.g. in propane and isopropane, only the location of the single oxygen atom differs, but that alters both the topology and the chemical properties of the molecule).[13] He also pointed out the relationship between the number of discrete changeable elements and information capacity, anticipating Shannon [1948]. Some complex molecules with quantum-

[11] E.g. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html

[12] https://en.wikipedia.org/wiki/Meccano, https://en.wikipedia.org/wiki/Tinkertoy and https://en.wikipedia.org/wiki/Lego

[13] E.g. see James Ashenhurst's tutorial: http://www.masterorganicchemistry.com/2011/11/10/dont-be-futyl-learn-the-butyls/

based structural stability are simultaneously capable of *continuous* deformations, e.g. folding, twisting, coming together, moving apart, etc., all essential for the role of DNA and other molecules in reproduction, and many other biochemical processes. This combination of discrete topological structure (forms of connectivity), used for storing very precise information for extended periods, and non-discrete spatial flexibility, used in assembling, replicating and extracting information from large structures, is unlike anything found in digital computers, although it can to some extent be approximated in digital computer models of molecular processes.

Highly deterministic, very small-scale, discrete interactions between very complex, multi-stable, enduring molecular structures, combined with continuous deformations (folding, etc.) that alter opportunities for the discrete interactions, may have hitherto unnoticed roles in brain functions, in addition to their profound importance for reproduction and growth. Much recent AI and neuroscience uses statistical properties of complex systems with many continuous scalar quantities changing randomly in parallel, unlike symbolic mechanisms used in logical and symbolic AI, though the latter are still far too restricted to model animal minds. The Meta-Morphogenesis project has extended a set of examples studied four decades earlier (e.g. in Sloman [1978]) of types of mathematical discovery and reasoning that use perceived *possibilities* and *impossibilities* for change in geometrical and topological structures. Further work along these lines may help to reveal biological mechanisms that enabled the great discoveries by Euclid and his predecessors that are still unmatched by AI theorem provers (as discussed in Sect. 5).

2.6 Thermodynamic Issues

The question sometimes arises whether formation of life from non-living matter violates the second law of thermodynamics, because life increases the amount of order or structure in the physical matter on the planet, reducing entropy. The standard answer is that the law is applicable only to closed systems, and the earth is not a closed system, since it is constantly affected by solar and other forms of radiation, asteroid impacts, and other external influences. The law implies only that our planet could not have generated life forms without energy from non-living sources, e.g. the sun (though future technologies may reduce or remove such dependence). Some of the ways in which pre-existing dispositions can harness external sources of energy to increase local structure are discussed in a separate collection of thoughts on entropy, evolution, and construction kits.[14]

---

14 http://www.cs.bham.ac.uk/research/projects/cogaff/misc/entropy-evolution.html This was partly inspired by memories of Lionel Penrose in Oxford around 1960 giving a talk about devices he called *droguli* – singular *drogulus*. Such naturally occurring multi-stable physical structures seem to me to render redundant the apparatus proposed in Deacon [2011] to explain how life apparently goes against the second law of thermodynamics. See https://en.wikipedia.org/wiki/Incomplete_Nature

Our discussion so far suggests that the FCK has two sorts of components: (a) a generic framework including space-time and generic constraints on what can happen in that framework, and (b) components that can be non-uniformly and dynamically distributed in the framework. The combination makes possible formation of galaxies, stars, clouds of dust, planets, asteroids, and many other lifeless entities, as well as supporting forms of life based on derived construction kits (DCKs) that exist only in special conditions. Some local conditions e.g. extremely high pressures, temperatures, and gravitational fields (among others), can mask some parts of the FCK, i.e. prevent them from functioning. So, even if all sub-atomic particles required for earthly life exist at the centre of the sun, local factors can rule out earth-like life forms. Moreover, if the earth had been formed from a cloud of particles containing no carbon, oxygen, nitrogen, iron, etc., then no DCK able to support life as we know it could have emerged, since that requires a region of space-time with a specific manifestation of the FCK, embedded in a larger region that can contribute additional energy (e.g. solar radiation), and possibly other resources.

As the earth formed, new physical conditions created new DCKs that made the earliest life forms possible. Ganti [2003], usefully summarised in Korthof [2003] and Fernando [2008], presents an analysis of requirements for a minimal life form, the "chemoton", with self-maintenance and reproductive capabilities. Perhaps still unknown DCKs made possible formation of pre-biotic chemical structures, and also the *environments* in which a chemoton-like entity could survive and reproduce. Later, conditions changed in ways that supported more complex life forms, e.g. oxygen-breathing forms. Perhaps attempts to identify the first life form in order to show how it could be produced by the FCK are misguided, because several important pre-life construction kits were necessary: i.e. several DCKs made possible by conditions on earth were necessary for precursors. Some of the components of the DCKs may have been more complex than their living products, including components providing *scaffolding* for constructing life forms, rather than materials.

2.7 Scaffolding in Construction Kits

An important feature of some construction kits is that they contain parts that are used during assembly of products of the kit, but are not included in the products. For example, Meccano kits include spanners and screwdrivers, used for manipulating screws and nuts during assembly and disassembly, though they are not normally included in the models constructed. Similarly, kits for making paper dolls and their clothing[15] may include pencils and scissors, used for preparing patterns and cutting them out. But the pencils and scissors are not parts of the dolls or their clothing. When houses are built, many items are used that are not part of the completed house, including tools and scaffolding frameworks to support incomplete structures. A loose

[15] https://en.wikipedia.org/wiki/Paper_doll

analogy can be made with the structures used by climbing plants, e.g. rock faces, trees, or frames provided by humans: these are essential for the plants to grow to the heights they need but are not parts of the plant. More subtly, rooted plants that grow vertically make considerable use of the soil penetrated by their roots to provide not only nutrients but also the stability that makes tall stalks or trunks possible, including in some cases the ability to resist strong winds most of the time. The soil forms part of the scaffolding. A mammal uses parts of its mother as temporary scaffolding while developing in the womb, and continues to use the mother during suckling and later when fed portions of prey caught by parents. Other species use eggs with protective shells and food stores. Plants that depend on insects for fertilization can be thought of as using scaffolding in a general sense.

This concept of scaffolding may be crucial for research into origins of life. As far as I know, nobody has found candidate non-living chemical substances made available by the FCK that have the ability spontaneously to assemble themselves into primitive life forms. It is possible that the search is doomed to fail because there never were such substances. Perhaps the earliest life forms required not only materials but also scaffolding — e.g. in the form of complex molecules that did not form parts of the earliest organisms but played an essential causal role in assembly processes, bringing together the chemicals needed by the simplest organisms. Evolution might then have produced new organisms without that reliance on the original scaffolding. The scaffolding mechanisms might later have ceased to exist on earth, e.g. because they were consumed and wiped out by the new life forms, or because physical conditions changed that prevented their formation but did not destroy the newly independent organisms. A similar suggestion was recently made by Mathis et al. [2015]. So it is quite possible that many evolutionary transitions, including transitions in information processing, our main concern, depended on forms of scaffolding that later did not survive and were no longer needed to maintain what they had helped to produce. So research into evolution of information processing, our main goal, is inherently partly speculative.

2.8 Biological Construction Kits

How did the FCK generate complex life forms? Is the Darwin-Wallace theory of natural selection the whole answer, as suggested in Bell [2008]: *"Living complexity cannot be explained except through selection and does not require any other category of explanation whatsoever"*. No: the explanation must include both *selection* mechanisms and *generative* mechanisms, without which selection processes will not have a supply of new viable options. Moreover, insofar as environments providing opportunities, challenges and threats are part of the selection process, the construction kits used by evolution include mechanisms not intrinsically concerned with life, e.g. volcanoes, earthquakes, asteroid impacts, lunar and solar tides, and many more.

The idea of evolution producing construction kits is not new, though they are often referred to as "toolkits". Coates et al. [2014] ask whether there is "a genetic toolkit for multicellularity" used by complex life-forms. Toolkits and construction kits normally have *users* (e.g. humans or other animals), whereas the construction kits we have been discussing (FCKs and DCKs) do not all need separate users.

Both generative mechanisms and selection mechanisms change during evolution. Natural selection (blindly) uses the initial enabling mechanisms provided by physics and chemistry not only to produce new organisms, but also to produce new richer DCKs, including increasingly complex information-processing mechanisms. Since the mid 1900s, spectacular changes have also occurred in human-designed computing mechanisms, including new forms of hardware, new forms of virtual machinery, and networked social systems all unimagined by early hardware designers. Similar changes during evolution produced new biological construction kits, e.g. grammars, planners, and geometrical constructors, not well understood by thinkers familiar only with physics, chemistry and numerical mathematics.

Biological DCKs produce not only a huge variety of physical forms and physical behaviours, but also forms of *information-processing* required for increasingly complex control problems, as organisms become more complex and more intelligent in coping with their environments, including interacting with predators, prey, mates, offspring, conspecifics, etc. In humans, that includes abilities to form scientific theories and discover and prove theorems in topology and geometry, some of which are also used unwittingly in practical activities.[16] I suspect many animals come close to this in their *systematic* but unconscious abilities to perform complex actions that use mathematical features of environments. Abilities used unconsciously in building nests or in hunting and consuming prey may overlap with topological and geometrical competences of human mathematicians. (See Sect. 6.2.) For example, search for videos of weaver birds building nests.

3 Concrete (Physical), Abstract and Hybrid Construction Kits

Products of a construction kit may be concrete, i.e. physical; or abstract, like a theorem, a sentence, or a symphony; or hybrid, e.g. a written presentation of a theorem or poem.

Concrete Kits Construction kits for children include physical parts that can be combined in various ways to produce new physical objects that are not only larger than the initial components but also have new shapes and new behaviours. Those are *concrete* construction kits. The FCK is (arguably?) a concrete construction kit.

[16] Such as putting a shirt on a child: http://www.cs.bham.ac.uk/research/projects/cogaff/misc/shirt.html I think Piaget noticed some of the requirements.

Lego, Meccano, twigs, mud, and stones, can all be used in construction kits whose constructs are physical objects occupying space and time: *concrete* construction kits.

Abstract Kits There are also non-spatial *abstract* construction kits, for example components of languages, such as vocabulary and grammar, or methods of construction of arguments or proofs. Physical *representations* of such things, however, can occupy space and/or time, e.g. a spoken or written sentence, a diagram, or a proof presented on paper. Using an abstract construction kit, e.g. doing mental arithmetic or composing poetry in your head, requires use of one or more physical construction kits, directly or indirectly implementing features of the abstract kit.

There are (deeply confused) fashions emphasising "embodied cognition" and "symbol grounding" (previously known as "concept empiricism" and demolished by Immanuel Kant and twentieth century philosophers of science). These fashions disregard many examples of thinking, perceiving, reasoning and planning that require abstract construction kits. For example, planning a journey to a conference does not require physically trying possible actions, like water finding a route to the sea. Instead, you may use an abstract construction kit able to *represent* possible options and ways of combining them. Being able to talk requires use of a grammar specifying abstract structures that can be assembled using a collection of grammatical relationships, in order to form new abstract structures with new properties relevant to various tasks involving information. Sentences allowed by a grammar for English are abstract objects that can be instantiated physically in written text, printed text, spoken sounds, Morse code, etc.; so a grammar is an abstract construction kit whose constructs can have concrete (physical) instances. The idea of a grammar is not restricted to verbal forms: it can be extended to many complex structures, e.g. grammars for sign languages, circuit diagrams, maps, proofs, architectural layouts and even molecules.

A grammar does not fully specify a language: a structurally related *semantic* construction kit is required for building possible *meanings*. Use of a language depends on language users, for which more complex construction kits are required, including products of evolution, development and learning. Evolution of various types of language, including languages used only *internally*, is discussed in Sloman [2015].

In computers, digital circuitry implements abstract construction kits via intermediate abstract kits—virtual machines—presumably also required in brains.

Hybrid Abstract+Concrete kits These are combinations, e.g. physical chess board and chess pieces combined with the rules of chess, lines and circular arcs on a physical surface instantiating Euclidean geometry, puzzles like the mutilated chess-board puzzle, and many more. A particularly interesting hybrid case is the use of physical objects (e.g. blocks) to instantiate arithmetic, which may lead to the discovery of prime numbers when certain attempts at rearrangement fail—and an explanation of the impossibility is found. [17]

In some hybrid construction kits such as games like chess, the concrete (physical) component may be redundant for some players, e.g. chess experts who can play without physical pieces on a board. But communication of moves needs physical mechanisms, as does the expert's brain (in ways that are not yet understood). Related abstract structures, states and processes can also be implemented in computers, which can now play chess better than most humans, without replicating human brain mechanisms. In contrast, physical components are indispensable in hybrid construction kits for outdoor games, such as cricket Wilson [2015]. (I don't expect to see good robot cricketers soon.)

Physical computers, programming languages, operating systems and virtual machines form hybrid construction kits that can make things happen when they run. A logical system with axioms and inference rules can be thought of as an abstract kit supporting construction of logical proof-sequences, usually combined with a physical notation for written proofs. A purely logical system cannot have physical causal powers, whereas its concrete instances can, e.g. teaching a student to distinguish valid from invalid proofs. Natural selection "discovered" the power of hybrid construction kits using virtual machinery long before human engineers did. In particular, biological virtual machines used by animal minds outperform current engineering designs in some ways, but they also generate much confusion in the minds of philosophical individuals who are aware that something more than purely physical machinery is at work, but don't yet understand how to implement virtual machines in physical machines Sloman and Chrisley [2003], Sloman [2010, 2013b].

Animal perception, learning, reasoning, and intelligent behaviour require *hybrid* construction kits. Scientific study of such kits is still in its infancy. Work done so far on the Meta-Morphogenesis project suggests that natural selection "discovered" and used a staggering variety of types of hybrid construction kit that were essential for reproduction, for developmental processes (including physical development and learning), for performing complex behaviours, and for social/cultural phenomena.

3.1 Kits Providing External Sensors and Motors

Some construction kits can be used to make toys with moving parts, e.g. wheels or grippers, that interact with the environment. A toy car may include a spring, whose potential energy can be transformed into mechanical energy via gears, axles and wheels in contact with external surfaces. Further interactions, altering the direction

[17] A possibility discussed in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/toddler-theorems.html#primes Contributions from observant parents and child-minders are welcome. Much deeper insights come from extended individual developmental trajectories than from statistics of snapshots of many individuals.

of motion, may result from collisions with fixed or mobile objects in the environment or from influence of some control device.

As noted in [Sloman, 1978, Chap. 6], the distinction between internal and external components is often arbitrary. For example, a music box may play a tune under the control of a rotating disc with holes or spikes. The disc can be thought of either as part of the music box or as part of a changing environment.

If a toy train set has rails or tracks used to guide the motion of the train, then the wheels can be thought of as sensing the environment and causing changes of direction. This is partly like and partly unlike a toy vehicle that uses an optical sensor linked to a steering mechanism, so that the vehicle can follow a line painted on a surface. The railway track provides both information and the forces required to change direction. A painted line, however, provides only the information, and other parts of the vehicle have to supply the energy to change direction, e.g. an internal battery that powers sensors and motors. Evolution uses both sorts, e.g. wind blowing seeds away from parent plants and a wolf following a scent trail left by its prey. An unseen wall uses force to stop your forward motion in a dark room, whereas a visible, or lightly touched, wall provides only information Sloman [2011]. More sophisticated kits use sensors, e.g. optical, auditory, tactile, inertial, or chemical sensors, providing information that internal mechanisms can use to evaluate and select goals, control actions, interact with conspecifics, predict events in the environment, evaluate hypotheses, and other functions.

3.2 Mechanisms for Storing, Transforming and Using Information

Often information is acquired, used, and then lost because it is overwritten, e.g. sensor information in simple servo-control systems with "online intelligence", where only the latest sensed state is used for deciding whether to speed something up, change direction, etc. In more complex control systems, with "offline intelligence", sensor information is saved, possibly combined with previously stored information, and remains available for use on different occasions for different purposes.[18]

In the "offline" case, the underlying construction-kit needs to be able to support stores of information that grow with time and can be used for different purposes at different times. A control decision at one time may need items of information obtained at several different times and places, for example information about properties of a material, where it can be found, and how to transport it to where it is needed. Sensors used online may become faulty or require adjustment. Evolution may provide mechanisms for testing and adjusting. When used offline, stored information may need to be checked for falsity caused by the environment changing, as opposed to sensor faults. The offline/online use of visual information has caused

[18] Trehub [1991] proposed an architecture for vision that allows snapshots from visual saccades to be integrated in a multi-layer fixation-independent visual memory.

much confusion among researchers, including muddled attempts to interpret the difference in terms of "what" and "where" information.[19] Contrast Sloman [1983].

Ways of acquiring and using information have been discovered and modelled by AI researchers, psychologists, neuroscientists, biologists and others. However, evolution has produced many more. Some of them require not additional storage space but very different sorts of information-processing architectures. Different possible architectures of varying complexity and evolutionary age are discussed in Sloman [1978, 1983, 1993, 2003, 2006]. In contrast, AI engineers typically seek one architecture for a project. A complex biological architecture may use sub-architectures that evolved at different times, meeting different needs in different niches.

This raises the question whether evolution produced "architecture kits" able to combine evolved information-processing mechanisms in different ways, long before software engineers discovered the need. Such a kit could be particularly important for species that produce new subsystems, or modify old ones, during individual development, e.g. during different phases of learning by apes, elephants, and humans, as described in Sect. 2.3, contradicting the common assumption that a computational architecture must remain fixed.[20]

3.3 Mechanisms for Controlling Position, Motion and Timing

All concrete construction kits (and some hybrid kits) share a deep common feature insofar as their components, their constructs and their construction processes involve space and time, both during assembly and while working. Those behaviours include both relative motion of parts, e.g. wheels rotating, joints changing angles, and also motion of the whole object relative to other objects, e.g. an ape grasping a berry. *A consequence of spatiality is that objects built from different construction kits can interact, by changing their spatial relationships (e.g. if one object enters, encircles or grasps another), applying forces transmitted through space, and using spatial sensors to gain information used in control.* Products of different kits can interact in varied ways, e.g. one being used to assemble or manipulate another, or one providing energy or information for the other. Contrast the problems of getting software components available on a computer to interact sensibly: merely locating them in the same virtual or physical machine will not suffice. Some rule-based systems are composed of condition-action rules, managed by an interpreter that constantly checks for satisfaction of conditions. Newly added rules may then be invoked simply because

[19] http://en.wikipedia.org/wiki/Two-streams_hypothesis

[20] The BICA society aims to bring together researchers on biologically inspired cognitive architectures. Some examples are here: http://bicasociety.org/cogarch/

their conditions become satisfied, though "conflict resolution" mechanisms may be required if the conditions of more than one rule are satisfied.[21]

New concrete kits can be formed by combining two or more kits. In some cases this will require modification of a kit, e.g. combining Lego and Meccano by adding pieces with Lego studs or holes alongside Meccano-sized screw holes. In other cases, mere spatial proximity and contact suffices, e.g. when one construction kit is used to build a platform and others are used to assemble a house on it. Products of different biological construction kits may also use complex mixtures of juxtaposition and adaptation.

Objects that exist in space and/or time often need timing mechanisms. Organisms use "biological clocks" operating on different time-scales controlling repetitive processes, including daily cycles, heart-beats, breathing, and wing or limb movements required for locomotion. More subtly, there are adjustable speeds and adjustable rates of change: e.g. a bird in flight approaching a perch; an animal running to escape a predator and having to decelerate as it approaches a tree it needs to climb; a hand moving to grasp a stationary or moving object, with motion controlled by varying coordinated changes of joint angles at waist, shoulder, elbow and finger joints so as to bring the grasping points on the hand into suitable locations relative to the intended grasping points on the object. (This can be very difficult for robots, when grasping novel objects in novel situations, if they use ontologies that are too simple.) There are also biological mechanisms for controlling or varying rates of production of chemicals (e.g. hormones).

So biological construction kits need many mechanisms able to measure time intervals and to control rates of repetition or rates of change of parts of the organism. These kits may be combined with other sorts of construction kit that combine temporal and spatial control, e.g. changing speed and direction.

3.4 Combining Construction Kits

At the molecular level there is now a vast, and rapidly growing, amount of biological research on interacting construction kits, for example interactions between different parts of the reproductive mechanism during development of a fertilised egg, interactions between invasive viral or bacterial structures and a host organism, and interactions with chemicals produced in medical research laboratories. In computers, the ways of combining different toolkits include the application of functions to arguments, although both functions and their arguments can be far more complex than the cases most people encounter when learning arithmetic. A function could be a compiler, its arguments could be arbitrarily complex programs in a high-level

[21] Our SimAgent toolkit http://www.cs.bham.ac.uk/research/projects/poplog/packages/simagent.html [1996b] is an example.

programming language, and the output of the function might be either a report on syntactic errors in the input program, or a machine code program ready to run.

Applying functions to arguments is very different from assembling structures in space-time, where inputs to the process form parts of the output. If computers are connected via digital-to-analog interfaces linking them to surrounding matter, or if they are mounted on machines that allow them to move around in space and interact, that adds a kind of richness that goes beyond application of functions to arguments.

The additional richness is present in the modes of interaction of chemical structures that include both digital (on/off chemical bonds) and continuous changes in relationships, as discussed in Turing [1952], the paper on chemistry-based morphogenesis that inspired this Meta-Morphogenesis project Sloman [2013a].

3.5 Combining Abstract Construction Kits

Section 2.1 showed how a new DCK using combinations of old components can make some new developments very much quicker to reach — fewer steps are required, and the total search space for a sequence of steps to a solution may be dramatically reduced. Combining *concrete* construction kits uses space-time occupancy. Combining *abstract* construction kits is less straightforward. Sets of letters and numerals are combined to form labels for chess board squares, e.g. “a2”, “c5”, etc. A human language and a musical notation can form a hybrid system for writing songs. A computer operating system (e.g. Linux) can be combined with programming languages (e.g. Lisp, Java). In organisms, as in computers, products of different kits may share *information*, e.g. information for sensing, predicting, explaining or controlling, including information about information Sloman [2011].

Engineers combining different kinds of functionality find it useful to design re-usable information-processing *architectures* that provide frameworks for combining different mechanisms and information stores[20], especially in large projects where different teams work on sensors, learning, motor systems, reasoning systems, motivational systems, various kinds of metacognition, etc., using specialised tools. The toolkit mentioned in footnote 21 is an example framework. It is often necessary to support different sorts of **virtual** machinery interacting simultaneously with one another and with internal and external physical environments, during perception and motion. This may require new general frameworks for assembling complex *information-processing architectures*, accommodating multiple interacting virtual machines, with different modifications developed at different times Minsky [1987, 2006], Sloman [2003]. Self-extension is a topic for further research—See footnote 17.

Creation of new construction kits may start by simply recording parts of successful assemblies, or, better still, parametrized parts, so that they can easily be reproduced in modified forms—e.g. as required for organisms that change size and shape while developing. Eventually, parametrized stored designs may be combined to form a *"meta-construction kit"* able to extend, modify or combine previously created construction kits, as human engineers have recently learnt to do in software development environments. Evolution needs to be able to create new meta-construction kits using natural selection. Natural selection, the great creator/meta-creator, is now spectacularly aided and abetted by its products, especially humans!

4 Construction Kits Generate Possibilities and Impossibilities

Explanations of how things are possible (Section 1) can refer to construction kits, either manufactured, e.g. Meccano and Lego, or composed of naturally occurring components, e.g. boulders, mud, or sand. (Not all construction kits have sharp boundaries.) Each kit makes possible certain types of construct, instances of which can be built by assembling parts from the kit. Some construction kits use *products of products of* biological evolution, e.g. birds' nests assembled from twigs or leaves.

In some kits, features of components, such as shape, are inherited by constructed objects. E.g. objects composed only of Lego bricks joined in the "standard" way have external surfaces that are divisible into faces parallel to the surfaces of the first brick used. However, if two Lego bricks are joined at a corner only, using only one stud and one socket, it is possible to have continuous relative rotation (because studs and sockets are circular), violating that constraint, as Ron Chrisley pointed out in a conversation. This illustrates the fact that constructed objects can have "emergent" features none of the components have, e.g. a hinge is a non-rigid object that can be made from two rigid objects with aligned holes through which a screw is passed.

So, a construction kit that makes some things possible and others impossible can be extended so as to remove some of the impossibilities, e.g. by adding a hinge to Lego, or adding new parts from which hinges can be assembled.

4.1 Construction Kits for Making Information Users

Not everything that can play a role in acquisition, storage or transfer of information has information-processing capabilities. Consider a lump of plasticine or damp clay that can be deformed under pressure, then retains the deformation.

If a coin is pressed against it the lump will change its shape. Entities with information-processing capabilities (e.g. archaeologists) can use the depression as a source of information about the coin. But the deformed lump of material is not an information user. If the depression is used to control a process, e.g. making copies of the coin, or to help a historian years later, then the deformed material is used as a source of information about the coin. The fact that some part of a brain is changed by perceptual processes in an organism does not imply that that portion of the brain is an information user.

It may play a role analogous to the lump of clay, or a footprint in soil. Additional mechanisms are required if the information is to be *used*: different mechanisms for different types of use. A photocopier acquires information from a sheet of paper, but all it can do with the information is produce a replica (possibly after slight modifications such as changes in contrast, intensity or magnification). Different mechanisms are required for recognising text, correcting spelling, analysing the structure of an image, interpreting it as a picture of a 3-D scene, using information about the scene to guide a robot, building a copy of the scene, or answering a question about which changes are possible. Thinking up ways of using the impression as a source of information about the coin is left as an exercise for the reader.

Biological construction kits for producing information-processing mechanisms evolved at different times. Sloman [1993] discusses the diversity of uses of information from sensors, including sharing of sensor information between different uses, concurrently or sequentially. Subsystems can compete for sensors (e.g. concentrating on the road or admiring the scenery). Information vehicles such as sound or light provide multi-purpose information about the source or reflector of the sound or light, e.g. used for deciding whether to flee, or for controlling actions such as grasping or avoiding the information-source.

Some information-using mechanisms are direct products of biological evolution, e.g. reflex protective blinking mechanisms. Others are grown by epigenetic mechanisms influenced by context. For example, humans in different cultures start with a generic language construction kit (sometimes misleadingly labelled a “universal grammar”) which is extended and modified to produce locally useful linguistic mechanisms. Language-specific mechanisms, such as mechanisms for acquiring, producing, understanding and correcting textual information, must have evolved long after mechanisms shared between many species that can use visual information for avoiding obstacles or grasping objects. In some species, diversity in the construction kits produced by individual genomes, can lead to even greater diversity in adults, especially if they develop in different physical and cultural environments using the epigenetic mechanisms suggested in Section 2.3 and Fig. 3.

4.2 Different Roles for Information

Despite huge diversity in biological construction kits and the mechanisms in individual organisms, some themes recur, such as functions of different sorts of information in control: e.g. information about how things actually are or might be (“belief-like” information contents), information about how things need to be or might need to be for the individual information user (“desire-like” information contents), and information about how to achieve or avoid certain states ("procedural" information contents).

Each type has different subtypes: across species, across members of a species and across developmental stages in an individual. How a biological construction kit supports all those requirements depends on the environment, the animal's sensors, its needs, the local opportunities, and the individual's history. Different mechanisms performing such functions may share a common evolutionary precursor after which they diverged. Moreover, mechanisms with similar functions can evolve independently: convergent evolution.

Information relating to targets and how to achieve or maintain them is *control* information: the most basic type of biological information, from which all others are derived. A simple case is a thermostatic control, discussed in McCarthy [1979]. It has (at least) two sorts of information: (a) a *target* temperature ("desire-like" information) (b) *current temperature* ("belief-like" information). A discrepancy between them causes the thermostat to select between turning a heater on, or off, or doing nothing.

This very simple homeostatic mechanism uses information and a source of energy to achieve or maintain a target state. There are very many variants on this schema, based on the type of target (e.g. a measured state or some complex relationship), the type of control (on, off, or variable, with single or multiple effectors), and the mechanisms by which targets and control actions are selected, which may be modified by learning, and may use simple actions or complex plans.

As Gibson [1966] pointed out, acquisition of information often requires cooperation between processes of sensing and acting. Saccades are visual actions that constantly select new information samples from the environment (or the optic cone). Uses of the information vary widely according to context, e.g. controlling grasping, controlling preparation for a jump, controlling avoidance actions, or sampling text to be read. A particular sensor can therefore be shared between many control subsystems (Sloman [1993]), and the significance of the sensor state will depend partly on which subsystems are connected to the sensor at the time and partly on which other mechanisms receive information from the sensor (which may change dynamically — a possible cause of some types of "change blindness").

The study of varieties of use of information in organisms is exploding, and includes many mechanisms on molecular scales as well as many intermediate levels of informed control, including sub-cellular levels (e.g. metabolism), physiological processes of breathing, temperature maintenance, digestion, blood circulation, control of locomotion, feeding and mating of large animals and coordination in communities, such as collaborative foraging in insects and trading systems of humans. Slime moulds include spectacular examples in which modes of acquisition and use of information change dramatically.[22]

[22] http://www.theguardian.com/cities/2014/feb/18/slime-mould-rail-road-transport-routes

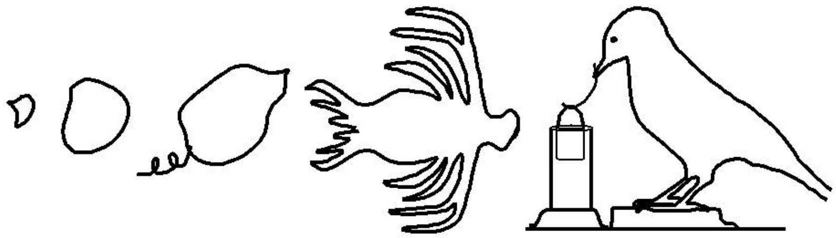

Fig. 4 Between the simplest and most sophisticated organisms there are many intermediate forms with very different information processing requirements and capabilities.

The earliest organisms must have acquired and used information about things inside themselves and in their immediate vicinity, e.g. using chemical detectors in an enclosing membrane. Later, evolution extended those capabilities in dramatic ways (crudely indicated in Fig. 4). In the simplest cases, local information is used immediately to select between alternative possible actions, as in a heating control or a trail-following mechanism. Uses of motion in haptic and tactile sensing and uses of saccades, changing vergence, and other movements in visual perception all depend on the interplay between sensing and doing, characteristic of "online intelligence".

But there are cases ignored by Gibson and anti-cognitivists, namely organisms that exhibit "offline intelligence", using perceptual information for tasks other than controlling immediate reactions, for example, reasoning about remote future possibilities or attempting to explain something observed, or working out that bending a straight piece of wire will enable a basket of food to be lifted out of a tube as in Fig. 4 Weir et al. [2002], or correcting a stored generalisation. Doing that (namely "using perceptual information for tasks other than controlling immediate reactions)" requires use of previously acquired information about the environment, including particular information about individual objects and their locations or states, general information about learnt laws or correlations and information about what is and is not possible. (Compare footnote [11].)

An information-bearing structure (e.g. the impression of a foot, the shape of a rock) can provide very different information to different information users, or to the same individual at different times, depending on (a) what kinds of sensors they have, (b) what sorts of information-processing (storing, analysing, comparing, combining, synthesizing, retrieving, deriving, using, etc.) mechanisms they have, (c) what sorts of needs or goals they can serve by using various sorts of information (knowingly or not), and (d) what information they already have. So, from the fact that changes in some portion of a brain correlate with changes in some aspect of the environment, we cannot conclude much about what information about the environment the brain acquires and uses or how it does that, since typically that will depend on context.

4.3 Motivational Mechanisms

It is often assumed that every information user, U, constantly tries to achieve rewards or avoid punishments (negative rewards), and that each new item of information, I, will make some actions more likely for U, and others less likely, on the basis of what U has previously learnt about which actions increase positive rewards or decrease negative rewards under conditions indicated by I. But animals are not restricted to acting on motives selected *by them* on the basis of expected rewards. They may also have motive generators that are simply triggered as "internal reflexes" just as evolution produces phototropic reactions in plants without giving plants any ability to anticipate benefits to be gained from light. Some reflexes, instead of directly triggering *behaviour*, trigger *new motives*, which may or may not lead to behaviour, depending on the importance of other competing motives. For example, a kind person watching someone fall may acquire a motive to rush to help, which is not acted on if competing motives are too strong. It is widely believed that all motivation is reward-based. But a new motive triggered by an internal reflex need not be associated with some reward. It may be "architecture-based motivation" rather than "reward-based motivation" Sloman [2009]. Triggering of architecture-based motives in playful intelligent young animals can produce kinds of delayed learning that the individuals could not possibly anticipate, and therefore cannot be motivated by Karmiloff-Smith [1992].

Unforeseeable biological benefits of automatically triggered motives include acquisition of new information by sampling properties of the environment. The new information may not be immediately usable, but in combination with information acquired later and genetic tendencies activated later, as indicated in Fig. 3, it may turn out to be important during hunting, caring for young, or learning a language. A toddler may have no conception of the later potential uses of information gained in play, though the ancestors of that individual may have benefited from the presence of the information-gathering reflexes. In humans this seems to be crucial for mathematical development.

During evolution, and also during individual development, the sensor mechanisms, the types of information-processing, and the uses to which various types of information are put, become more diverse and more complex, while the information-processing architectures allow more of the processes to occur in parallel (e.g. competing, collaborating, invoking, extending, recording, controlling, redirecting, enriching, training, abstracting, refuting, or terminating). Without understanding how the architecture grows, which information-processing functions it supports, and how they diversify and interact, we are likely to reach wrong conclusions about biological functions of the parts: e.g. over-simplifying the functions of sensory subsystems or over-simplifying the variety of concurrent control mechanisms involved in producing behaviours. Moreover, the architectural knowledge about how such a system works, like information about the architecture of a computer operating system, may not be expressible in sets of equations, or statistical learning mechanisms and relationships. (Ideas about architectures for human information-processing can be found in Simon [1967], Minsky [1987], Laird et al. [1987], Sloman [2003], Minsky [2006], Sun [2006], Sloman [2013b], among many others.)

Construction kits for building information-processing architectures with multiple sensors and motor subsystems in complex and varied environments differ widely in the designs they can produce. Understanding that variety is not helped by disputes about which architecture is best. A more complete discussion would need to survey the design options and relate them to actual choices made by evolution or by individuals interacting with their environments.

5 Mathematics: Some Constructions Exclude or Necessitate Others

Physical construction kits (e.g. Lego, plasticine, or a combination of paper, scissors and paste) have parts and materials with physical properties (e.g. rigidity, strength, flexibility, elasticity, adhesion, etc.), possible relationships between parts and possible processes that can occur when the parts are in those relationships (e.g. rotation, bending, twisting and elastic or inelastic resistance to deformation).

Features of a physical construction kit—including the shapes and materials of the basic components, ways in which the parts can be assembled into larger wholes, kinds of relationships between parts and the processes that can occur involving them—explain the possibility of *entities* that can be constructed and the possibility of *processes*, including processes of construction and behaviours of constructs.

Construction kits can also explain necessity and impossibility. A construction kit with a large initial set of generative powers can be used to build a structure realising some of the kit's possibilities, in which some further possibilities are excluded, namely all extensions that do not include what has so far been constructed. If a Meccano construction has two parts in a substructure that fixes them a certain distance apart, then no extension can include a new part that is wider than that distance in all dimensions and is in the gap. Some extensions to the part-built structure that were previously possible become impossible unless something is undone. That example involves a limit produced by a gap size. There are many more examples of impossibilities that arise from features of the construction kit.

Euclidean geometry includes a construction kit that enables construction of closed planar polygons (triangles, quadrilaterals, pentagons, etc.), with interior angles whose sizes can be summed. If the polygon has three sides, i.e. it is a triangle, then the interior angles must add up to exactly half a rotation. Why? In this case, no physical properties of a structure (e.g. rigidity or impenetrability of materials) are involved, only spatial relationships. Figure 5 provides one way to answer the question, unlike the standard proofs, which use parallel lines. It presents a proof, found by Mary

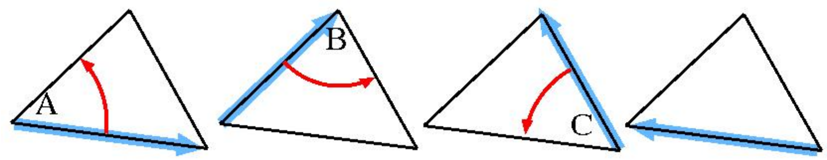

Fig. 5 The sequence demonstrates how the three-cornered shape has the consequence that summing the three angles necessarily produces half a rotation (180°). Since the position, size, orientation, and precise shape of the triangle can be varied without affecting the possibility of constructing the sequence, this is a proof that generalises to any planar triangle. It nowhere mentions Euclid's parallel axiom, used by "standard" proofs. This unpublished proof was reported to me by Mary Pardoe, a former student who became a mathematics teacher, in the early 1970s.

Pardoe, that internal angles of a planar triangle sum to a straight line, or 180°. (I am ignoring the question how to verify that the surface is planar.)

Most humans are able to look at a physical situation, or a diagram representing a class of physical situations, and reason about constraints on a class of possibilities sharing a common feature. This may have evolved from earlier abilities to reason about changing affordances in the environment (Gibson [1979]). Current AI perceptual and reasoning systems still lack most of these abilities, and neuroscience cannot yet explain what's going on (as opposed to where it's going on?). (See footnote [11].)

These illustrate mathematical properties of construction kits (partly analogous to mathematical properties of formal deductive systems and AI problem solving systems). As parts (or instances of parts) of the FCK are combined, structural relations between components of the kit have two opposed sorts of consequences: they make some further structures *possible* (e.g. constructing a circle that passes through all the vertices of the triangle), and other structures *impossible* (e.g. relocating the corners of the triangle so that the angles add up to 370°). These possibilities and impossibilities are *necessary* consequences of previous selection steps. The examples illustrate how a construction kit with mathematical relationships can provide the basis for necessary truths and necessary falsehoods in some constructions (as argued in [Sloman, 1962, Chap. 7]).[23] Being able to think about and reason about possible and impossible alterations in some limited portion of the environment is a very common requirement for intelligent action Sloman [1996a]. It seems to be partly shared with other intelligent species, e.g. squirrels, nest-builders, elephants, apes, etc. Since our examples of making things possible or impossible, or changing ranges of possibilities, are examples of causation (mathematical causation), this also provides the basis for a Kantian notion of causation based on mathematical

[23] Such relationships between possibilities provide a deeper, more natural basis for understanding modality (necessity, possibility, impossibility) than so-called "possible world semantics". I doubt that most normal humans who can think about possibilities and impossibilities base that ability on thinking about truth in the whole world, past, present and future, and in the set of alternative worlds.

necessity Kant [1781], so that not all uses of the notion of "cause" are Humean (i.e. based on empirical correlations), even if some are. Compare Section 5.3.[24]

Neuroscientific theories about information-processing in brains currently omit the processes involved in such mathematical discoveries, so AI researchers influenced too much by neuroscience may fail to replicate important brain functions. Progress may require major conceptual advances regarding what the problems are and what sorts of answers are relevant.

We now consider ways in which evolution itself can be understood as discovering mathematical proofs—proofs of possibilities.

5.1 Proof-Like Features of Evolution

A subset of the FCK produced fortuitously as a side effect of formation of the earth supported (a) primitive life forms and (b) processes of evolution that produced more and more complex forms of life, including new, more complex, derived, DCKs. New products of natural selection can make more complex products more reachable, as with toy construction kits and mathematical proofs. However, starting from those parts will make some designs unreachable except by disassembling some parts.

Moreover, there is not just one sequence: different evolutionary lineages evolving in parallel can produce different DCKs. According to the "Symbiogenesis" theory, different DCKs produced independently can sometimes merge to support new forms of life combining different evolutionary strands.[25] Creation of new DCKs in parallel evolutionary streams with combinable products can hugely reduce part of the search space for complex designs, at the cost of excluding parts of the search space reachable from the FCK. For example, use of DCKs in the human genome may speed up development of language and typical human cognitive competences, while excluding the possibility of "evolving back" to microbe forms that might be the only survivors after a cataclysm.

5.2 Euclid's Construction Kit

An old example, of great significance for science, mathematics, and philosophy, is the construction kit specified in Euclidean geometry, starting with points, lines, surfaces and volumes, and methods of constructing new more complex geometrical configurations using a straight edge for drawing straight lines in a plane surface, and a pair of compasses for drawing circular arcs. This construction kit makes it

[24] For more on Kantian vs. Humean causation, see the presentations on different sorts of causal reasoning in humans and other animals by Chappell and Sloman at the Workshop on Natural and Artificial Cognition (WONAC, Oxford, 2007): http://www.cs.bham.ac.uk/research/projects/cogaff/talks/wonac Varieties of causation that do not involve mathematical necessity, only probabilities (Hume?) or propensities (Popper), will not be discussed here.

[25] http://en.wikipedia.org/wiki/Symbiogenesis

possible to bisect, but not trisect, an arbitrary planar angle. A slight extension, the "Neusis construction", known to Archimedes, allows line segments to be translated and rotated in a plane while preserving their length, and certain incidence relations. This allows arbitrary angles to be trisected! (See http://www.cs.bham.ac.uk/research/projects/cogaff/misc/trisect.html.)

The ability of (at least some) humans to discover such things must depend on evolved information-processing capabilities of brains that are as yet unknown and not yet replicated in AI reasoning systems. The idea of a space of possibilities generated by a physical construction kit may be easier for most people to understand than the comparison with generative powers of grammars, formal systems, or geometric constructions, though the two must be connected, since grammars and mathematical systems are abstract construction kits that can be parts of hybrid construction kits.

Concrete construction kits corresponding to grammars can be built out of physical structures. For example, a collection of small squares with letters and punctuation marks, and some blanks, can be used to form sequences that correspond to the words in a lexicon. A cursive ("joined up") script requires a more complex physical construction kit. Human sign-languages are far more demanding, since they involve multiple body parts moving concurrently.

Some challenges for construction kits used by evolution, and also challenges for artificial intelligence and philosophy, arise from the need to explain both how natural selection makes use of mathematical properties of construction kits related to geometry and topology, in producing organisms with spatial structures and spatial competences, and also how various subsets of those organisms (e.g. nest-building birds) developed specific topological and geometrical reasoning abilities used in controlling actions or solving problems; and finally how at least one species developed abilities to reflect on the nature of those competences and eventually, through unknown processes of individual development and social interaction, using unknown representational and reasoning mechanisms, managed to produce the rich, deep and highly organised body of knowledge published as Euclid's *Elements* (see footnote 1). There are important aspects of those mathematical competences that, as far as I know, have not yet been replicated in Artificial Intelligence or Robotics.[26] Is it possible that currently understood forms of digital computation are inadequate for the tasks, whereas chemistry-based information-processing systems used in brains are richer, because they combine both discrete and continuous operations, as discussed in Sect. 2.5? (That's not a rhetorical question: I don't know the answer.)

5.3 Mathematical Discoveries Based on Exploring Construction Kits

Some mathematical discoveries result from observation of naturally occurring physical construction kits and noticing how constraints on modes of composition of

[26] Several are listed at http://www.cs.bham.ac.uk/research/projects/cogaff/misc/mathstuff.html

components generate constraints on resulting constructs. E.g. straight line segments on a surface can be joined end to end to enclose a finite region, but that is impossible with only two lines, as noted by Kant in [1781]. Likewise, flat surfaces can be combined to enclose a volume, such as a tetrahedron or cube, but it is impossible for only three flat surfaces to enclose a finite space. It is not clear how humans detect such impossibilities: no amount of trying and failing can establish impossibility. Kant had no access to a twentieth-century formal axiomatisation of Euclidean geometry. What he, and before him Euclid, Archimedes and others had were products of evolution. What products?

Many mathematical domains (perhaps all of them) can be thought of as sets of possibilities generated by construction kits. Physicists and engineers deal with hybrid concrete and abstract construction kits. The space of possible construction kits is also an example. As far as I know, this domain has not been explored systematically by mathematicians, though many special cases have.

In order to understand biological evolution on this planet, we need to understand the sorts of construction kits made possible by the existence of the physical universe, and in particular the variety of construction kits inherent in the physics and chemistry of the materials of which our planet was formed, along with the influences of its environment (e.g. solar radiation, asteroid impacts). An open research question is whether a construction kit capable of producing all the non-living structures on the planet would also suffice for evolution of all the forms of life on this planet, or whether life and evolution have additional requirements, e.g. cosmic radiation.

5.4 Evolution's (Blind) Mathematical Discoveries

Insofar as construction kits have mathematical properties, life and mathematics are closely interconnected, as we have already seen. More complex relationships arise after evolution of mathematical metacognitive mechanisms. On the way to achieving those results, natural selection often works as "a blind theorem prover". Many of the "theorems" are about new *possible* structures, processes, organisms, ecosystems, etc. The proofs that they are possible are implicit in the evolutionary trajectories that lead to occurrences. Proofs are often thought of as abstract entities that can be represented physically in different ways (using different formalisms) for communication, persuasion (including self-persuasion), predicting, explaining and planning. A physical sequence produced unintentionally, e.g. by natural selection or by plant growth, that leads to a new sort of entity is a proof that some construction kit makes that sort of entity possible. The evolutionary or developmental trail answers the question: how is that sort of thing possible? So biological evolution can be construed as a "blind theorem prover", despite there being no intention behind the proof. Proofs of *impossibility* (or *necessity*) raise more complex issues, to be discussed elsewhere.

These observations seem to support a new kind of "Biological-evolutionary" foundation for mathematics[27] that is closely related to Immanuel Kant's philosophy of mathematics in his *Critique of Pure Reason* (1781). I attempted to defend his ideas in Sloman [1962]. This answers questions like "How is it possible for things that make mathematical discoveries to exist?", an example of explaining a possibility (See footnote 3). Attempting to go too directly from hypothesized properties of the primordial construction kit (or the physical universe) to explaining advanced capabilities such as human self-awareness, without specifying all the relevant construction kits, including required temporary scaffolding, will fail, because short-cuts omit essential details of both the problems and the solutions, like mathematical proofs with gaps.

Many of the "mathematical discoveries" (or inventions?) produced (blindly) by evolution depend on mathematical properties of physical structures or processes or problem types, whether they are specific solutions to particular problems (e.g. use of negative feedback control loops) or new construction-kit components that are usable across a very wide range of different species (e.g. the use of a powerful "genetic code", the use of various kinds of learning from experience, the use of new forms of representation for information, the use of new physical morphologies to support sensing, or locomotion, or consumption of nutrients, etc.)

These mathematical "discoveries" started happening long before there were any humans doing mathematics (refuting claims that humans create mathematics). Many of the discoveries were concerned with what is possible, either absolutely or under certain conditions, or for a particular sort of construction-kit. Other discoveries, closer to what are conventionally thought of as mathematical discoveries, are concerned with limitations on what is possible, i.e. necessary truths. Some discoveries are concerned with probabilities derived from statistical learning, but I think the relative importance of statistical learning in biology has been vastly overrated because of misinterpretations of evidence (to be discussed elsewhere). In particular, the discovery that something important is possible does not require statistical evidence: a single instance suffices. No amount of statistical evidence can show that something is impossible: structural constraints need to be analysed. For human evolution, a particularly important subtype of mathematical discovery was the unwitting discovery and use of mathematical (e.g. topological) structures in the environment, a discovery process that starts in human children before they are aware of what they are doing, and in some species without any use of language for communication. Examples are discussed in the "Toddler Theorems" document referenced in footnote 17.

[27] More fine-grained varieties of evolutionary foundation for mathematics are summarised in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/maths-multiple-foundations.html

6 Varieties of Derived Construction Kit

DCKs may differ (a) at different evolutionary stages within a lineage, (b) across lineages (e.g. in different coexisting organisms), and (c) during development of an individual that starts as a single cell and produces mechanisms that support different kinds of growth, development and information processing at different stages of development (Sect. 2.3). New construction kits can also be produced by cultures or ecosystems (e.g. human languages) and applied sciences (e.g. bioengineering, computer systems engineering). New cases build on what was previously available. Sometimes separately evolved DCKs are combined, for instance in symbiosis, sexual reproduction, and individual creative learning.

What sort of kit makes it possible for a child to acquire competence in any one of the thousands of different human languages (spoken or signed) in the first few years of life? Children do not merely learn pre-existing languages: they *construct* languages that are new for them, constrained by the need to communicate with conspecifics, as shown dramatically by Nicaraguan deaf children who developed a sign language going beyond what their teachers understood Senghas [2005]. There are also languages that might have developed but have not (yet). Evolution of human spoken language may have gone from purely internal languages needed for perception, intention, etc., through collaborative actions, then (later) signed communication, then spoken communication, as argued in Sloman [2015].

If language acquisition were mainly a matter of learning from expert users, human languages could not have existed, since initially there were no expert users to learn from, and learning could not get started. This argument applies to any competence thought to be based entirely on learning from experts, including mathematical expertise. So data mining in samples of expert behaviours will never produce AI systems with human competences—only inferior subsets at best.

The history of computing since the earliest calculators illustrates changes that can occur when new construction kits are developed. There have not only been changes in size, speed and memory capacity: there have also been profound qualitative changes, e.g. in new layers of virtual machinery, such as new sorts of mutually interacting causal loops linking virtual machine control states with portions of external environments, as in the use of GPS-based navigation. Long before that, evolved virtual machines provided semantic contents referring to non-physical structures and processes, e.g. mathematical problems, rules of games, and mental contents referring to possible future mental contents (e.g. "What will I see if...?"), including contents of other minds.

I claim, but will not argue here, that some new machines cannot be *fully described* in the language of the FCK even though they are *fully implemented* in physical mechanisms. (See Sect. 2.2 on ontologies.) We now understand many key components and many modes of composition that provide platforms on which *human-designed* layers of computation can be constructed, including subsystems closely but not rigidly coupled to the environment (e.g. a hand-held video camera).

Several different "basic" abstract construction kits have been proposed as sufficient for the forms of (discrete) computation required by mathematicians: namely Turing machines, Post's production systems, Church's Lambda Calculus, and several more, each capable of generating the others. The Church-Turing thesis claims that each is sufficient for all forms of computation.[28] There has been an enormous amount of research in computer science and computer systems engineering on forms of computation that can be built from such components. One interpretation of the Church-Turing thesis is that these construction kits generate all *possible* forms of information-processing. But it is not at all obvious that those discrete mechanisms suffice for all biological forms of information-processing. For example, chemistry-based forms of computation include both discrete mechanisms (e.g. forming or releasing chemical bonds) of the sort Schrödinger discussed, and continuous process, e.g. folding, twisting, etc. used in reproduction and other processes. Ganti [2003] shows how a chemical construction-kit can support forms of biological information processing that don't depend only on external energy sources (a fact that's also true of battery-powered computers), and can also support growth and reproduction using internal mechanisms, which human-made computers cannot do (yet).

There seem to be many different sorts of construction-kit that allow different sorts of information processing to be supported, including some that we don't yet understand. In particular, the physical/chemical mechanisms that support the construction of both physical structures and information-processing mechanisms in living organisms may have abilities not available in digital computers.[29]

6.1 A New Type of Research Project

Very many biological processes and associated materials and mechanisms are not well understood, though knowledge about them is increasing rapidly. It is hard to know how many of the derived construction kits have not yet been identified and studied. I am not aware of any systematic attempt to identify features of the FCK that suffice to explain the possibility of all known evolved biological DCKs. Researchers in fundamental physics and cosmology do not normally attempt to ensure that their theories explain the many materials and process types that have been explored by natural selection and its products, in addition to known facts about physics and chemistry. Schrödinger [1944] pointed out that a theory of the physical basis of life should explain such phenomena, though he could not have appreciated some of the requirements for sophisticated forms of information processing, because, at the time

[28] For more on this see: http://en.wikipedia.org/wiki/Church-Turing_thesis

[29] Examples of human mathematical reasoning in geometry and topology that have, until now, resisted replication on computers are presented in http://www.cs.bham.ac.uk/research/projects/cogaff/misc/torus.html and http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-sum.html

he wrote, scientists and engineers had not learnt what we now know. Curiously, although he mentioned the need to explain the occurrence of metamorphosis in organisms, the example he gave was the transformation from a tadpole to a frog. He could have given more spectacular examples, such as the transformation of a caterpillar to a butterfly via an intermediate stage as a chemical soup in an outer case, from which the butterfly later emerges.[30]

Penrose [1994] attempted to show how features of quantum physics explain obscure features of human consciousness, especially mathematical consciousness, but he ignored all the intermediate products of biological evolution on which animal mental functions build. Human mathematics, at least the ancient mathematics done before the advent of modern algebra and logic, seems to build on animal abilities, for instance abilities to see various types of affordance. The use of diagrams and spatial models by Penrose could be an example of that. It is unlikely that there are very abstract human mathematical abilities that somehow grow directly out of quantum mechanical aspects of the FCK, without depending on the mostly unknown layers of perceptual, learning, motivational, planning, and reasoning competences produced by billions of years of evolution.

Twentieth-century biologists understood some of the achievements of the FCK in meeting physical and chemical requirements of various forms of life, though they used different terminology from mine, e.g. Haldane.[31] However, the task can never be finished, since the process of construction of new derived biological construction kits may continue indefinitely, producing new kits with components and modes of composition that allow production of increasingly complex types of *structure* and *behaviour* in organisms. That idea is familiar to computer scientists and engineers since very many new sorts of computational construction kit (new programming languages, new operating systems, new virtual machines, new development toolkits) have been developed from old ones in the last half century, making possible new kinds of computing system that could not previously be built from the original computing machinery, without introducing new intermediate layers, including new virtual machines that are able to detect and record their own operations, a capability that is often essential for debugging and extending computing systems. Sloman [2013b] discusses the importance of virtual machinery in extending what information-processing systems can do, and the properties they can have, including radical self-modification while running.

6.2 Construction Kits for Biological Information Processing

Each newly evolved mechanism provides opportunities for yet more control at higher levels: a recurring process that can repeatedly generate opportunities for *additional*

[30] See http://en.wikipedia.org/wiki/Pupa, and http://en.wikipedia.org/wiki/Holometabolism

[31] http://en.wikipedia.org/wiki/J._B._S._Haldane

mechanisms for (information-based) control of recently evolved mechanisms, for example choosing between competences or “tuning” them dynamically, often on the basis of their mathematical properties.

Implicit mathematical discovery processes enable production of competences used in interpretation of sensory information, e.g. locating perceived objects, events and processes in 2-D or 3-D space and time, and deriving coherent wholes from separate information fragments. Further enhancements may include: new mechanisms for prediction; for motive generation and selection; for construction, comparison, selection and control of new plans, with resulting new behaviours.

Many of evolution’s new mathematical discoveries, e.g. use of negative feedback control loops, were used in new designs producing useful behaviours, e.g. controlling temperature, osmotic pressure and other states; use of geometric constraints by bees whose cooperative behaviours produce hexagonal cells in honeycombs; and use of new ontologies for separating situations requiring different behaviours, e.g. manipulating different materials or hunting different kinds of prey.

As a result, construction kits used by evolution produced metacognitive mechanisms enabling individuals to notice and reflect on their own discoveries (enabling some of them to notice and remove flaws in their reasoning). Such metacognitive capabilities are required for abilities to communicate discoveries to others, discuss them, use them in shared practical tasks (e.g. making tools, clothes or weapons, building shelters, planning routes, discussing what will or will not work, and why), then later organising them into complex, highly structured bodies of shared knowledge, such as Euclid’s *Elements* (footnote 1). I don’t think anyone knows how long all of this took, what the detailed evolutionary changes were, or how the required mechanisms of perception, motivation, intention formation, reasoning, planning and retrospective reflection actually evolved. The brain mechanisms involved are also mostly unknown.

Explaining how all that could happen, and what it tells us about the nature of mathematics and biological/evolutionary foundations for mathematical knowledge, is a long-term goal of the Meta-Morphogenesis project. That includes seeking unnoticed overlaps between the human competences discovered by metacognitive mechanisms, and similar competences in animals that lack the metacognition, such as young humans making and using mathematical discoveries, on which they are unable to reflect because the required architecture has not yet developed.

Other intelligent species make and use similar “proto-mathematical” discoveries, without the meta-cognitive abilities required to notice what they are doing.

This could stimulate new research in robotics attempting to replicate such competences. Most of the naturally occurring mathematical abilities have not yet been replicated in Artificial Intelligence systems or robots, unlike logical, arithmetical, and algebraic competences that are relatively new to humans and (paradoxically?) easier to replicate on computers. Examples of topological reasoning about equivalence classes of closed curves not yet modelled in computers (as far as

I know) are referenced in footnote 29. Even the ability to reason about alternative ways of putting a shirt on a child (footnote 16) is still lacking. It is not clear whether the difficulty of replicating such mathematical reasoning processes is due to the need for a kind of construction kit that digital computers (e.g. Turing machines) cannot support, or due to our lack of imagination in using computers to replicate some of the products of biological evolution, or both! Perhaps there are important forms of representation or types of information-processing architecture still waiting to be discovered by AI researchers. Alternatively, the gaps may be connected with properties of chemistry-based information-processing mechanisms combining discrete and continuous interactions, or other physical properties that cannot be replicated exactly (or even approximately) in familiar forms of computation. (This topic requires more detailed mathematical analysis.)

6.3 Representational Blind Spots of Many Scientists

Although I cannot follow all the details of writings of physicists, I think it is clear that most debates regarding what should go into a fundamental theory of matter ignore most of the biological demands on such a theory. For example, presentations on dynamics of physical systems make deep use of branches of mathematics concerned with sets of numerical values, and the ways in which different measurable or hypothesized physical values do or do not covary, as expressed in (probabilistic or non-probabilistic) equations of various sorts. But the biological functions of complex physiological structures, especially structures that change in complexity as they develop, don't necessarily have those forms.

Biological mechanisms include: digestive mechanisms; mechanisms for transporting chemicals; mechanisms for detecting and repairing damage or infection; mechanisms for storing reusable information about an extended structured environment; mechanisms for creating, storing and using complex percepts, thoughts, questions, values, preferences, desires, intentions and plans, including plans for cooperative behaviours; and mechanisms that transform themselves into new mechanisms with new structures and functions.

Forms of mathematics used by physicists are not necessarily useful for studying such biological mechanisms. Logic, grammars and map-like representations are sometimes more appropriate, though I think little is actually known about the variety of forms of representation (i.e. encodings of information) used in human and animal minds and brains. We may need entirely new forms of mathematics for biology, and therefore for specifying what physicists need to explain.

Many physicists, engineers and mathematicians who move into neuroscience assume that states and processes in brains need to be expressed as collections of numerical measures and their derivatives plus equations linking them, a form of representation that is well supported by widely used tools such as Matlab, but is not necessarily best suited for the majority of types of mental content (e.g. grammatical and semantic structures of thoughts like those expressed here). Related challenges are posed by attempts to model chemical processes, where complex molecules form and interact with multiple changing chemical bonds along with changing geometrical and topological relationships–one of the reasons for the original invention of symbolic chemical notations now being extended in computer models of changing interacting molecular structures. (There are many online videos of computer simulations of chemical reactions including protein folding processes.)

6.4 Representing Rewards, Preferences, Values

It is often assumed that all intelligent decision making uses positive or negative scalar rewards or utility values that are comparable across options Luce and Raiffa [1957]. But careful attention to consumer magazines, political debates, and the varieties of indecision in normal human life shows that reality is far more complex. For example, many preferences are expressed in rules about how to choose between certain options. Furthermore, preferences can be highly sensitive to changes in context. A crude example is the change in preference for type of car after having children. Analysis of examples in consumer reports led to the conclusion that "better" is a complex, polymorphic, logical concept with a rich structure that cannot be reduced to simple comparisons of numerical values Sloman [1969, 1970]. Instead of a linear reward or utility metric, choices for intelligent individuals, or for natural selection, involve a complex network of partial orderings, with "annotated" links between nodes (e.g. "better" qualified by conditions: "better for", "better if", "better in respect of"...). In the Birmingham CogAff project Sloman [2003], those ideas informed computational models of simple agents with complex choices to be made under varying conditions, but the project merely scratched the surface, as reported in Beaudoin and Sloman [1993], Beaudoin [1994], Wright et al. [1996], Wright [1997].[32] Most AI/Cognitive Science models use much shallower notions of motivation.

Despite all the sophistication of modern psychology and neuroscience, I believe they currently lack the conceptual resources required to describe either functions of brains in dealing with these matters, including forms of development and learning required, or the mechanisms implementing those functions. In particular, we lack deep explanatory theories about mechanisms that led to mathematical discoveries over thousands of years, including brain mechanisms producing mathematical conjectures, proofs, counter-examples, proof-revisions, new scientific theories, new works of art and new styles of art. In part that's because models considered so far lack both sufficiently rich forms of information processing (computation) and sufficiently deep methodologies for identifying what needs to be explained. There

[32] Further information is available here: http://www.cs.bham.ac.uk/research/projects/cogaff/

are other unexplained phenomena concerned with artistic creation and enjoyment, and the mechanisms involved in finding something funny.

7 Computational/Information-Processing Construction Kits

Since the mid-twentieth century, we have been learning about abstract construction kits whose products are machines that can be used for increasingly complex tasks. Such construction kits include programming languages, operating systems, software development tools and environments, and network technology that allows ever more complex information-processing machines to be constructed by combining simpler ones. A crucial, but poorly understood, feature of that history is the growing use of construction kits based on virtual machinery, mentioned in Sect. 2. A complete account of the role of construction kits in biological evolution would need to include an explanation of how the fundamental construction kit (FCK) provided by the physical universe could be used by evolution to produce an increasing variety of types of *virtual* machinery as well as increasingly varied *physical* structures and mechanisms.

7.1 Infinite, or Potentially Infinite, Generative Power

A construction kit implicitly specifies a large, in some cases infinite, set of possibilities, though as an instance of the kit is constructed, each addition of a new component or feature changes the set of possibilities accessible in later steps of that construction process. For example, as you construct a sentence or phrase in a language, at each state in the construction there are alternative possible additions (not necessarily at the end) and each of those additions will alter the set of possible further additions consistent with the vocabulary and grammar of the language. When use of language is embedded in a larger activity, such as composing a poem, that context can modify the constraints that are relevant. Chemistry does something like that for types of molecule, types of process involving molecular changes, and types of structure made of multiple molecules. Quantum mechanics added important constraints to nineteenth century chemistry, including both the possibility of highly stable structures (resistant to thermal buffetting) and also locks and keys as in catalysis. All of that is essential for life as we know it, and also for forms of information processing produced by evolution (mostly not yet charted).

Research in fundamental physics is a search for the construction kit that has the generative power to accommodate all the varieties of matter, structure, process, and causation that can exist in our universe. However, physicists generally seek only to ensure that their construction kits are capable of accounting for phenomena observed in the physical sciences, most of which do not include production of living matter, or processes of evolution, development, learning, and mathematical discovery found in living organisms. Most do not try to ensure that their fundamental theories can account for those features also. There are notable exceptions, including Schrödinger's 1944 book, but most physicists (understandably) ignore most of the details of life, including the variety of forms it can take, the variety of environments coped with, the different ways in which individual organisms cope and change, the ways in which products of evolution become more complex and more diverse over time, and the many kinds of information processing and control both in individuals and in colonies (e.g. ant colonies), societies, and ecosystems.

If cosmologists and other theoretical physicists attempted to account for a wider range of biological phenomena, including the phenomena discussed here in connection with the Meta-Morphogenesis project, they would find considerable explanatory gaps between current physical theories and the diversity of phenomena of life, not because there is something about life that goes beyond what science can explain, but because we do not yet have a sufficiently rich theory of the constitution of the universe, including the Fundamental Construct Kit. In part that seems to be a consequence of the forms of mathematics known to physicists. The well known challenge presented by Anderson [1972] discussed in Sect. 10, below, supports this.

It may take many years of research to find out what is missing from current physics. Collecting phenomena that need to be explained, and trying as hard as possible to construct *detailed* explanations of those phenomena, including working models, is one way to make progress. That may pinpoint gaps in our theories and stimulate development of new, more powerful, theories. Compare the profound ways in which our understanding of possible forms of computation has been extended by unending attempts to put computation to new uses. Collecting examples of such challenges helps us assemble tests to be passed by future proposed theories: samples of possibilities that a deep physical theory needs to be able to explain.

Perhaps the most tendentious proposal here is that an expanded physical theory, instead of being expressed mainly in terms of equations relating measures, may need formalisms better suited to specification of a construction kit, perhaps sharing features of grammars, programming languages, partial orderings, topological relationships, architectural specifications, and the structural descriptions in chemistry. The theory will need to use appropriate kinds of mathematics for drawing out implications of the theories, including explanations of possibilities, both observed and unobserved, including possible future forms of intelligence. Theories of utility measures may need to be replaced, or enhanced with new theories of how benefits, evaluations, comparisons and preferences can be expressed (attempted in Sloman [1969]).

We must also avoid assuming optimality. Evolution produces designs as diverse as microbes, cockroaches, elephants and orchids, none of which is optimal or rational in any simple sense, yet many of them survive and sometimes proliferate, because they are lucky, at least for a while, as with human decisions, policies, preferences, cultures, etc.

8 Types and Levels of Explanation of Possibilities

Suppose someone uses a Meccano kit to construct a toy crane, with a jib that can be moved up and down by turning a handle, and a rotating platform on a fixed base that allows the direction of the jib to be changed. What's the difference between explaining how that is possible and how it was done? First of all, if nobody actually builds such a crane then there is no actual crane-building to be explained. Yet, insofar as the Meccano kit makes such cranes possible it makes sense to ask *how* it is possible. This has several types of answer, including answers at different levels of abstraction, with varying generality and economy of specification.

More generally, the question "How is it possible to create X using construction kit Y?", or, simply, "How is X possible?", has several types of answer, including answers at different levels of abstraction, with varying generality. I'll assume that a particular construction kit is referred to either explicitly or implicitly. The following is not intended to be an exhaustive survey of the possible types of answer. It is merely a first experimental foray, preparing the ground for future work:

1. **Structural conformity:** The first type of answer, structural conformity (grammaticality), merely identifies the parts and relationships between parts that are supported by the kit, showing that X (e.g. a crane of the sort in question) could be composed of such parts arranged in such relationships. An architect's drawings for a building, specifying materials, components, and their spatial and functional relations, would provide such an explanation of how a proposed building is possible, including, perhaps, answering questions about how the construction would make the building resistant to very high winds, or to earthquakes up to a specified strength. This can be compared with showing that a sentence is acceptable in a language with a well-defined grammar by showing how the sentence would be parsed (analysed) in accordance with the grammar of that language. A parse tree (or graph) also shows how the sentence can be built up piecemeal from words and other grammatical units by assembling various substructures and using them to build larger structures. Compare this with using a chemical diagram to show how a collection of atoms can make up a particular molecule, e.g. the ring structure of C_6H_6 (Benzene).

   Some structures are specified in terms of piecewise relations, where the whole structure cannot possibly exist, because the relations cannot hold simultaneously, e.g. X is above Y, Y is above Z, Z is above X. It is possible to depict such objects, e.g. in pictures of impossible objects by Reutersvard, Escher, Penrose, and others.[33] Some logicians and computer scientists have attempted to design languages in which specifications of impossible entities are necessarily syntactically ill-formed. This leads to impoverished languages with restricted practical uses, e.g. strongly typed programming languages. For some purposes less restricted languages, needing greater care in use, are preferable, including human languages, as I tried to show in Sloman [1971].

2. **Process possibility:** The second type of answer demonstrates constructability by describing a sequence of spatial trajectories by which such a collection of parts could be assembled. This may include processes of assembly of temporary scaffolding (Sect. 2.7) to hold parts in place before the connections have been made that make them self-supporting or before the final supporting structures have been built (as often happens in large engineering projects, such as bridge construction). Many different possible trajectories can lead to the same result. Describing (or demonstrating) any such trajectory explains both how that construction process is possible and how the end result is possible. There may be several different routes to the same end result.

   In some cases, a complex object has type 1 possibility although not type 2. For example, from a construction kit containing several rings it is possible to assemble a *pile* of three rings, but not possible to assemble a *chain* of three rings even though each of the parts of the chain is exactly like the parts of the pile.

3. **Process abstraction:** Some possibilities are described at a level of abstraction that ignores detailed routes through space, and covers *many* possible alternatives. For example, instead of specifying precise trajectories for parts as they are assembled, an explanation can specify the initial and final state of each trajectory, where each state-pair may be shared by a vast, or even infinite, collection of different possible trajectories producing the same end state, e.g. in a continuous space.

   In some cases, the possible trajectories for a moved component are all continuously deformable into one another (i.e. they are topologically equivalent); for example the many spatial routes by which a cup could be moved from a location where it rests on a table to a location where it rests on a saucer on the table, without leaving the volume of space above the table. Those trajectories form a continuum of possibilities that is too rich to be captured by a parametrized equation for a line with a number of variables. If trajectories include passing through holes, or leaving and entering the room via different doors or windows, then the different possible trajectories will not all be continuously deformable into one another: there are different equivalence classes of trajectories sharing common start and end states, for example, the different ways of threading a shoe lace with the same end result.

   The ability to abstract away from detailed differences between trajectories sharing start and end points, thereby implicitly recognizing invariant features of an infinite collection of possibilities, is an important aspect of animal intelligence

[33] http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html

that I don't think has been generally understood. Many researchers assume that intelligence involves finding *optimal* solutions. So they design mechanisms that search using an optimisation process, ignoring the possibility of mechanisms that can find sets of possible solutions (e.g. routes) initially considered as a class of *equivalent* options, leaving questions about optimal assembly to be settled later, if needed. I suspect these remarks are closely related to the origins of abilities to reason about geometry and topology.[34]

4. Grouping: Another form of abstraction is related to the difference between **1** (structures) and **2** (processes). If there is a sub-sequence of assembly processes whose order makes no difference to the end result, they can be grouped to form an unordered "composite" move containing an unordered set of moves. If N components are moved from initial to final states in a sequence of N moves, and it makes no difference in what order they are moved, merely specifying the set of N possibilities without regard for order collapses N factorial sets of possible sequences into one composite move. If N is 15, that will collapse 1,307,674,368,000 different sequences into one.

Sometimes a subset of moves can be made in parallel. For example someone with two hands can move two or more objects at a time while transferring a collection of items from one place to another. Parallelism is particularly important in many biological processes where different processes occurring in parallel constrain one another so as to ensure that instead of all the possible states that could occur by moving or assembling components separately, only those end states occur that are consistent with parallel constructions. In more complex cases, the end state may depend on the relative speeds of sub-processes and also on continuously changing spatial relationships. This is important in epigenesis, since all forms of development from a single cell to a multi-celled structure depend on many mutually constraining processes occurring in parallel.

For some construction kits, certain constructs made of a collection of subassemblies may require different subassemblies to be constructed in parallel if completing some too soon could make the required final configuration unachievable. For example, rings being completed before being joined could prevent formation of a chain.

5 Iterative or recursive abstraction: Some process types involve unspecified numbers of parts or steps, although each instance of the type has a definite number, for example a process of moving chairs by repeatedly carrying a chair to the next room until there are no chairs left to be carried, or building a tower from a collection of bricks, where the number of bricks can be varied. A specification that abstracts from the number can use a notion like "repeat until", or a recursive specification: a very old idea in mathematics, such as Euclid's

[34] Illustrated in these discussion notes:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/changing-affordances.html
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/triangle-theorem.html
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/torus.html

algorithm for finding the highest common factor of two numbers. Production of such a generic specification can demonstrate a large variety of possibilities inherent in a construction kit in an extremely powerful and economical way. Many new forms of abstraction of this type have been discovered by computer scientists developing programming languages, for operating not only on numbers but many other structures, e.g. trees and graphs.

Evolution may also have "discovered" many cases long before humans existed by taking advantage of mathematical structures inherent in the construction kits available and the trajectories by which parts can be assembled into larger wholes. This may be one of the ways in which evolution produced powerful new genomes, and reusable genome components that allowed many different biological assembly processes to result from a single discovery, or a few discoveries, at a high enough level of abstraction.

Some related abstractions may have resulted from parametrisation: processes by which details are removed from specifications in genomes and left to be provided by the context of development of individual organisms, including the physical or social environment. (See Sect. 2.3 on epigenesis.)

6 **Self-assembly:** If, unlike construction of a toy Meccano crane or a sentence or a sorting process, the process to be explained is a self-assembly process, like many biological processes, then the explanation of how the assembly is possible will not merely have to specify trajectories through space by which the parts become assembled, but also:

- what causes each of the movements (e.g. what manipulators are required?);
- where the energy required comes from (an internal store, or external supply?);
- whether the process involves pre-specified information about required steps or required end states, and, if so, what mechanisms can use that information to control the assembly process;
- how that prior information structure (e.g. specification of a goal state to be achieved, or plan specifying actions to be taken) came to exist, e.g. whether it was in the genome as a result of previous evolutionary transitions, or whether it was constructed by some planning or problem-solving mechanism in an individual, or whether it was provided by a communication from an external source;
- how these abilities can be acquired or improved by learning or reasoning processes or by random variation (if they can).

7. **Use of explicit intentions and plans:** None of the explanation-types above presupposes that the possibility being explained has ever been represented explicitly by the machines or organisms involved. Explaining the possibility of some structure or process that results from intentions or plans would require specifying pre-existing information about the end state and in some cases also intermediate states, namely information that existed before the process began — information that can be used to control the process (e.g. intentions, instructions, or

sub-goals, and preferences that help with selections between options). It seems that some of the reproductive mechanisms that depend on parental care make use of mechanisms that generate intentions and possibly also plans in carers, for instance intentions to bring food to an infant, intentions to build nests, intentions to carry an infant to a new nest, intention to migrate to another continent when the temperature drops, and many more. Use of intentions that can be carried out in multiple ways selected according to circumstances rather than automatically triggered reflexes could cover a far wider variety of cases, but would require provision of greater intelligence in individuals.

Sometimes an explanation of possibility prior to construction is important for engineering projects where something new is proposed and critics believe that the object in question could not exist, or could not be brought into existence using available known materials and techniques. The designer might answer sceptical critics by combining answers of any of the above types, depending on the reasons for the scepticism.

Concluding Comment on Explanations of Possibilities: Those are all examples of components of explanations of assembly processes, including self-assembly. In biological reproduction, growth, repair, development, and learning there are far more subdivisions to be considered, some of them already studied piecemeal in a variety of disciplines. In the case of human development, and to a lesser extent development in other species, there are many additional sub-cases involving construction kits both for creating information structures and for creating information-processing mechanisms of many kinds, including perception, learning, motive formation, motive comparison, intention formation, plan construction, plan execution, language use, and many more. A subset of cases, with further references can be found in Sloman [2006].

The different answers to “How is it possible to construct this type of object?” may be correct as far as they go, though some provide more detail than others. More subtle cases of explanations of possibility include differences between reproduction via egg-laying and reproduction via parturition, especially when followed by caring for offspring. The latter allows a parent’s influence to continue during development, as does teaching of younger individuals by older ones. This also allows development of cultures suited to different environments.

To conclude this rather messy section: the investigation of different types of generality in modes of explanation for possibilities supported by a construction kit is also relevant to modes of specification of new designs based on the kit. Finding economical forms of abstraction may have many benefits, including both reducing search spaces when trying to find a new design and also providing a generic design that covers a broad range of applications tailored to detailed requirements. Of particular relevance in a biological context is the need for designs that can be adjusted over time, e.g. during growth of an organism, or shared across species with slightly different physical features or environments. Many of the points made here are also related to structural changes in both computer programming languages and software design specification languages. Evolution may have beaten us to important ideas. That these levels of abstraction are possible is a metaphysical feature of the universe, implied by the generality of the FCK.

9 Alan Turing's Construction Kits

Turing [1936] showed that a rather simple sort of machine, now known as a Turing machine, could be used to specify an infinite set of constructions with surprisingly rich mathematical features. The set of possibilities was infinite because a Turing machine is defined to have an infinite (or indefinitely extendable) linear "tape" divided into discrete locations in which symbols can be inserted. A feature of a Turing machine that is not in most other construction kits is that it can be set up and then started, after which it will modify initial structures and build new ones, possibly indefinitely, though in some cases the machine will eventually halt.

Another type of construction kit with related properties is Conway's Game of Life,[35] a construction kit that creates changing patterns in 2D regular arrays. Stephen Wolfram has written a great deal about the diversity of constructions that can be explored using such cellular automata. Neither a Turing machine nor a Conway game has any external sensors: once started they run according to their stored rules and the current (changing) state of the tape or grid cells. In principle, either of them could be attached to external sensors able to produce changes to the tape of a Turing machine or the states of some of the cells in the Life array. However, any such extension would significantly alter the powers of the machine, and theorems about what such a machine could or could not do would change.

Modern computers use a variant of the Turing machine idea, where each computer has a finite memory but with the advantage of much more direct access between the central computer mechanism and the locations in the memory (a von Neumann architecture). Increasingly, computers have also been provided with a variety of external interfaces connected to sensors or motors so that while running they can acquire information (e.g. from keyboards, buttons, joysticks, mice, electronic piano keyboards, network connections, and many more) and can also send signals to external devices. Theorems about disconnected Turing machines may not apply to machines with rich two-way interfaces connected to the environment.

Turing machines and Game of Life machines can be described as "self-propelling" because, once set up, they can be left to run according to the general instructions they have and the initial configuration on the tape or in the array. But they are not really self-propelling: they have to be implemented in physical machines with an external power supply. In contrast, Ganti [2003] shows how the use of chemistry as a construction kit provides "self-propulsion" for living things, though every now and

[35] http://en.wikipedia.org/wiki/Conway.27s.Game.of.Life

again the chemicals need to be replenished. A battery driven computer is a bit like that, but someone else has to make the battery.

Living things make and maintain themselves, at least after being given a kick-start by their parent or parents. They do need constant, or at least frequent, external inputs, but, for the simplest organisms, those are only chemicals in the environment and energy from chemicals or heat-energy via radiation, conduction or convection. John McCarthy pointed out in a conversation that some animals also use externally supplied mechanical energy, e.g. rising air currents used by birds that soar. Unlike pollen grains, spores, etc. propagated by wind or water, the birds use internal information-processing mechanisms to control how the wind energy is used, as does a human piloting a glider.

9.1 Beyond Turing Machines: Chemistry

Turing also explored other sorts of construction kits, including types of neural net and extended versions of Turing machines with "oracles" added. Shortly before his death (in 1954), he published Turing [1952], in which he explored a type of pattern-forming construction kit in which two chemical substances can diffuse through the body of an expanding organism and interact strongly wherever they meet. He showed that that sort of construction kit could generate many of the types of surface pattern observed on plants and animals. I have been trying to show how that can be seen as a very simple example of something far more general.

One of the important differences between types of construction kit mentioned above is the difference between kits supporting only discrete changes, e.g. to a first approximation Lego and Meccano (ignoring variable length strings and variable angle joints), and kits supporting continuous variation, e.g. plasticine and mud (ignoring, for now, the discreteness at the molecular level).

One of the implications of such differences is how they affect abilities to search for solutions to problems. If only big changes in design are possible, the precise change needed to solve a problem may be inaccessible (as many who have played with construction kits will have noticed). On the other hand, if the kit allows arbitrarily small changes, it will, in principle, permit exhaustive searches in some sub-spaces. The exhaustiveness comes at the cost of a very much larger (infinite, or potentially infinite!) search-space. That feature could be useless, unless the space of requirements has a structure that allows approximate solutions to be useful. In that case, a mixture of big jumps to get close to a good solution, followed by small jumps to home in on a (locally) optimal solution can be very fruitful—a technique that has been used by Artificial Intelligence researchers, called "simulated annealing".[36]

[36] One of many online explanations is at
http://www.theprojectspot.com/tutorial-post/simulated-annealing-algorithm-for-beginners/6

Wagner [2014] claims that the structure of the search space generated by the molecules making up the genome increases the chance of useful approximate solutions to important problems to be found with *relatively* little searching (compared with other search spaces), after which small random changes allow improvements to be found. I have not yet read the book, but it seems to illustrate the importance for evolution of the types of construction-kit available.[37] I have not yet had time to check whether the book discusses uses of abstraction and the evolution of mathematical and meta-mathematical competences discussed here. Nevertheless, it seems to be an (unwitting) contribution to the Meta-Morphogenesis project. Recent work by Jeremy England at MIT[38] may turn out also to be relevant.

9.2 Using Properties of a Construction-Kit to Explain Possibilities

A formal axiomatic system can be seen as an abstract construction kit with axioms and rules that support construction of proofs that end in theorems. The theorems are formulae that can occur at the end of a proof using only axioms and inference rules in the system. The kit explains the possibility of some theorems based on the axioms and rules. The non-theorems of an axiomatic system are formulae for which no such proof exists. Proving that something is a non-theorem can be difficult, and requires a proof in a meta-system.

Likewise, a physical construction kit can be used to demonstrate that some complex physical objects can occur at the end of a construction process. In some cases there are objects that are describable but cannot occur in a construction using that kit: e.g. an object whose outer boundary is a surface that is everywhere curved cannot be produced in a construction based on Lego bricks or a Meccano set, though one could occur in a construction based on plasticine or soap-film.

9.3 Bounded and Unbounded Construction Kits

A rectangular grid of squares combined with the single digit numbers, 0,1,...,9 (strictly numerals representing numbers), allows construction of a set of configurations in which numbers are inserted into the squares subject to various constraints, e.g. whether some squares can be left blank, whether certain pairs of numbers can be adjacent, or whether the same number can occur in more than one square. For a given grid and a given set of constraints, there will be a finite set of possible configurations (although it may be a very large set). If, in addition to insertion of a number, the "construction kit" allows extra empty rows or columns to be added to the grid no matter how large it is, then the set of possible configurations

[37] An interview with the author is online at https://www.youtube.com/watch?v=wyQgCMZdv6E

[38] https://www.quantamagazine.org/20140122-a-new-physics-theory-of-life/

becomes infinite. Many types of infinite construction kit have been investigated by mathematicians, logicians, linguists, computer scientists, musicians and other artists.

Analysis of chemistry-based construction kits for information-processing systems would range over a far larger class of possible systems than Turing machines (or digital computers), because of the mixture of discrete and continuous changes possible when molecules interact, e.g. moving together, moving apart, folding, and twisting, but also locking and unlocking, using catalysts Kauffman [1995]. I don't know whether anyone has a deep theory of the scope and limits of chemistry-based information processing.

Recent discoveries indicate that some biological mechanisms use quantum-mechanical features of the FCK that we do not yet fully understand, providing forms of information processing that are very different from what current computers do. For example a presentation by Seth Lloyd summarises quantum phenomena used in deep-sea photosynthesis, avian navigation, and odour classification.[39] This may turn out to be the tip of the iceberg of quantum-based information-processing mechanisms.

There are some unsolved, very hard, partly ill-defined, problems about the variety of functions of biological vision, e.g. simultaneously interpreting a very large, varied and changing collection of visual fragments, perceived from constantly varying viewpoints as you walk through a garden with many unfamiliar flowers, shrubs, bushes, etc. moving irregularly in a changing breeze. Could some combination of quantum entanglement and non-local interaction play a role in rapidly and simultaneously processing a large collection of mutual constraints between multiple visual fragments? The ideas are not yet ready for publication, but work in progress is recorded here: http://www.cs.bham.ac.uk/research/projects/cogaff/misc/quantum-evolution.html.

Some related questions about perception of videos of fairly complex moving plant structures are raised here: http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vision/plants/.

10 Conclusion: Construction Kits for Meta-Morphogenesis

A useful two-part survey by Keller of previous attempts to show how life and its products relate to the physical world [2008, 2009], concluded that attempts so far have not been successful. Keller ends with the suggestion that the traditional theory of dynamical systems is inadequate for dealing with constructive processes and needs to be expanded to include "objects, their internal properties, their construction, and their dynamics", i.e. a theory of *"Constructive dynamical systems"*. This paper outlines a project to do that and more, including giving an account of branching layers of new *derived* construction kits produced by evolution, development and other processes. The physical world clearly provides a very powerful (chemistry-based) fundamental

[39] https://www.youtube.com/watch?v=wcXSpXyZVuY

construction kit that, together with natural selection processes and processes within individuals as they develop, produced an enormous variety of organisms on this planet, based on additional derived construction kits (DCKs), including concrete, abstract and hybrid construction kits and, most recently, new sorts of construction kit used as a toy or an engineering resource.

The idea of a construction kit is offered as a new unifying concept for Philosophy of mathematics, Philosophy of science, Philosophy of biology, Philosophy of mind and Metaphysics. The aim is to explain how it is possible for minds to exist in a material world and to be produced by natural selection and its products. Related questions have been raised about the nature of mathematics and its role in life. The ideas are still at an early stage of development and there are probably many more distinctions to be made, and a need for a more formal, mathematical presentation of properties of and relationships between construction kits, including the ways in which new derived construction kits can be related to their predecessors and their successors. The many new types of computer-based *virtual* machinery produced by human engineers since around 1950 provide examples of non-reductive supervenience (as explained in Sloman [2013b]). They are also useful as relatively simple examples to be compared with far more complex products of evolution.

In Esfeld et al. [2015], a distinction is made between two "principled" options for the relationship between the basic constituents of the world and their consequences. In the "Humean" option there is nothing but the distribution of structures and processes over space and time, though there may be some empirically discernible patterns in that distribution. The second option is "modal realism", or "dispositionalism", according to which there is something about the primitive stuff and its role in space-time that constrains what can and cannot exist, and what types of process can or cannot occur. I am arguing for a "multi-layer" version of the modal realist option (developing ideas in Sloman [1962, 1996a, 2013b]).

I suspect that a more complete development of this form of modal realism can contribute to answering the problem famously posed by Anderson in [1972], namely how we should understand the relationships between different levels of complexity in the universe (or in scientific theories). The reductionist alternative claims that when the physics of elementary particles (or some other fundamental physical level) has been fully understood, everything else in the universe can be explained in terms of mathematically derivable consequences of the basic physics. Anderson contrasts this with the anti-reductionist view that different levels of complexity in the universe require "entirely new laws, concepts and generalisations" so that, for example, biology is not applied chemistry and psychology is not applied biology. He writes: "Surely there are more levels of organization between human ethology and DNA than there are between DNA and quantum electrodynamics, and each level can require a whole new conceptual structure". However, the structural levels are not merely in the concepts used by scientists, but actually in the world.

We still have much to learn about the powers of the fundamental construction kit (FCK), including: (1) the details of how those powers came to be used for life on earth, (2) what sorts of derived construction kit (DCK) were required in order to make more complex life forms possible, (3) how those construction kits support "blind" mathematical discovery by evolution, mathematical competences in humans and other animals and, eventually, meta-mathematical competences, then meta-meta-mathematical competences, at least in humans, (4) what possibilities the FCK has that have not yet been realised, (5) whether and how some version of the FCK could be used to extend the intelligence of current robots, (6) whether currently used Turing-equivalent forms of computation have at least the same information-processing potentialities (e.g. abilities to support all the biological information-processing mechanisms and architectures), and (7) if those forms of computation lack the potential, then how are biological forms of information processing different? Don't expect complete answers soon.

In future, physicists wishing to show the superiority of their theories should attempt to demonstrate mathematically and experimentally that they can explain more of the potential of the FCK to support varied construction kits required for, and produced by, biological evolution than rival theories can. Will that be cheaper than building bigger, better colliders? Will it be harder?[40]

End Note

As I was finishing off this paper, I came across a letter Turing wrote to W. Ross Ashby in 1946 urging Ashby to use Turing's ACE computer to implement his ideas about modelling brains. Turing expressed a view that seems to be unfashionable among AI researchers at present (2015), but accords with the aims of this paper:

> *"In working on the ACE I am more interested in the possibility of producing models of the actions of the brain than in the practical applications to computing."*
> http://www.rossashby.info/letters/turing.html

It would be very interesting to know whether he had ever considered the question whether digital computers might be incapable of accurately modelling brains making deep use of chemical processes. He also wrote in Turing [1950], *"In the nervous system chemical phenomena are at least as important as electrical."* But he did not elaborate on the implications of that claim.[41]

Acknowledgements My work on the Meta-Morphogenesis project, including this paper, owes a great dept to Barry Cooper, who unfortunately died in October 2015.

[40] Here's a cartoon teasing particle physicists: http://www.smbc-comics.com/?id=3554

[41] I think it will turn out that the ideas about "making possible" used here are closely related to Alastair Wilson's ideas about grounding as "metaphysical causation". Wilson [2015].

I had never met him until we both contributed chapters to a book published in 2011 on Information and Computation. Barry and I first met, by email, when we reviewed each others' chapters. Later, out of the blue, he invited me to contribute to the Turing centenary volume he was co-editing Cooper and van Leeuwen [2013]. I contributed three papers. He then asked me for a contribution to Part 4 (on Emergence and Morphogenesis) based on Turing's paper on morphogenesis published in 1952, 2 years before he died. That got me wondering what Turing might have done if he had lived another 30-40 years. So I offered Barry a paper proposing "The Meta-Morphogenesis Project" as an answer. He accepted it (as the final commentary paper in the book) and ever since then I have been working on the project. He later encouraged me further by inviting me to give talks and to contribute a chapter to this book. As a result we had several very enjoyable conversations. He changed my life by giving me a new research direction, which does not often happen to 75-year old retired academics! (Now 5 years older.) I wish we could continue our conversations. I also owe much to the highly intelligent squirrels and magpies in our garden, who have humbled me.

Note added May 2017
My ideas have probably been influenced in more ways than I recognise by my colleague Margaret Boden, whose work has linked AI/Cognitive Science to Biology since the 1960s, notably in her extraordinary two volume survey published in 2006 by OUP, *Mind As Machine: A history of Cognitive Science*.

Ongoing work
This paper is a snapshot (late in 2016) of a large and complex project, that continues to grow. Examples of continuing questioning, conjecturing, exploring, etc. can be found in

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/construction-kits.html

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/maths-multiple-foundations.html

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/anger.html

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/family-resemblance-vs-polymorphism.html

http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html

References

P.W. Anderson. More is different. *Science, New Series*, 177(4047):393–396, Aug 1972. http://robotics.cs.tamu.edu/dshell/cs689/papers/anderson72more_is_different.pdf.

W. Ross Ashby. *Design for a Brain*. Chapman and Hall, London, 1952.

W. Ross Ashby. *An Introduction to Cybernetics,*. Chapman & Hall, 1956. http://pespmc1.vub.ac.be/ASHBBOOK.html.

L.P. Beaudoin. *Goal processing in autonomous agents*. PhD thesis, School of Computer Science, The University of Birmingham, Birmingham, UK, 1994. http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#38.

L.P. Beaudoin and A. Sloman. A study of motive processing and attention. In A. Sloman, D. Hogg, G. Humphreys, D. Partridge, and A. Ramsay, editors, *Prospects for Artificial Intelligence*, pages 229–238. IOS Press, Amsterdam, 1993. http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#16.

Graham Bell. *Selection The Mechanism of Evolution*. OUP, 2008. Second Edition.

Karen Bennett. Construction Zone: No Hard Hat Required. *Philosophical Studies*, 154:79–104, 2011. doi: 10.1007/s11098-011-9703-8.

J.M. Chappell and A. Sloman. Natural and artificial meta-configured altricial information-processing systems. *International Journal of Unconventional Computing*, 3(3):211–239, 2007. http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#717.

N. Chomsky. *Aspects of the theory of syntax*. MIT Press, Cambridge, MA, 1965.

Juliet Coates, Umm-E-Aiman, and Benedicte Charrier. Understanding ”green” multicellularity: do seaweeds hold the key? *Frontiers in Plant Science*, Dec 2014. doi: 10.3389/fpls.2014.00737.

S. Barry Cooper and J. van Leeuwen, editors. *Alan Turing: His Work and Impact*. Elsevier, Amsterdam, 2013. ISBN 9780123869807.

Terence W. Deacon. *Incomplete Nature: How Mind Emerged from Matter*. W. W. Norton & Company, 2011. ISBN 978-0-393-04991-6.

Michael Esfeld, Dustin Lazarovici, Vincent Lam, and Mario Hubert. The Physics and Metaphysics of Primitive Stuff. *British Journal for the Philosophy of Science*, 67(4):1–29, 2015. doi = 10.1093/bjps/axv026.

Chrisantha Fernando. Review of “*The Principles of Life*” by Tibor Ganti. (2003, Oxford University Press.). *Artificial Life*, 14(4):467–470, Fall 2008. doi: 10.1162/artl.2008.14.4.14404.

Tibor Ganti. *The Principles of Life*. OUP, New York, 2003. Eds. Eörs Szathmáry & James Griesemer, Translation of the 1971 Hungarian edition.

J. J. Gibson. *The Senses Considered as Perceptual Systems*. Houghton Mifflin, Boston, 1966.

J. J. Gibson. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA, 1979.

I. Kant. *Critique of Pure Reason*. Macmillan, London, 1781. Translated (1929) by Norman Kemp Smith.

A Karmiloff-Smith. *Beyond Modularity: A Developmental Perspective on Cognitive Science*. MIT Press, Cambridge, MA, 1992.

S. Kauffman. *At home in the universe: The search for laws of complexity*. Penguin Books, London, 1995.

Evelyn Fox Keller. Organisms, Machines, and Thunderstorms: A History of Self-Organization, Part One. *Historical Studies in the Natural Sciences,*, 38(1 (Winter)): 45–75, 2008. http://www.jstor.org/stable/10.1525/hsns.2008.38.1.45.

Evelyn Fox Keller. Organisms, Machines, and Thunderstorms: A History of Self-Organization, Part Two. Complexity, Emergence, and Stable Attractors. *Historical Studies in the Natural Sciences*, 39(1 (Winter)):1–31, 2009. http://www.jstor.org/stable/10.1525/hsns.2009.39.1.1.

Gert Korthof. Review of *The Principles of Life* by Tibor Ganti, Dec 2003. Updated 6 Oct 2014, http://wasdarwinwrong.com/korthof66.htm.

J.E. Laird, A. Newell, and P.S. Rosenbloom. SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33:1–64, 1987.

I. Lakatos. Falsification and the methodology of scientific research programmes. In J. Worrall and G. Currie, editors, *Philosophical papers, Vol I*, pages 8–101. Cambridge University Press, Cambridge, 1980. ISBN 0-521-28031-1.

R.D. Luce and H. Raiffa. *Games and Decisions: Introduction and CriticalSurvey*. John Wiley and Sons, Inc.; Chapman and Hall, New York; London, 1957.

C. Mathis, T. Bhattacharya, and S.I. Walker. The Emergence of Life as a First Order Phase Transition. Technical report, Arizona State University, Oct 2015. http://arxiv.org/pdf/1503.02776.

J. McCarthy. Ascribing mental qualities to machines. In M. Ringle, editor, *Philosophical Perspectives in Artificial Intelligence*, pages 161–195. Humanities Press, Atlantic Highlands, NJ, 1979. http://www-formal.stanford.edu/jmc/ascribing/ascribing.html.

J. McCarthy and P.J. Hayes. Some philosophical problems from the standpoint of AI. In B. Meltzer and D. Michie, editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press, Edinburgh, Scotland, 1969. http://www-formal.stanford.edu/jmc/mcchay69/mcchay69.html.

M. L. Minsky. *The Society of Mind*. William Heinemann Ltd., London, 1987.

M. L. Minsky. *The Emotion Machine*. Pantheon, New York, 2006.

Roger Penrose. *Shadows of the mind: A Search for the Missing Science of Consciousness*. OUP, Oxford, 1994.

K.R. Popper. *The logic of scientific discovery*. Routledge, London, 1934.

K.R. Popper. *Unended Quest*. Fontana/Collins, Glasgow, 1976.

K.R. Popper. Natural Selection and the Emergence of Mind. *Dialectica*, 32(3-4): 339–355, 1978. ISSN 1746-8361. http://dx.doi.org/10.1111/j.1746-8361.1978.tb01321.x.

W. T. Powers. *Behavior, the Control of Perception*. Aldine de Gruyter, New York, 1973.

Erwin Schrödinger. *What is life?* CUP, Cambridge, 1944.

Ann Senghas. Language Emergence: Clues from a New Bedouin Sign Language. *Current Biology*, 15(12):R463–R465, 2005. http://dx.doi.org/10.1016/j.cub.2005.06.018.

Claude Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 and 623–656, July and October 1948.

H. A. Simon. Motivational and emotional controls of cognition. In H. A. Simon, editor, *reprinted in Models of Thought*, pages 29–38. Yale University Press, Newhaven, CT, 1967.

A. Sloman. *Knowing and Understanding: Relations between meaning and truth, meaning and necessary truth, meaning and synthetic necessary truth (DPhil Thesis)*. PhD thesis, Oxford University, 1962. http://www.cs.bham.ac.uk/research/projects/cogaff/07.html#706.

A. Sloman. How to derive "better" from "is". *American Phil. Quarterly*, 6:43–52, Jan 1969. http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1969-02.

A. Sloman. "Ought" and "better". *Mind*, LXXIX(315):385–394, July 1970. http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1970-01.

A. Sloman. Tarski, Frege and the Liar Paradox. *Philosophy*, 46(176):133–147, April 1971. http://www.cs.bham.ac.uk/research/projects/cogaff/62-80.html#1971-03.

A. Sloman. *The Computer Revolution in Philosophy*. Harvester Press (and Humanities Press), Hassocks, Sussex, 1978. http://www.cs.bham.ac.uk/research/cogaff/62-80.html#crp, Revised 2015.

A. Sloman. Image interpretation: The way ahead? In O.J. Braddick and A.C. Sleigh., editors, *Physical and Biological Processing of Images (Proceedings of an international symposium organised by The Rank Prize Funds, London, 1982.)*, pages 380–401. Springer-Verlag, Berlin, 1983. http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#57.

A. Sloman. The mind as a control system. In C. Hookway and D. Peterson, editors, *Philosophy and the Cognitive Sciences*, pages 69–110. Cambridge University Press, Cambridge, UK, 1993. http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#18.

A. Sloman. Actual possibilities. In L.C. Aiello and S.C. Shapiro, editors, *Principles of Knowledge Representation and Reasoning: Proc. 5th Int. Conf. (KR '96)*, pages 627–638, Boston, MA, 1996a. Morgan Kaufmann Publishers. http://www.cs.bham.ac.uk/research/cogaff/96-99.html#15.

A. Sloman. The SimAgent TOOLKIT – for Philosophers and Engineers (And Some Biologists, Psychologists and Social Scientists), 1996b. http://www.cs.bham.ac.uk/research/projects/poplog/packages/simagent.html.

A. Sloman. The Cognition and Affect Project: Architectures, Architecture-Schemas, And The New Science of Mind. Technical report, School of Computer Science, University of Birmingham, Birmingham, UK, 2003. http://www.cs.bham.ac.uk/research/projects/cogaff/03.html#200307, (Revised August 2008).

A. Sloman. Requirements for a Fully Deliberative Architecture (Or component of an architecture). Research Note COSY-DP-0604, School of Computer Science, University of Birmingham, Birmingham, UK, May 2006. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/fully-deliberative.html.

A. Sloman. Architecture-Based Motivation vs Reward-Based Motivation. *Newsletter on Philosophy and Computers*, 09(1):10–13, 2009. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/architecture-based-motivation.html.

A. Sloman. How Virtual Machinery Can Bridge the "Explanatory Gap", In Natural and Artificial Systems. In S. Doncieux and et al., editors, *Proceedings SAB 2010, LNAI 6226*, pages 13–24, Heidelberg, August 2010. Springer. http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#sab.

A. Sloman. What's information, for an organism or intelligent machine? How can a machine or organism mean? In G. Dodig-Crnkovic and M. Burgin, editors, *Information and Computation*, pages 393–438. World Scientific, New Jersey, 2011. http://www.cs.bham.ac.uk/research/projects/cogaff/09.html#905.

A. Sloman. Virtual machinery and evolution of mind (part 3) meta-morphogenesis: Evolution of information-processing machinery. In S. B. Cooper and J. van Leeuwen, editors, *Alan Turing - His Work and Impact*, pages 849–856. Elsevier, Amsterdam, 2013a. ISBN 9780123869807. http://www.cs.bham.ac.uk/research/projects/cogaff/11.html#1106d.

A. Sloman. Virtual Machine Functionalism (The only form of functionalism worth taking seriously in Philosophy of Mind and theories of Consciousness). Research note, School of Computer Science, The University of Birmingham, 2013b. http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html.

A. Sloman. Evolution of minds and languages. What evolved first and develops first in children: Languages for communicating, or languages for thinking (Generalised Languages: GLs)?, 2015. http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang.

A. Sloman and R.L. Chrisley. Virtual machines and consciousness. *Journal of Consciousness Studies*, 10(4-5):113–172, 2003. http://www.cs.bham.ac.uk/research/projects/cogaff/03.html#200302.

R. Sun. The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In Ron Sun, editor, *Cognition and Multi-Agent Interaction*, pages 79–99. Cambridge University Press, New York, 2006. http://www.cogsci.rpi.edu/~rsun/sun.clarion2005.pdf.

D'Arcy W. Thompson. *On Growth and Form*. Cambridge University Press, Cambridge, 1917. (Revised Edition 1948).

A. Trehub. *The Cognitive Brain*. MIT Press, Cambridge, MA, 1991. http://people.umass.edu/trehub/.

A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proc. London Math. Soc.*, 42(2):230–265, 1936. http://www.thocp.net/biographies/papers/turing_oncomputablenumbers_1936.pdf.

A. M. Turing. Computing machinery and intelligence. *Mind*, 59:433–460, 1950. (reprinted in E.A. Feigenbaum and J. Feldman (eds) *Computers and Thought* McGraw-Hill, New York, 1963, 11–35).

A. M. Turing. The Chemical Basis Of Morphogenesis. *Phil. Trans. R. Soc. London B 237*, 237:37–72, 1952.

C. H. Waddington. *The Strategy of the Genes*. MacMillan, 1957.

Andreas Wagner. *Arrival of the Fittest: Solving Evolution's Greatest Puzzle*. Published by: Oneworld Publications, 2014.

A A S Weir, J Chappell, and A Kacelnik. Shaping of hooks in New Caledonian crows. *Science*, 297(9 August 2002):981, 2002.

N. Wiener. *Cybernetics: or Control and Communication in the Animal and the Machine*. The MIT Press, Cambridge, MA, 1961. 2nd edition.

Alastair Wilson. Metaphysical Causation. Technical report, University of Birmingham, Department of Philosopy, July 2015. http://alastairwilson.org/files/mc120715.pdf, (Video lecture: https://www.youtube.com/watch?v=j2l1wKrtlxs).

I.P. Wright. *Emotional agents*. PhD thesis, School of Computer Science, The University of Birmingham, 1997. http://www.cs.bham.ac.uk/research/projects/cogaff/96-99.html#2.

I.P. Wright, A. Sloman, and L.P. Beaudoin. Towards a design-based analysis of emotional episodes. *Philosophy Psychiatry and Psychology*, 3(2):101–126, 1996. http://www.cs.bham.ac.uk/research/projects/cogaff/96-99.html#22.