



A Possible Genome To Architecture Project (GenToA) [The Meta-Genome Project?]

Installed: 2 Aug 2010

Last updated: 4 Aug 2010; 15 Aug 2010; 16 Aug 2010; 17 Aug 2010; 25 Oct 2017(format)

Also available as PDF file [here](#).

How can a genome specify an information-processing architecture that grows itself guided by interaction with the environment?

CONTENTS

- [Prologue \(Added 17 Aug 2010\)](#)
 - [Transitions to be understood](#)
 - [More on the book-decoding analogy](#)
 - [Virtual machinery](#)
 - [Some notes on the question](#)
 - [First meeting: Wed 18th August, room 124 School of Computer Science 2pm.](#)
 - [The Problem\(s\): what can biological reproduction produce, and how?](#)
 - [A KEY IDEA -- Development of layers of competence, using a layered architecture](#)
 - [Updated version of the 2007 diagram, 2017](#)
 - [Triggering reorganization](#)
 - [The need for new formalisms and new mathematics](#)
 - [Notes on varieties of products of reproduction](#)
 - [Reproductive and other pathologies](#)
 - [Related work \(To be Extended\)](#)
 - [Background](#)
-

Prologue (Added 17 Aug 2010) (Draft, to be rewritten)

One of several inspiring people I had the privilege of meeting during the 27 years I worked at the University of Sussex was John Maynard Smith -- one of the leading theoretical biologists of our time.

I once heard him express concern about the Human Genome project, comparing it with buying a book written in a language nobody understands. A slightly different expression of doubt would be a comparison with paying for a large computer program for which there is no compiler or interpreter that can be used to run the program and experiment with effects of changing it.

However, comparing a genome with a book or with what most people think computer programs are, may be seriously misleading, if the comparison assumes that books are sequences of meaningful chunks whose interpretation can be assembled to form a story or argument or theory, or assumes that a program is a sequence of instructions to be run (possibly including loops and jumps) to produce some behaviour.

[Verónica Arriola Ríos](#) added that trying to understand the genome is more like trying to reverse engineer a compiled complex software system, without knowing in which language it was written nor the details of the architecture on which it runs. This is far more complex than trying to decipher a book, because not only is the language unknown, but it can not be interpreted in a linear way. (I am not suggesting that Maynard Smith would have disagreed with that.)

Transitions to be understood

Maynard Smith and Szathm'ary wrote a book proposing that there are eight major transitions in evolution, shown here: http://en.wikipedia.org/wiki/The_Major_Transitions_in_Evolution

I think that by extending the enquiry we can find principles for seeking out additional transitions that need to be understood as achievements of evolution, which added extra complexity to the capabilities of organisms, and in many cases extra complexity to the relationship between genome and individual organism.

E.g. We can identify many discontinuities of **function** and discontinuities of **designs** (or implementations) for achieving those functions among biological organisms. Some of those discontinuities occur in evolutionary trajectories and some in development of individuals. Perhaps some occur in both.

Studying those discontinuities, and, where appropriate, sequences of small changes that produce large qualitative differences in function or in design may help us to understand better what relationships can hold between a genome, a developmental context, and the information processing architectures in individuals with the genome.

In general those differences will relate to differences between environments in which organisms live, which can include results of previous evolutionary or developmental changes. So we need to study environments and the information processing problems they pose.

(Often AI research, or experiments on animals or humans, abstract away from most of the complexities of the environment, leading to discrepancies between claims and actual achievements.)

(Slides enlarging on this are in preparation.)

More on the book-decoding analogy

The book analogy could be strengthened (with inevitable obfuscation) by comparing the genome with a book that does not tell a single story, but can generate a host of different stories by generating, in collaboration with other sources of information, a host of different books with instructions for composing various story fragments, composed in collaboration with other information sources, and also books with instructions for collaboratively creating additional books for composing story fragments and books that give instructions for putting all the fragments together, where the instructions are also produced collaboratively, using not only the original book but also many of its products and many other things encountered in the environment.

In all these cases the collaboration is not merely with other books but with all sorts of things that already exist, including products of previous construction processes producing books that previously led to books that had inspired previous construction processes!

An important factor in the analogy is that the process of using the starting book (the genome) does not merely produce various intermediate books and eventually a mobile machine -- but also produces other things in the environment, referred to as 'The extended phenotype' by Dawkins. This can include nests built by birds, dams built by beavers, cathedrals built by termites and many more. In some cases the constructions produced by different individuals or groups of individuals sharing a genome all have common features (like termite cathedrals). In other cases the changes to the environment are different for different individuals, like the different sorts of houses humans choose to build, even in the same location.

Understanding what is expressed in a book that plays such a complex, multi-faceted, multi-stage role in construction processes is much more complex than decoding something like a part of the Rosetta stone by using knowledge of the portion of the world referred to and comparing it with what is said, in other known languages, on other parts of the stone.

It is also more complex than decoding a complex 'legacy' software system whose source code has been lost and whose designers are no longer available. "De-compiling" program code is often intractable.

How can we make progress in the task of finding out about how such a system works? The sort of thing I have described would be far too complex to 'decode' by looking for statistical correlations between occurrences of fragments in the initial 'book' and various things that are created (or bungled) when something uses the book to make what the book specifies. (Which is not to say that looking for such correlations is a complete waste of time.)

A way to facilitate progress is to spend much effort looking not just at the book (or the chemical expression of the genome) nor at the products of the processes instigated by such a book, but also at many aspects of the environment at different times in the evolutionary history of the genome, in particular the many challenges that those environments posed for individuals, and various groups of individuals (including in some cases individuals of different species that interact -- whether competitively, as predators and prey, or symbiotically) in the ancestry of the genome under investigation. We also need to look at other books and their products to find out what sorts of

differences can occur in both genomes of different organisms and also what differences occur in the physical forms, observable behaviours, and conjectured competences of those organisms.

In particular, it may turn out that certain features of the environment cause the same 'discoveries' to be made, or problems to be solved on different occasions, in different evolutionary lineages. An example would be multiple evolution of mechanisms for acquiring, storing then later using information about where things are located in a large and complex space.

A prerequisite for this investigation is learning how to think about complex designs for working systems with complex constraints and requirements to be met. However, in addition to the modes of design and construction and the types of requirements engineers already know about, we shall have to learn about new sorts of requirement, often subtle and unobvious functions that organisms or parts of organisms, or behaviours of organisms can perform. And we shall certainly have to learn about new sorts of mechanism, including the extraordinary information processing mechanisms implemented in chemical subsystems in living things -- for instance the mechanisms used to construct brains which cannot use information processing in a brain, if the required portions of the brain have not yet been built.

Virtual machinery

One of the most important ideas that has come out of research and development work on many aspects of hardware and software required for complex computing systems in the last half century is the idea of a **virtual machine**, that 'runs' on a physical machine, making use of resources provided by the physical machine, and which can perform complex tasks (often tasks with parallel subtasks) whose nature is completely different from the tasks of the physical machinery. Virtual machinery can handle communications with other machines, send and receive email, check email and web pages for malware, run spelling correctors, organise file systems, check for intruders, monitor activities by different users of the system, play chess, and detect and compensate for certain sorts of hardware faults, learn how to improve their use of resources, and many more.

There are several reasons why the use of virtual machinery can enhance the power and flexibility of complex information processing systems, including the fact that they are in many cases easier to modify, monitor and control, than the far more complex underlying physical machinery. It seems likely (though research on this is needed) that that was 'discovered' by biological evolution long before human scientists and engineers understood the issues. If that is right, part of the answer to our main question will be that the genome needs to specify features of virtual as well as physical machinery, including virtual machines that grow themselves in ways that depend on the environment with which they interact.

So one of the implications is that there is a need to relate the many advances in research on epigenetic processes at physical, chemical and neural levels to research on virtual information-processing machinery that depends on and uses those processes.

Despite the use of the unfortunate word 'virtual', these machines are **real**, and can perform real tasks, and have real effects, including controlling physical devices like computer memory, robot arms, and visual displays. So they are not epiphenomena. Talking about them is not just loose metaphorical talk. Software engineers do no work on imaginary systems. ('Virtual reality' systems are another topic for discussion in this context.)

For a discussion of some implications of this see

<http://www.cs.bham.ac.uk/research/projects/cogaff/10.html#sab>

How Virtual Machinery Can Bridge the "Explanatory Gap", In Natural and Artificial Systems (Paper for SAB2010).

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#darwin>

Talk 80: Helping Darwin: How to Think About Evolution of Consciousness

Or: "How could evolution get ghosts into machines?"

Added 2013:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/vm-functionalism.html>

Virtual Machine Functionalism (VMF)

(The only form of functionalism worth taking seriously in Philosophy of Mind and theories of Consciousness)

Some notes on the question (first draft, to be reorganised)

All living organisms are informed control systems: they use information to control their own growth and repair and to control their behaviours, whether interacting with merely physical, biological, social, or other types of environment.

What the information is about, how it is represented, where it comes from, how it is processed, whether it is stored for future use, how it is stored, how it is used, what mechanisms are needed for all this --- can vary enormously between different types of organism, and to some extent between individual organisms of the same time, or even an individual organism at different times. (Think of a human baby that grows up to be a quantum physicist, a professional baseball player, a ballet dancer, a shepherd, ...)

Some of the information can be described as **somatic**: it is information about states, processes, signals, patterns, correlations, etc. **within** the organism, including its appendages, sensors, effectors, etc.

Some organisms acquire and use information about objects, object parts, structures, events, processes, relationships, and causal interactions **outside** themselves, i.e. they acquire and use **exosomatic** information.

Organisms that use information can be described as having **semantic competences**, of varying kinds depending on what information they use, how they use it and what they use it for.

Some are able to use information about things that use information: this requires **meta-semantic competences**, e.g. used for acquiring information about what something perceives, learns, wants, attempts to do, plans, predicts, reasons about, etc. A special case is **self-directed** meta-semantic competence, i.e. being able to use information about one's own information processing.

There are many unknowns about which semantic competences, meta-semantic competences, self-directed meta-semantic competences exist in which organisms, when they evolved, why they evolved, what mechanisms evolved to support those competences, and what the consequences were of those evolutionary transitions.

There are also many unknowns about which organisms have all their semantic or meta-semantic competences pre-specified by the genome, and which do not and what the costs and benefits are of the different options.

Information, of different kinds, can play many roles in such control processes: e.g. information about what is possible, about what constraints exist, about where things are, about events and processes occurring, about causal interactions, about intentions, plans and actions of others, about the content of communications, about certain or likely consequences of various possibilities, about the merits of alternatives, about the locations of resources and dangers, about how close an action is to its goal, about whether something is moving in the right direction, or with the right spatial orientation, to achieve a goal, about whether a goal has been achieved, about how securely a goal has been achieved, and many many more.

Not all information-using organisms have self-awareness: many use information without having access to information about what they are doing -- they just do it. Others have additional self-monitoring capabilities that enhance those capabilities. Biological evolution has produced (not by design, of course, except where there has been deliberate breeding) an enormous variety of kinds of self-monitoring, some restricted to very local spatial and temporal regions others to large historical viewpoints and planning for long term achievements. Some forms of self-monitoring provide only information about physical relationships and processes, while others provide meta-semantic information: information about information processing. We do not yet (as far as I know) have a systematic taxonomy of types of information processing including self-monitoring, found in organisms. (A small start has been made [here](#).)

A more detailed, careful, wording of the question should refer not just to **one** type of information-processing architecture -- e.g. the (constantly developing) architectures of human minds -- but to many types, on different scales (e.g. microbes upwards), and different levels of abstraction (compare the physical, digital, and various software levels of abstraction in computing systems), including socio-economic information-processing architectures whose possibility depends on the evolved information-processing architectures in individual organisms.

The environment in question can include the physical environment (on different spatial and temporal scales) but also social, linguistic, and more generally cultural environments), as well as the information-processing architectures of other individual organisms -- conspecifics, competitors, collaborators, prey, predators, and underdeveloped offspring.

Don't assume that 'information-processing' refers only to what brains do: brains are physical machines that can, like computers (as Freud observed before the advent of computers?), support complex virtual machines made of coexisting interacting subsystems performing a host of mutually-dependent functions, as explained in presentations on virtual machinery available [here](#).

Emil Toescu raised some useful questions about the scope of this project, to which I have provided some tentative answers [here](#).

[See the background section for an explanation of how the idea of arranging a discussion of the question arose.](#)

First meeting: Wed 18th August, School of Computer Science 2pm.

This is a meeting to discuss this question, raised at the end of the recent research awayday.

Draft agenda for meeting (or series of meetings perhaps):

- Brief personal introductions, followed by first attempts:
 - to introduce and clarify the problem
 - to identify sub-problems on which progress can be made, e.g.
 - using biological evidence of various kinds about what animals of different sorts, or at different stages of development, can or cannot do, in combination with problems found in attempts to design machines with various competences, to analyse the problems that evolution had to solve -- e.g. producing a taxonomy of biological design problems?
 - finding evidence for the evolutionary history of various solutions
 - including evolution of different solutions to similar problems, and convergent independent evolution of similar solutions in different lineages e.g. solutions to problems of perception, learning, navigation, planning, communicating, motivation, decision-making, looking after and teaching offspring, ...
 - analysing requirements for the various intermediate stages of production of working systems, e.g. stages (before and after birth/hatching) between fertilisation of egg and production of fully competent adults,
 - comparing different kinds of developmental trajectory
 - (a) to understand the different requirements that they address
 - (b) to understand the different mechanisms underpinning those developments (including both physical machinery and virtual information-processing machinery).
 - (c) comparing information-processing requirements for different sorts of processes of development and learning
 - analysing mutual-dependencies between processes of physical development and development of information-processing mechanisms e.g. implications for control mechanisms of changes in physical size, shape, weight, strength, speed, physical needs (including energy requirements).
 - analysing information-processing requirements related to different social/cultural patterns and different roles in social/cultural systems (learner, teacher, protector, leader, challenger, competitor, collaborator, potential or actual mate, ...)
 - identifying physical, chemical, neural mechanisms that can support the various information-processing requirements
 - investigation the need for, and biological forms of, 'virtual' information-processing mechanisms and their physical underpinnings (compare the relationships between operating systems and file managements systems and their implementations in electronic mechanisms).
 -
- to discuss what is already known about the topic (And who is doing directly relevant research, here and elsewhere). Including 'position statements' [*]
- to find out who is interested in being involved
- to discuss ways of making progress
- Discussion of requirements for new forms of education/training for researchers able to work on such problems?
- Are there institutional obstacles to progress on such problems?
- Consideration of what, if anything, to do next.

[*] If you wish to offer a 'position statement' please email A.Sloman@cs.bham.ac.uk

The Problem(s): what can biological reproduction produce, and how?

It is generally accepted by the scientific community that something like Darwinian natural selection, coupled with epigenetic processes (occurring during individual development), including social/cultural factors in some cases, all responding to changing environmental opportunities, threats and constraints, can account for the huge diversity of living things on Earth, although there are some disputes about the full range of influences in the process, and about what can be explained by currently known influences.

A crucial part of this process is the production of new individuals from old ones. However, the results of such reproduction are of different kinds, requiring different sorts of explanation.

For example, [Steve Busby](#) (Biosciences) drew my attention to this useful paper about bacterial reproduction:

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2704931/?tool=pubmed>

Bacteria as computers making computers

Antoine Danchin

FEMS Microbiol Rev. 2009 January; 33(1): 3-26. doi:10.1111/j.1574-6976.2008.00137.x

Abstract:

Various efforts to integrate biological knowledge into networks of interactions have produced a lively microbial systems biology. Putting molecular biology and computer sciences in perspective, we review another trend in systems biology, in which recursivity and information replace the usual concepts of differential equations, feedback and feedforward loops and the like. Noting that the processes of gene expression separate the genome from the cell machinery, we analyse the role of the separation between machine and program in computers. However, computers do not make computers. For cells to make cells requires a specific organization of the genetic program, which we investigate using available knowledge. Microbial genomes are organized into a paleome (the name emphasizes the role of the corresponding functions from the time of the origin of life), comprising a constructor and a replicator, and a genome (emphasizing community-relevant genes), made up of genes that permit life in a particular context. The cell duplication process supposes rejuvenation of the machine and replication of the program. The paleome also possesses genes that enable information to accumulate in a ratchet-like process down the generations. The systems biology must include the dynamics of information creation in its future developments.

Compare:

Peter Hoffmann (video lecture), Life's Ratchet: How Molecular Machines Extract Order from Chaos, November 19, 2012,

<https://www.microsoft.com/en-us/research/video/lifes-ratchet-how-molecular-machines-extract-order-from-chaos/>

Peter M Hoffmann, 2016, How molecular motors extract order from chaos (a key issues review) 10 February 2016, *Reports on Progress in Physics*, Volume 79, Number 3, IOP Publishing Ltd.

<https://iopscience.iop.org/article/10.1088/0034-4885/79/3/032601/meta>

Bacteria illustrate only a subset of the problems our project is concerned with, since bacterial information-processing capabilities do not include most of the competences observed in vertebrates, and among vertebrates there are also considerable variations in information-processing capability. There is also considerable variation in competences available at birth: compare a new born human infant with a new born foal, struggling to its feet and going to suck its mother's nipple soon after.

At this stage we do not know the details of how humans differ from other animals, partly because we know little about what sorts of information-processing go on at various stages of human development (especially the information-processing required for developmental transitions), and partly because we don't know what the capabilities are of other intelligent species (including, for example, elephants, corvids, orangutans, ...).

Experimental observations may reveal animal (including human) behaviours in various conditions, but not the information-processing architectures, forms of representation, and forms of processing that produce those behaviours.

Attempts (over about six decades) to produce robots with similar competences have made only slight progress so far (in comparison with animals), but have taught us a lot about forms of processing that do not suffice, and about what some of the problems are, and have suggested some ways of thinking about the processing mechanisms that may also be relevant to understanding natural systems -- though we can be sure that all current working models are wrong, because they are all seriously limited in comparison with animals, or in some cases have competences (e.g. arithmetical competences, file searching and sorting competences) that surpass anything found in nature by a huge margin: i.e. so far artificial systems tend to do biologically old things very badly and some other things extremely well.

I have been discussing some of the issues with various colleagues, here and elsewhere, over several years, especially [Jackie Chappell](#) (Biosciences) since early 2005. We have developed some sketchy ideas about patterns of development of information-processing, but they are still too sketchy to be implemented, though it may be possible to implement various simplified fragments. Perhaps a combination of empirical biological research, attempted modelling, and theoretical analysis across a broad front will produce more rapid progress.

Some of the ideas were reported in this (invited) 2007 journal paper:

Jackie Chappell and Aaron Sloman
Natural and artificial meta-configured altricial information-processing systems
<http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609>
International Journal of Unconventional Computing, Vol 3, No 3, pp. 211--239, 2007

This built on and extended an earlier paper:
Aaron Sloman and Jackie Chappell,
The Altricial-Precocial Spectrum for Robots,
<http://www.cs.bham.ac.uk/research/cogaff/05.html#200502>
Proceedings Nineteenth International Joint Conference on Artificial Intelligence
[IJCAI'05](#), pp. 1187--1192, ([Online Proceedings](#))

Additional materials are available in PDF presentations [here](#).

Others involved in our discussions:

[Susannah Thorpe \(Biosciences\)](#)

[Nick Hawes](#) (Computer Science/Robotics).

A proposal abstract for a project to investigate a subset of the issues, led by Jackie, with collaborators in Scotland, Sweden, Germany and Italy, was submitted to the EU (FET-Open), without success.

We organised an interdisciplinary symposium on some of the issues [AI-Inspired Biology \(AIIB\)](#) at the AISB2010 conference this year.

A KEY IDEA -- Development of layers of competence, using a layered architecture

A key idea in the above papers is that for some species the information-processing architecture is not fully assembled at birth, but has layers of functionality that are added later, in a complex process that is partly controlled by

- results of previous experiments in the environment
- new meta-competences that observe and experiment with previously developed competences

where the growth of those meta-competences has to be delayed until the earlier competences have reached an appropriate level of maturity to meet the requirements of later developments. The following diagram was published in an invited journal paper in 2007. A later version, below, adds more detail.

Multiple routes from genome to behaviours

(Environment affects all embedded processes)

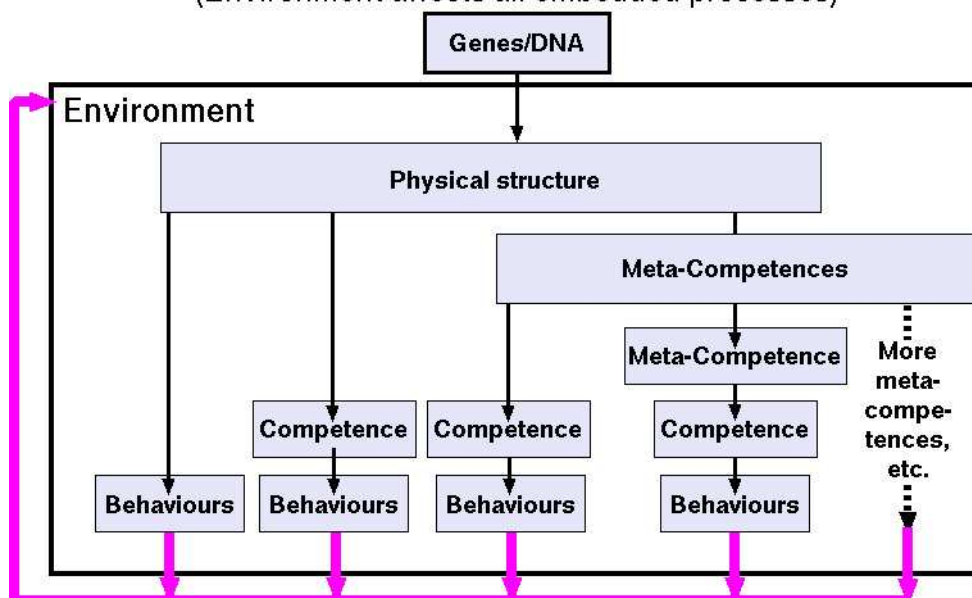
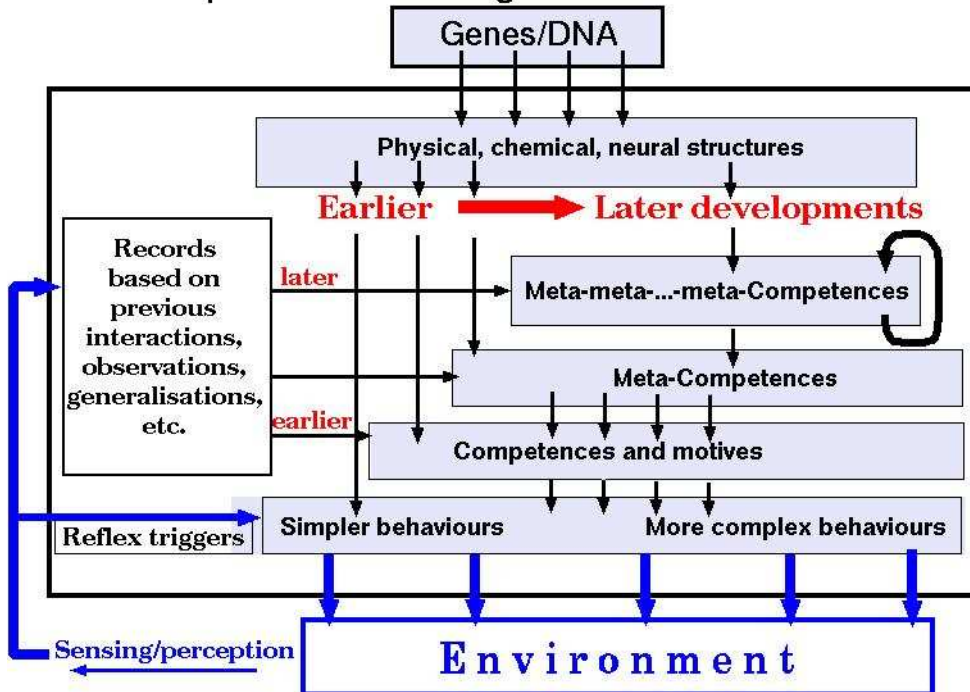


Figure from Chappell and Sloman (IJMC 2007) indicating different routes from genome to behaviours, via increasingly layered mechanisms developing at different stages (not found in all organisms).

Chris Miall helped with the diagram.

Updated version of the 2007 diagram, 2017
 The Meta-Configured Genome

Multiple routes from genome to behaviours



In part the changes correspond to ideas developed in the Meta-Morphogenesis project (since 2011):

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>

Added 21 May 2019: There is now a short (9minute) online video presentation of the ideas in this diagram (extracted from a longer online lecture), available on youtube

<https://youtu.be/G8jNdBCAxVQ>, and here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/movies/#meta-config-genome>

We don't know how many such layers there are, nor how they differ. One example seems to be development of a layer of processing that can learn how to plan sequences of actions by making use of previously acquired knowledge about types of situations, types of action, types of constraints on action, and types of consequences various sorts of action can have.

The ability to make plans in ones head may presuppose prior learning in which plans were constructed by trying them out in the environment. Doing that mentally may be important for actions that are either very expensive and time consuming or potentially dangerous. We don't know how many other species are capable of learning to make plans in advance of acting, but it seems clear that humans are not alone. (Even a spider, 'portia', seems to have this sort of ability as reported [here](#) and [here](#).)

Another example may be development of linguistic competences.

These seem to depend on prior abilities to interact both with the physical environment and with conspecifics.

(Children achieve a lot of interaction prior to use of language -- as do many animal pets that never learn a human language!)

That provides a substratum of competences, in which motives for enriched communication can be generated.

Such motives (illustrated by increasingly sophisticated questions and requests addressed by toddlers to parents and others) could help to drive processes that create a language-using layer in the architecture.

Such a layer might combine

- new linguistic and social information from the environment, with
- previously developed competences in interacting with the physical environment, and
- genetically provided generic mechanisms for constructing and manipulating information structures needed for language use, which the system tailors to the specific features of the local linguistic environment.

An intermediate transition after learning many useful verbal patterns seems to be the discovery of higher level patterns -- patterns linking patterns.

This amounts to the discovery of syntactic patterns in the language learnt that seem to switch the learning process to develop **powerful generative rules** instead of merely storing **useful verbal patterns**.

This famously seems to depend on mechanisms in the genome since other animals cannot do it (and some forms of genetic abnormality in humans prevent such learning).

But the details of what is learnt also depend on rich information provided by the environment about what syntactic patterns actually produce successful communication.

Such claims were made half a century ago by Noam Chomsky, e.g. in *Aspects of the theory of syntax*, MIT Press, 1965.

However, most of the details of his theories are almost certainly wrong because they attempt to explain linguistic development in isolation from the rich surrounding processes of cognitive development.

Triggering reorganization

Something similar to the switch from (a) learning useful patterns to (b) developing a unifying generative understanding, may happen in learning to interact with the physical environment: empirical discovery of useful regularities forming a rich enough web of relationships could trigger a high level cognitive system (mostly genetically determined?) to start looking for patterns in what has been learnt, so as to produce a more principled and 'generative' understanding of spatial, temporal, structures, and processes.

Euclidean geometry and the topology that it subsumes could be 'discovered' in this way as part of the biological process of developing mastery of the environment. Something like this seems to occur in other intelligent species with creative problem solving capabilities, not just humans.

However, there is a difference between making and using these discoveries and being able to notice that the discoveries have occurred, so as to reflect on them, describe them and communicate them to others (e.g. in teaching grammar or teaching mathematics).

Although many species perform actions that help others to learn, only humans seem to have the sort of meta-semantic cognitive competence that allows what is learnt to be explicitly articulated, as opposed to merely being demonstrated or shaped by modifying the environment.

I doubt that anyone knows what the genetic differences are (between humans and other species) that account for such differences in developmental potential, nor **how** the genetic mechanisms produce such effects.

Neither does anyone know exactly what changes within an individual human child produce the ability to observe, reflect on, describe and communicate abstract learnt competences.

If the above suggestions are close to the truth, how could a genome specify the powerful abstraction mechanisms that lie dormant until enough has been constructed by more concrete learning mechanisms to provide a basis for a new level of 'generative' understanding?

The main advantage of having that over not having it is the ability to be able to **work out** solutions to novel problems in unfamiliar situations instead of always having to use trial and error (statistical learning) to discover useful regularities.

This idea was proposed long ago by Kenneth Craik in
The Nature of Explanation, Cambridge University Press, 1943,

In humans there may be similar, but less obvious, layers of information-processing competence involved in progressing through different kinds of toys, games and non-verbal social interaction.

In other species such layers of competence may be related to basic mobility, various kinds of play and exploration, then more 'serious' hunting, fighting, escaping from predators, nest-building, etc.

Mobility in water, on land and in the air will pose different challenges and dangers, requiring different kinds of information to be precompiled into the genome in cases where individual learning would be impossibly dangerous. An over-adventurous **walking** neonate may slightly hurt itself, whereas similar exploration in a **flying** animal could be fatal.

This raises questions both about different evolutionary trajectories and about the differences in the information that needs to be encoded in the genome and the ways in which that generic genetic information is combined with the information acquired by the individual during exploration and learning.

The need for new formalisms and new mathematics

It is pretty certain that even though thousands of programming languages have been produced by computer scientists and software engineers, varying in both shallow ways (different syntax) and deep ways (e.g. different tradeoffs between what happens before and after programs start running), it is likely that new forms of 'programming language' (or something more general) will be required to explain the varieties of programming of individuals by the genome.

Reasoning about the different sorts of process specified by these new formalisms may require the development of new fields of mathematics.

Notes on varieties of products of reproduction

If a project of the sort envisaged is to get off the ground it will be necessary to have clear ideas about the different kinds of products produced by biological reproduction.

What follows is not a complete survey, only an indication of some of the diversity of types of product of reproduction -- very much a very unscholarly first draft requiring much expansion and refinement:

- production of new physical body, including sensors and effectors, with overall externally visible morphology and an intricate, mostly microscopic, network of physical/chemical support mechanisms dealing with: growth; repair; digestion; distribution of chemicals required for construction; growth; repair/reconstruction; reproduction; provision of energy; control signals of many sorts; detection of infections and fighting them; removal of waste products through various routes; maintaining various physiological states, e.g. temperature, blood pressure, supply of oxygen, and others.
- initiation, control, and termination of various kinds of externally visible behaviours, of varying functionality, varying spatial and temporal scope, and varying information-processing requirements
- production of long term growth and change in behavioural capabilities, based on changes in information-processing capabilities in some more complex organisms, e.g. learning routes, learning what items are nourishing and how to obtain them, learning to perceive 3-D structures and processes that may have to be interacted with, learning about dangers and how to avoid or reduce them, learning to communicate.
- various forms of perception, producing various kinds of usable information about the environment, including information that is only transiently available and usable only at the time and information may be acquired and stored over a period of time, and put to different uses at different times.
- learning to think and reason about abstractions (e.g. numbers, theories, causal laws, forms of communication); learning to think about what might have happened but did not, about constraints on what can happen, about possible future actions and their possible consequences, about unobserved entities and processes that explain observable ones (e.g. atoms, molecules, infectious agents).
- learning to think about thinking and other mental states and processes, in oneself and in others (e.g. offspring, competitors, potential mates, prey, predators), and using that to help communicate effectively with younger individuals whose learning can benefit from being accelerated through teaching.
- Learning to engage in various forms of collaboration, bartering, discussing, collaborative problem-solving, collaborative plan execution, ... all of which require the ability to think (to some extent) about what's going on in the minds of others.
- in the case of humans, learning to take part in academic studies, e.g. mathematics, philosophy, linguistics, theoretical physics, chemistry, ...
- [Dawkins Notion of an "Extended Phenotype"](#) would also have to be accommodated in a full theory how a genome does what it does. Ron Chrisley has suggested that there is also an 'extended genotype' where what is replicated is a combination of organism and niche, where the niche contributes information helping to control the replication of individual organisms and thereby of the niche. The cathedrals built by termites would be examples, insofar as they play a crucial role in the reproductive process.

However, as far as I know, nobody has attempted to survey the variety of types of result of such processes with a view to understanding what needs to be explained, and attempting to find explanations.

Reproductive and other pathologies

The reproductive processes discussed here are so complex that there are very many different ways in which they can go wrong -- some concerned with physical 'mistakes' and their consequences, others concerned with information-processing 'mistakes' and their consequences.

Some of both types may be fixable/reversible, others not.

When we have a good idea of how the genome produces the adult architecture, we shall be in a much better position to understand the various possible pathologies (which I suspect are more diverse than anyone has noticed), including perhaps finding ways of detecting them and dealing with them.

Likewise there are potentially very deep implications for educational policies and strategies.

TO BE EXTENDED.

Related work (To be Extended)

Jean Piaget

There are some analogies between these ideas and Piaget's ideas about developmental stages. But he was writing at a time when most of the relevant information-processing ideas had not been developed.

C.H. Waddington

Waddington had some relevant ideas. His metaphor of an 'epigenetic landscape' would need to be modified to fit our ideas: instead of the landscape constituting a fixed, genetically determined, set of possible trajectories for an individual to follow, we need a landscape that is modified to alter the possibilities as more is learnt, and instead of only one trajectory per individual there would be different aspects of the individual moving in parallel down different parts of the landscape modifying it on the way so as to influence other trajectories.

But that metaphor is too 'mechanical' to meet the needs of a theory about the development of an information-processing architecture, in which conventional ideas about wholes and parts have to be abandoned.

Steve Burbeck's website

has some relevant observations: <http://evolutionofcomputing.org/>

Multicellular Life as a Metaphor for the Future of Computing

We can add: "And vice versa". Perhaps the relationship is stronger than a metaphor.

Two books by Cohen and Stewart

These claim that any sort of complexity, eventually generates a new kind of structure that is relatively simple and allows predictions and explanations to be made relatively economically.

Jack Cohen, Ian Stewart,
The collapse of chaos,
Penguin Books, New York, 1994

Ian Stewart, Jack Cohen,
Figments of Reality: The Evolution of the Curious Mind,
Cambridge University Press, Cambridge, 1997,

John Holland makes some similar points in his writings, e.g. [At Home in the Universe: The Search for the Laws of Self-Organization and Complexity](#)

[Thomas Ray](#)

The Meta-Morphogenesis project (2012--)

Some time after this document was written, an invitation to contribute to a book commemorating Turing's centenary led to a proposal for a new Turing-inspired project, the [Meta-Morphogenesis project](#), which led to a theory of evolved construction kits of many types. See:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>

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

Background

A research awayday was held by two of the Colleges (EPS and LES) at this University on 21st July 2010 to enable academics to come together and share ideas about possible new research projects.

At the end of the day I asked if anyone would be interested in a meeting to discuss the question

How can a genome specify an information-processing architecture that grows itself guided by interaction with the environment?

(This is not meant to be a definitive formulation of the problem, but a pointer to a collection of ideas and problems explained more fully elsewhere -- some of them still to be clarified.)

A number of people expressed interest and I promised to send round an email message about arranging such a meeting.

I sent several messages to people in both colleges and to others in the Medical School and there seemed to be enough interest, among members of several schools and departments, to arrange at least one meeting, which may or may not lead on to subsequent activities.

The problem is very difficult, requiring collaboration between a number of disciplines (e.g. to specify in some detail exactly what needs to be explained).
It may require development of new kinds of mathematics to describe the processes.

As far as I know, nobody is working on this, although simplified versions are studied by some researchers in evolutionary computation, and there are many people investigating varieties of innate knowledge and competence in humans and other organisms, but not how the genome produces (or transfers) the knowledge and competences.

An updated version of my slides for the awayday on 21st July can be found here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#talk82>

Steps Towards a 21st Century University: Planting Seeds ... for a unified science of information

Maintained by [Aaron Sloman](#)
[School of Computer Science](#)
[The University of Birmingham](#)