Why reductionism does not work
Essay for the Kurt Gödel Prize

September 16, 2019

1 Emergence, Reductionism, and Causation

Kurt Gödel opposed the reductionist viewpoint of logical positivism ([149]:173) For example he wrote "*Even if we adopt positivism, it seems to me that the assumption of such entities as concepts is quite [as] legitimate as the assumption of physical objects and that there is quite as much reason to believe in their existence*" ([149]:174). He asked more specifically as regards biology, "*Is there enough specificity in the enzymes to allow for a mechanical interpretation of all functions of the mind? .. I believe that mechanism in biology is a prejudice of our time which will be disproved* ([149]:192). The arguments I give below show he is correct in both cases. The reductionist explanation he opposed is doomed to failure.

Life emerges out of physics in a bottom-up way: atoms are made of electrons and protons, molecules of atoms, cells of molecules, physiological systems (including brains) out of cells, and organisms out of physiological systems. The issue is whether higher levels have causal powers, or not: are they just epiphenomena, in view of the alleged causal completeness of the underlying physics?

The thesis of this essay is that, (i) Reductionism does not work because strong emergence occurs in many important cases. In particular in biology, "More is different" [5], since the whole is more than just the sum of its parts. (ii) This emergence is possible because downwards causation takes place right down to the lower physical levels, hence arguments from the alleged causal completeness of physics and supervenience are wrong. Lower levels, including the underlying physical levels, are conscripted to higher level purposes; the higher levels are thereby causally effective, so strong emergence occurs.

No violation of physical laws is implied. The key point is that outcomes of universally applicable generic physical laws depend on the context when applied in specific real world biological situations [8]. The same is true for example in the case of digital computers.

In this essay, I look at the nature of strong emergence, the fact that it occurs, and how it is enabled by downward causation (Section 2); the ways that downwards causation is possible (Section 3); physics examples (Section 4); digital computers as a very clear example (Section 5); and biology examples, including the brain (Section 6). I end with a discussion of how top down action causes branching of physics at the lower levels, and hence undermines the argument from supervenience against strong emergence (Section 7). This makes clear key questions reductionists would have to answer, but cannot.

2 Strong emergence occurs

In this section, I look at the definition of strong ontological emergence, and its relation to ontology (Section 2.1); its outcome, the existence of modular hierarchical structures (Section 2.2), which is the proper context to consider strong emergence (Section 2.3). It is useful to distinguish different types of strong emergence (Section 2.4). But does strong emergence occur? I argue that it does (Section 2.5), and particularly that abstract entities have causal powers (Section 2.6). They clearly cannot be explained in a reductionist way: they have a completely different nature than physical variables. The crucial point is that it is downward causation that enables strong emergence to occur (Section2.7),

2.1 Strong ontological emergence

Emergence is usually classified firstly into ontological and epistemological emergence, and secondly into strong emergence and weak emergence. This paper is concerned with ***strong ontological emergence***, that is, firstly it does not "*characterize the concept of emergence strictly in terms of limits on human knowledge of complex systems*" [114], rather it considers emergence as a phenomenon that exists in its own right (whether or not humans know about it, and independent of whether it concerns issues to do with the mind and brain.)

Ontology My take on ontology is as follows:

(i) Physical objects exist at all scales, so for example a desk exists just as much as the atoms out of which it is made (cf Eddington [38]), and that is true whether humans know about it or not. In that sense, this agrees with a materialist position;

(ii) Any entity that can be demonstrated in either an experimental or counterfactual way [96] to have a causal effect on physical entities that exist (item (i)) must also be said to exist, else we will have uncaused events occurring in the physical universe. This leads to the conclusion that for example algorithms, ideas, and social conventions are abstract entities that ontologically exist, as discussed below (Section 2.6). In that sense this disagrees with a materialist position.

Strong emergence Secondly, I follow Chalmer's definition of Strong Emergence [28]:

> **Strong Emergence (Chalmers):** *"A high-level phenomenon is strongly emergent with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are not deducible even in principle from truths in the low-level domain"*

This is a clear statement of the principle "More is different" [5]. By contrast, Chalmers states as regards weak emergence [28],

> **Weak Emergence (Chalmers)**: "*We can say that a high-level phenomenon is weakly emergent with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are unexpected given the principles governing the low-level domain...... It often happens that a high-level phenomenon is unexpected given principles of a low-level domain, but is nevertheless deducible in principle from truths concerning that domain."*

My position, in common with Leggett [83], is that such an "in principle" deduction is almost always illusory: it cannot in fact be done in a way that depends only on lower level quantities. As stated by Leggett [83],

> *"No significant advance in the theory of matter in bulk has ever come about through derivation from microscopic principles. (...) I would confidently argue further that it is in principle and forever impossible to carry out such a derivation. (...) The so-called derivations of the results of solid state physics from microscopic principles alone are almost all bogus, if 'derivation' is meant to have anything like its usual sense. Consider as elementary a principle as Ohms law. As far as I know, no-one has ever come even remotely within reach of deriving Ohm's law from microscopic principles without a whole host of auxiliary assumptions ('physical approximations'), which one almost certainly would not have thought of making unless one knew in advance the result one wanted to get, (and some of which may be regarded as essentially begging the question)."*

Essentially the same is stated by Laughlin [82] in the context of superconductivity, and Scott [136] in the case of the brain. One must look at the issue in the relevant hierarchical context (see Section 2.4), and then such emergence at some higher level L_2 from a lower level L_1 almost always in fact depends on concepts and entities at the higher level L_2.

The third possibility is [13]

> **Nominal emergence (Bedau)**: *The notion of a macro property that is the kind of property that cannot be a micro property, and is not strongly emergent.*

Thus this is the kind of emergence in the mind of reductionists. It is the case where more is just the sum of the parts: there are no surprises.

2.2 The Outcome: Modular Hierarchical Structures

The outcome of emergent processes is the existence of Modular Hierarchical Structures, with very different kinds of causation occurring at each level of the hierarchy. The way this works out is radically different in the cases of the the natural sciences and human sciences hierarchies (see Table 1).

	Inanimate matter	**Living Matter**
Level 10	Cosmology	Sociology/Economics/Politics
Level 9	Astronomy, astrophysics	Psychology, Rationality
Level 8	Space & planetary Science	Physiological systems
Level 7	Geology, Earth Science	Cell Biology, Cell signalling networks
Level 6	Materials, Structures	Molecular biology, Supramolecular chemistry
Level 5	Physical chemistry, crystals	Biochemistry
Level 4	Atomic physics: elements	Atomic physics: elements
Level 3	Nuclear physics	Nuclear physics
Level 2	Particle physics	Particle physics
Level 1	Fundamental theory	Fundamental theory

Table 1 *The emergent hierarchy of structure and causation for inanimate matter (left) and life (right) as characterised by academic discipline. Causality of appropriate kind occurs at each level in both cases, described by suitable variables for that level. The bottom four levels are common to both sides (life emerges from ordinary matter).*

Key to understanding emergence is conceptual clarity as to what variables belong to what levels in this hierarchy. Note that in the natural sciences hierarchy, higher levels correspond to larger scales (and lower energies), and only physical variables come into play. By contrast in the life sciences hierarchy, quite different kinds of variables come into play at higher levels: indeed they include non-physical variables (see Section 2.6).

The emergent higher level structures on the life sciences side are *Adaptive Modular Hierarchical Structures*, with specific functions at each level. Key features are,

- Structure, which underlies function, and shapes what happens [25] [103]
- Hierarchical: different levels of emergent complexity arise, each with appropriate emergent entities/variables and causation for that level [25]
- Modular: abstraction and information hiding occur with controlled interfaces, allowing module modification without destroying system function [18] [132]
- Networks: interactions between modules form causal networks [19], with preferred network motifs [97] [4] and perhaps hubs
- Adaptive: at each level entities adapt to the environment for that level, which includes higher levels. This takes place on evolutionary, developmental, and functional timescales.

The basis of life All living systems are of this nature [25] [129] [126], where the structural and functional details have been determined by Darwinian evolutionary processes [91] [25] in the sense of an Extended Evolutionary Synthesis [124] [26] [108].

Artificial systems Complex technological systems such as digital computers and aircraft have many of the characteristics of life, also being based in modular hierarchical structures [137] [18], except in general they are not as adaptive.

2.3 Strong emergence (Hierarchical Context)

In looking at what strong emergence happens, we must interpret Chalmers' "*truths concerning that phenomenon are not deducible even in principle from truths in the low-level domain*" [28] in the context of this hierarchy (Table 1). I claim it should be taken to mean

> **Strong Emergence (Hierarchical Context)**: *Truths concerning a phenomenon at a higher level are not deducible even in principle only using variables defined at a lower level in the hierarchy*

Variables that can only be defined in terms of relations at a higher level (e.g. crystal structure, which is level 5) are thus in this case by definition irreducible to variables defined at a lower level (e.g. the level of electrons and protons, Level 2).

Thus if a reductionist says "Yes but the crystal structure is nothing but an aggregation of electrons and ions", i.e. it is describable at Level 2, the response is that that statement tells you nothing about the specific crystal structure, for example whether it will support superconductivity or not (which is an ontological rather than epistemological issue: i.e. it is a matter of fact about the crystal, whether we know the answer or not). To answer that, you have to describe the details of the crystal structure itself, which is a Level 5 variable. Properties of electrons and ions *per se* simply does not determine anything about such macro level properties (superconductivity emerges at Level 5, and cannot even be described at Level 1). To say "the crystal is made of electrons and ions" does not in principle determine whether such properties will hold or not, in an ontological sense. It is an incomplete characterisation of the system. The issue is,

> **Higher level irreducible variables** *Does the outcome depend on the details of the higher level structure (as in the brain, where the detailed cortical connections matter, or in crystals, where the detailed lattice structure determines whether it is a superconducter) or not (as in a gas, where the details of what molecule is where is immaterial)*

An example that makes the point is that a protein may have the chemical formula $H_{n1}C_{n2}O_{n3}S_{n4}N_{n5}$ which (for suitable numbers n_I) lists all its chemical components, but that tells you nothing about its primary, secondary, tertiary, or quaternary structures [122], which are all higher level variables. It therefore says nothing about its function.

Given this understanding, the claim I make is that that strong emergence does indeed occur, as discussed below.

2.4 Types of Strong emergence

It is useful to distinguish three types of strong emergence.

- ***Unitary emergence*** This name is a generalisation of the use of that term in quantum physics. The basic idea is that outcomes at level L_I are determined uniquely by variables at level L_I in terms of initial data given at that level alone. For example, the motion of an ideal pendulum is determined by its initial position and angular velocity. These are macro level variables (Level 6) which entrain billions of lower level variables (motions of particles, Level 2) in a downward way. Again setting the right conditions for superconductivity to emerge (a laboratory level exercise) guarantees that superconductivity will indeed occur (an outcome at both the laboratory and electron levels). Outcomes are reliably determined by initial data.
- ***Branching emergence*** This is the case where outcomes at a level L_I are not determined uniquely by initial data of variables at that level because branching dynamics takes place [41] with the specific branching that occurs being determined in a contextual way [108]. For example, the reading of genes at the cellular level is determined by epigenetic processes influenced by higher level variables [109] [71]. Full knowledge of variables at the molecular biology level (Level 6) at a time t_0 does not determine outcomes at a time $t_1 > t_0$ because of this contextual dependence.
- ***Logical emergence*** When intelligent life emerges, rational thought occurs and enables deductive causation ([41]:§6) which has causal powers [39], affecting physical outcomes in a downward way, as do computer algorithms [89]. This is a case of an entirely new kind of causal effect (the causal efficacy of abstract entities) emerging via the structure of the brain and the higher level dynamics it enables.

2.5 But does strong emergence happen?

Strong emergence occurs in physics (unitary), digital computers (branching and logical), biology (branching), and the brain (branching and logical), where I refer to the classification in Section 2.4.

Physics Strong emergence takes place in both classical and quantum physics. It occurs through *broken symmetry* effects [5] such as chirality and existence of crystal structures [123], leading to quasiparticles [58] [145] which underlie emergent properties of materials [141] and superconductivity [82], and *through topological* effects as occurs in the Fractional

Quantum Hall Effect [81], topological insulators [61] [125], and colloids [94] [93]. Strong emergence in physics is discussed in Section 4.

Digital computers Digital computers provide a definitive example of both branching emergence and logical emergence, through the way computer algorithms control outcomes at all levels right down to the electron level [42]. This is discussed in Section 5.

Biology Branching strong emergence takes place in biology [41] [92] through downward constraints, downward control, and downward selection enabled by a variety of physiological and molecular biology mechanisms [108] involving signalling molecules [14] and the causal efficacy of information [113] [148], for example that encoded in the gene [150]. It is also central to the way Darwinian Evolution (in the sense of an extended evolutionary synthesis [104] [116] [152]) takes place [24] and enable evolutionary innovation [146], including emergence of new levels in the hierarchy. This is discussed in section 6.

The brain Agency, based in rational thought, occurs via branching emergence and logical emergence in the brain. This is apparent in the power of thought: the chain of causation from purpose to planning to muscle movement, as discussed in Section 6.3. This takes place in a social context that enables abstract social constructions such as money and laws to have causal powers (Section 2.6). Logical emergence takes place [41].

2.6 Abstract entities can have causal powers

Abstract entities have causal powers. Ideas and plans, social agreements, and Platonic entities all have causal effects in the physical world via brain functioning at the psychological level, enabled by the underlying brain structure and function [73] [48] [72].

Ideas have causal powers Through the functioning of the brain, ideas and plans have causal powers in the physical world. These abstract entities act down to the physical levels to result in billions of atoms being configured as buildings, roads, aircraft, and so on so as to serve individual and social purposes. These outcomes are the result of intention and planning. The ideas and plans that are the key causal element leading to these results are not coarse grained lower level variables. Although they are realised via brain states, they are not the same as any individual's brain state, because they can be shared between people, written down on paper, embodied in computer files, presented in a lecture, and so on. A plan for a building for example is not the same as any particular one of these realisations: it may originate in one person's head, but then attains a life of its own independent of the person who first brought it into being, as it is shared with others and realised in print. Ideas and plans cannot be reduced to any physical variables: they are of a completely different kind because they have a logical and symbolic nature [31].

The social world The social world is shaped by many social agreements, which change what happens in social interactions and thus in physical reality.

- **Language** is a symbolic system [31] that is the foundation of social communication. In each society, it is an culturally developed feature that enables society, technology, and commerce to exist. It played a key role in the rise of civilisation [21].
- **Money** is a key enabler of commercial transactions, as are closed corporations [59]. It can be realised in many forms, e.g. paper, coins, or electronic.

- **Laws** and associated physical manifestations such as contract documents and passports order society [59] and so control physical outcomes, for example passports determine where you can travel.

A very clear example of top-down causation of abstract entities is the rules of games, which constrain what is allowed to happen in physical terms. The rules of chess for example are abstract social agreements that have evolved over centuries; they are not the same as any individual's brain state, although they can be realised in that way, or in verbal form, or in books, or in computer programs. They have causal powers in that they constrain the possible movements of the pieces on a physical chess board, as if they were a force field.

Platonic entities can have causal powers Mathematical equations (abstract entities) change the world through the human mind, which comprehends them and uses them in engineering design. But what is their nature?

Gödel was a strong supporter of the Platonic nature of mathematics [149], as are high level mathematicians such as Roger Penrose [118] and Alain Connes [29]. The reason is that mathematical facts, such as the distribution of prime numbers, the fact that the square root of 2 is irrational, and Gödel's incompleteness results [149], are discovered rather than invented. Competent mathematicians everywhere, whatever their culture, eventually discover them and agree on them. Thus they are universal in nature.

But a long time objection to this proposal has been a claim by some philosophers that if they were to exist, they would be irrelevant as there is no plausible way they could be acccessed by the human mind. However Paul Churchland [30] has given a full answer to this argument, through a study of the way the neural network structure of the brain can learn to recognise abstract patterns. For example the number π - an abstract entity - can be comprehended and calculated to high precision by engineers anywhere in the world, and thereby influence real world engineering outcomes such as the design of aircraft engines and chemical plant. It is not a social construction, it is a mathematical discovery.

2.7 Downward causation enables strong emergence

Strong emergence demands downward causation This is clear in the case of biology on the one hand, and digital computers on the other.

Biology If all the higher emergent levels in the biological hierarchy (Table 1 §2.2) are to each be emergent levels with causal powers in their own right, as claimed by Noble [109], the effective higher level interactions (such as the pumping of the heart as an integral whole [107]) must reach down to lower levels to constrain and shape what happens at those levels (see Figure 1), for instance synthesizing the haemoglobin that is needed for circulation of blood. This happens by cell signaling networks [14] driving the branching logic of metabolic networks and gene regulatory networks at the lower levels according to physiological needs at higher levels, such as the function of the heart [45] or mental processes [71]. Hence metabolites are produced and proteins synthesized according to higher level needs (Section 6). But this in turn requires that electron and ion motions at the underlying physical level take place so that these biological functions occur as needed.

Thus the need is that by a process of downwards causation, biology can conscript underlying physics to its purposes. How this happens is discussed in Section 6.

Computers conscript underlying physics to implement abstract logic Similarly the flow of electrons in digital computers at the level of transistors is driven by the details

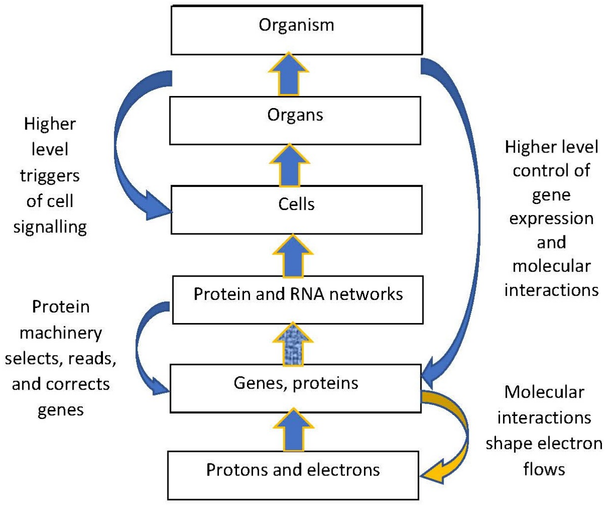

Figure 1: *Branching biology causes the underlying physics to branch, via time varying constraints altering the Hamiltonian.* From [41], extending a figure in [109]

.

of algorithms [80] as expressed in a high level computer programming language. Each level of the emergent hierarchy of virtual machines [139] has a precise effective logic at its own level, as described in the manual for the language (Python, Java, C, Assembly, etc.) used at that level, downwardly controlling electron flows at transistor level according to the digital logic of the corresponding machine code. We can understand every step of this process whereby abstract algorithms are causally effective in the physical world (Section 5) through existence of application programs[1] that can be contextually driven [42].

3 How is downward causation possible?

But how is downward causation possible? Many people deny it can happen (see [69] and [50]). In this section, I discuss various ways that downwards causation can take place. It occurs via time independent and time dependent contextual constraints (Section 3.1), homeostasis/feedback control (Section 3.2), downward emergence (Section 3.3), downward adaptation (Section 3.4), and downward selection (Section 3.5). I comment on the key features of existence of irreducible higher level variables and higher level organising principles (Section 3.6), and the multiple realisation of both at lower levels (3.7).

3.1 Contextual constraints

All outcomes depend on context. Constraints are a key part of context [12] [153] so a mechanism for downward causation is by changing that context [153] [17] Constraints can be physical, electromagnetic, chemical, biological, ecological, or social. A key distinction is whether they are time independent constraints, when outcomes are unitary, or time dependent constraints, when they are not.

[1] I do not enter here into the debate on the nature of computation [47]; the above statement is true whatever view one takes.

Time independent constraints A physical structure such as a dam by its overall construction constrains the position of each particle out of which it is made, but also acts to dam the water. Whether it will in fact successfully dam the water depends on the location of the dam in relation to the topography of the land, and whether it has leaks or not: all high level features relative to the particles making up the dam. They cannot be characterised at a lower level.

An electric circuit is made of a voltage source and wiring that is either open, and no current flows, or closed, and current flows. This is a topological feature of the wiring that cannot be described at any lower level: cutting the loop prevents a current flow. The wiring of a digital computer is an immensely complex construction that links specific transistors, resistors, and capacitors to others in an extremely precise way, determining how electrons can travel between components. The computer will function if and only if this higher level structuring (relative to the level of electrons) is correct.

Similarly the structure of neural networks in the cortex are a dense connectivity of neurons with each other via axons and dendrites [73]. The personality of each individual is determined by the details of that wiring - an immensely complex irreducible topological structure that determines what actions potentials can go where.

In all of these cases, emergence takes place because of the specific constraints arising from those structures. The emergent structures are much more than the particles out of which they are built.

Time independent constraints shape dynamical outcomes The dynamics of a Hamiltonian physical system $\mathcal{S}$ with state variables $\mathbf{r}_i$ depends on $\{\mathbf{r}_i, \dot{\mathbf{r}_i}\}$. If it is subject to time independent constraints

$$C(\mathbf{r}_i, \dot{\mathbf{r}_i}) = C_0,\ dC_0/dt = 0 \tag{1}$$

the evolution will be unitary: the initial data $\{\mathbf{r}_i(t_0), \dot{\mathbf{r}_i}(t_0)\}$ uniquely determines the state $\mathbf{r}_i(t)$ at all times. A simple example is a frictionless pendululm of length L with $X(t) = L\sin\theta(t)$, $Y(t) = -L\cos\theta(t)$. The constraint is

$$L^2 = X^2 + Y^2 = L_0^2. \tag{2}$$

The bob does not fall vertically to the ground because of the constraint (2), but rather moves on a circular arc. The equation of motion is

$$\frac{d^2\theta(t)}{dt^2} + \frac{g}{L}\sin\theta(t) = 0, \tag{3}$$

and the initial data $(\theta, \dot{\theta})(t_0)$ uniquely determines the outcome $\theta(t)$ at all times.

Time dependent constraints By contrast, if the constraints are time-dependent, it is their time dependence that controls what happens. A dam may have a valve that allows water to flow out when water is needed for agricultural purposes. Outcomes depend on when the valve is open. Similarly an electric circuit will have a switch that determines whether current flows or not, as will a digital computer (its ON/OFF switch); it functions when the switch is on. Brain plasticity at the macro level is based in plasticity at the micro level: details of neural network connectivity and weights in the cortex change all the time in response to interactions with the physical, ecological, and social environment [70] [48], changing electron flows in the brain and hence mental outcomes.

Time dependent constraints shape dynamical outcomes In the case of Hamiltonian systems, if there are time dependent constraints:

$$C(\mathbf{r}_i, \dot{\mathbf{r}_i}) = C(t), \tag{4}$$

it is they rather than initial data that determine outcomes. A simple model is a frictionless pendulum with time varying length $L(t)$. Equation (2) is replaced by

$$L^2 = X^2 + Y^2 = L(t)^2, \tag{5}$$

and the equation of motion (3) becomes

$$\frac{d^2\theta(t)}{dt^2} + 2\frac{\dot{L}(t)}{L(t)}\dot{\theta}(t) + \frac{g}{L(t)}\sin\theta(t) = 0, \tag{6}$$

In this case, the initial data $(\theta(t_0), \dot{\theta}(t_0))$ does not determine the solution $\theta(t)$ because of this time-variation of the constraint $L(t)$. It is $L(t)$ that controls the dynamical outcomes because of the second and third terms in (6).

This is a very powerful mechanism that underlies why Hamiltonian dynamics is not necessarily unitary dynamics. It occurs in the functioning of digital computers (Section 5) and in biology (Section 6); in both cases time dependent constraints conscript the underlying physics to higher level purposes (Figure 1).

Constraints on biological functioning are provided by the environment, for example global climate change (a time dependent large scale effect) has a serious effect on life in the sea [33] (an effect on the scale of animals). Global climate change cannot be characterized at any local scale, although its outcomes can.

3.2 Downward Control and Homeostasis

A more active process is downward control of lower levels, with feedback control being a key example in biology and engineering.

Downward control An engineering example is blasting at a quarry, where a radio transmitter (a holistic macro object) sends signal to a receiver (another emergent macro object) that triggers an ignition device which causes oxygen combination with nitroglicerine at the molecular level. Thus a top down signalling processes to the molecular level causes an explosion that causes rocks to be broken at the macro and micro levels.

In biology, many physiological systems at the macro level (Level 8 in Table 2.2) use cell signaling networks (Level 7) to control metabolic networks and gene regulatory networks shaping molecular biology processes (Level 6), thereby enabling higher emergent levels (Levels 8 and 9) to function according to the logic appropriate to that level (Figure 1). Examples are the heart [109] and memory [71].

Feedback control A particularly important form of downward control occurs in feedback control systems, whereby higher level goals determine lower level outcomes. The system state S is measured by a detector D and compared with a chosen goal G by a comparator C. If $S \neq G$ a signal is sent to a activator A that will alter S so as to move towards G. The set goal G is the causally effective factor in this situation, determining what happens both at the emergent macro level of the system and the underlying microlevels as this cycle is continually repeated. The initial data is irrelevant precisely because such a system is designed so as to give the desired output regardless of the initial data (provided the system's parameters are not exceeded). The total system $\mathcal{S} = \{S, D, G, A\}$ ia an emergent irreducible macro entity with a topological configuration physically (it forms a closed loop)

In engineering, this can be implemented in many ways: mechanically, electrically, electronically. The classic example was James Watts' governor for control of the speed of a steam engine. A common example is a thermostat, where the desired temperature T_0 is set on a dial. A thermometer measures the actual temperature T and compares it with T_0. If it is too low, a heater is activated, resulting in the temperature rising (at the macro level) and billions of molecules moving faster (at the micro level) - a classic case of top-down causation. If you set a different temperature on the dial at the macro level, a different outcome results at the micro level. The system as a whole is an irreducible emergent macro system. If you have all the components there but change the topology by undoing one connection, it no longer works. A more complex engineering example is an aircraft automatic landing system.

Homeostasis In biology, homeostasis is the process of feedback control at all levels whereby an organism can maintain a desired state of equilibrium despite all kinds of disturbances that may occur [20] [134]. It is a fundamental principle of physiology[126] [129], occurring at all levels in the emergent hierarchy: body temperature, blood pressure and so on are maintained at the macro level by using cell signalling networks [14] to control the needed processes at lower biological levels. Homeostasis at those levels, such as cross membrane electrical voltages, levels of potassium ions in axons, and so on operate via small scale localised feedback loops. The networks that implement such feedback in biology are irreducible emergent features [20]. An example of biological dynamics is a universal biomolecular integral feedback controller for robust perfect adaptation [7].

3.3 Downward emergence

The reductionist paradigm is based on the idea of the existence of lower level entities, such as billiard balls, that have fixed properties independent of the environment. This is often simply not the case when strong emergence occurs. Both the existence of lower level entities (this subsection) and their properties (next subsection) often depend on the environment in key ways. These are important cases of downward causation that strongly contrast to the billiard ball or particle based models.

Quasi-particles Properties of metals and semiconductors such as electrical and thermal conductivity depend on quasi-particles such as phonons. These occur because collective excitations of the crystal as a whole [138], which are by their nature irreducible emergent entities, lead to existence of quasi-particles at the electron level that are key players in condensed matter physics. Their existence (at the electron level) is only possible because the discrete symmetry (at the lattice level) of the crystal structure breaks the continuous symmetry of the underlying physical interactions [123]. Hence they come into being via a process of *downward emergence* (they would not otherwise exist). They are an excellent example of emergence and interactions between levels in physics [46] [58]: the crystal structure is an emergent higher level outcome of ions and electrons, that causes phonons to come into existence at the lower level (downward emergence) which underlie dispersion relations and band structure at the macro level, in turn determining optical absorption and electrical resistivity (macro properties).

Cooper pairs A similar example is the Cooper pairs of electrons that make superconductivity possible, which cannot even in principle be explained in a bottom up way [82]. At first glance they should not exists, because electrons repel each other; but lattice distortions at the crystal level (depending on the nature of the lattice) change the electric field (at the electron level) and so allow them to come into existence.

Gene expression Gene regulatory networks determine the set of proteins that are present in a cell by controlling what genes will get read when and where. This is a key aspect of developmental biology [154] [51] [52], as indicated in Figure 1. Obviously the physical state at the electron/ion level is changed by gene regulation processes.

Symbiosis An important feature of biology is symbiosis. In the case of obligatory symbiosis, organisms (for example specific birds and flowers exquisitely adapted to each other) are an emergent irreducible biological entity: its component members cannot exist on their own. This is for example why the death of bees is a threat to the existence of many plants.

A major example is multicellular organisms such as human beings. The individual cells in our bodies rely on the body as a whole for their continued existence. The lungs and circulatory system provides every single cell with oxygen and nutrients that are crucial to its metabolism, and take away waste material. Once the heart stops beating, blood no longer circulates and all the cells in the body (at the micro level) die within minutes because they cannot exist on their own; so the macro entity dies too. The circulatory system is of course itself an irreducible emergent entity: all its parts must be working and connected in a massively complex topologically connected network in order that the thing as a whole works. Thus for example, for survival, the state of the arteries is as important as the state of the heart.

3.4 Downward modification

Equally important, the nature of lower level entities - how they interact, which characterizes what they are - is often contextually dependent. Downward modification of properties takes place in physics, chemistry, and biology. Downwards selection occurs in engineering and biology, whereby lower level entities are selected to fulfil higher level purposes.

Physics The behaviour of neutrons is completely different when outside a nucleus than when bound in one [44]. Free neutrons decay with a half life of 10 minutes 11 seconds, whereas neutrons bound into a nucleus have half lives of billions of years (and if that were not so, we would not be here). Similarly free electrons interact with light in a completely different way than electrons bound in atoms, or in metals.

Chemistry The behaviour of a sodium atom is completely different when bound into a salt crystal with chlorine, than when free. This is true for all chemical compounds [62].

Development In developmental biology, cells originate as pluripotent (they can become anything). Their nature gets determined so as to meet specific higher level needs as developmental processes take place. Each cell has its fate determined (whether it becomes a muscle cell, blood cell, neuron, etc.) by positional indicators (morphogens) [154] [51].

3.5 Adaptive selection

Absolutely crucial to biology is the feature of downward selection, when the nature of lower level entities is determined by a selection process from an initial ensemble $\mathcal{S}$, with the outcome determined by some selection criterion $\mathcal{C}$. The result is a new ensemble $\mathcal{S}^+$ with members that are on average better adapted to the environment, as determined by the selection criterion, than those of $\mathcal{S}$. Thus it is a projection operation

$$\Pi_{\mathcal{C}} : \mathcal{S} \to \mathcal{S}^+ \tag{7}$$

with outcomes dependent on the selection criterin $\mathcal{C}$ (hence it is a top-down effect). The classic physics example is Maxwell's Demon. However it occurs in chemistry and engineering, biology in general and in evolution in particular, and in brain function.

Purification processes are key to the possibility of physics and chemistry experiments as well as engineering practice and medicine, because they all demand a supply of pure elements or compounds with specific well-defined properties. Thus there are many separation processes in chemistry, water purification, and in chemical engineering.

Biology Adaptative selection is a central process in biology. A central feature is variation [100] in order to create an ensemble from which a choice can be made.

> **Environmental adaptation** *It is a profound principle of biology that adaptation to the physical, ecological, and social environment takes place at all times and at all levels in a coordinated way. This happens on evolutionary, developmental, and functional timescales. It is a multi-level process whereby communal and individual needs drive adaptation not just of cells and biomolecules but also of developmental systems (cell signaling systems, metabolic networks, and gene regulatory networks) [116] [147]. This is a process of adaptive selection and hence a specific case of top-down causation [24]: different environments lead to different emergent outcomes.*

It has functional, developmental, and evolutionary aspects.

Functional Adaptation: Learning Animals adapt to their environment by learning processes. Brain plasticity at the macro level, entailing learning in response to interactions with the physical, ecological, and social environment, is a key feature of brain function [48] [72]. It is enabled by brain plasticity at the micro level where initially random synaptic connections in the neocortex are adjusted via gene regulation [71]: another case of top-down regulation of genes as indicated in Figure 1.

An example of a learning process is a brain that at time t_1 has neural connections encoding knowledge of Maxwell's equations, which were not there at time t_0. This is enabled by a social process of learning: a top-down process from society to detailed cortical connections, which cannot possibly have been determined in a bottom up way - the genome has no knowledge of Maxwell's equations [43].

Developmental Adaptation: Gene Regulation Developmental processes in biology take place in a environmentally dependent way [52], mediated by developmental systems [116] and gene regulation processes which are the explict mechanisms whereby downward causation takes place to the genome level as indicated in Figure 1. This enables for example acclimatization, whereby individual organisms adjust to changes in the environment.

Evolutionary Adaptation: the Extended Evolutionary Synthesis Evolution takes place over geological timescales, with adaptive selection taking place repeatedly after replication with variation [25] [91]. It occurs as described by the Extended Evolutionary Synthesis [26] [124], where evolutionary and developmental processes interact to shape outcomes ("EVO-DEVO" [26]), leading to emergence of and being shaped by physiology, epigenetics, and developmental systems as well as the genotype, with the developmental systems themselves being shaped by evolution [120]. For example, the existence of vision gives a great adaptive advantage. Evolution consequently leads to development of visual systems at the physiological level that require molecules such as rhodopsin at the molecular level, which would not exist apart from the macro level need of vision in a specific context (for example, eyes for use under water are different from those for use in air [55]).

It should be noted firstly, this is not a gene-centred process, it is much more than that, see *The Music of Life* [108] and Evo-Devo writings [26]. Secondly, it has major stochastic aspects [110], enabling organisms to adapt according to higher level needs [111], so its results are simply not predictable from specific initial conditions. This is also true because cosmic rays, determined by fundamentally random quantum processes, have influenced evolutionary history [119]. Thirdly, all this means it is highly misleading to describe evolution as an algorithmic process ([32]:50,63). It is nothing of the sort.

3.6 Irreducible higher level variables and organising principles

It is important that downward effects are driven by higher level variables that are irreducible, and are associated with higher level organising principles that are also irreducible. Higher organising principles are global states that cannot even in principle be described at any lower level. They reach down to shape what happens at all lower levels.

Physics Higher level variables in physics that are not reducible to lower level variables are related on the one hand to broken symmetries, such as chirality [5], and on the other to topological features such as occur in polymers and topological insulators [94]. These non-local states are discussed in Section 4.

Autocatalytic cycles and sets [143] "*An autocatalytic set is a collection of molecules and the chemical reactions between them, such that the set as a whole forms a functionally closed and self-sustaining system*" [67] An example is Bladderwort feeding [142]

A hiccup A hiccup is an involuntary spasm of the diaphragm that may occur once off or in a rhythmic series. It is a higher level integral process that can in severe cases have serious consequences such as fatigue and weight loss - clearly downwardly affecting physiological systems at both the macro and micro level.

Biological organisation Biological organisation [101] is very complex [126] [129]. However there are some fundamental underlying principles such as as closure of constraints [99] and of organisation [102] which cannot possibly be described or determined at any lower level than the organism as a whole. They are fundamental to the existence and functioning of life. Cellular organisation cannot be described at any lower level [65] [66].

Being alive "Alive" and "dead" are irreducible higher level variables, a state that organises all that happens in a biological system at each moment. This fundamental feature underlies the possibility of Darwininian evolutionary processes (understood in terms of an Extended Evolutionary Synthesis [104] [26]) that lead to the existence of life [91] [25].

Being conscious Consciousness is a global brain state, quite different than being asleep. This difference reaches down to affect all aspects of brain and body function [106]; the associated Circadian rhythms are a key feature of life [9]. Plans, ideas, and social constructions are irreducible higher level variables associated with consciousness (Section 2.6).

3.7 Multiple realisation

In real biology, higher-level functions, structure, and variables can be realised in multiple ways at lower levels [131]. This multiple realisability [15] causes major problems for any attempt to account for the higher level outcomes in terms of any lower level dynamics, because they cannot be naturally described at those levels [6].

Let higher level $L2$ variables V_I be realisable at lower level L_1 by any one of the combination of lower level variables u_i: $V_I = \cup u_i$ A behavioural law that can be simply stated in terms of the variables V_I at the higher level, such as

"IF {the sun is shining} THEN {the flower will open}"

can only be stated as a series of "OR" statements at the lower level:

"IF $\{u_1$ OR u_2 OR ..... $\}$ THEN $\{u_{1102}$ OR u_{1022} OR ... $\}$ "

for a vast number of combinations; one cannot even write them down at the molecular level, where they will number many billions. The latter statement is not a sensible scientific law (basically, it is not expressed in terms of 'natural kinds').

Whenever such multiple realization occurs, this is an indication that downward causation is occurring [11]: varying a macro variable causes selection of any one of the equivalence class of lower-level variations that correspond to this higher-level change. A key case is that a vast numbers of different genotypes can produce the same phenotype [147]. Darwinian selection takes place in terms of phenotype properties, which then chain down to select any one of the billions of genotypes that result in a better adapted phenotype. Consequently predictable convergence in function has unpredictable molecular underpinnings [105]. In the case of neural networks [16], there are many different detailed connectivity patterns that can result in the same higher-level outcome, such as face recognition. Training the neural network produces any one of those lower level networks that gives the desired macro level performance.

The true causal elements at lower levels are *equivalence classes* that all correspond to the same higher-level elements; these are the natural kinds in terms of which relationships between elements of a field can be defined [15]. The bottom line is that causation really happens in terms of the emergent dynamics at the higher level, such as Darwinian evolution (as just discussed in Section 3.5), whose dynamics cannot be described in lower level terms than survival of individuals. Equivalence classes at the biochemical and physical levels enable it to happen.

4 Physics examples

Although the focus of this essay is on life in general and the mind/brain in particular, it is useful to note that unitary strong emergence and associated top-down causation takes place in the case of physics.

Physics is based in unitary Hamiltonian dynamics at the fundamental level, with outcomes determined by the relevant interactions and the initial conditions, together with any constraints that may apply. Unitary emergence takes place through broken symmetries (Section 4.1) and topological effects (Section 4.2). In both cases irreducible higher level variables determine physical outcomes. A further key physics case is the emergence of the arrow of time, which cannot be determined in a bottom up way (Section 4.3).

4.1 Broken symmetries

The equations of fundamental physics are invariant under symmetries which are broken in real-world situations, which is why many emergent properties cannot be deduced in a bottom-up way from the foundational nature of the underlying physics [5]. This is a dominant feature of condensed matter physics [123] [138] and chemistry [88].

Chirality An emergent feature in physics and chemistry is chirality, that is, the handedness of an entity [5] such as the spin of a particle, the polarisation of a wave, or the handedness of the structure of a molecule. This has important outcomes in biology, where chirality affects biological activity because naturally occuring amino acids and sugars are chiral molecules. Thalidomide is a key case in point: the left-handed molecule was fine, but the right-handed one caused major abnormalities in babies. Chirality is an emergent property that cannot be determined locally: it needs some comparison reference object in order to be determined.

Quasiparticles Physics of quantum materials [76] is based in the way the continuous symmetry of the underlying fundamental theory is broken by the discrete crystal symmetry [138], leading to quasiparticles that control electrical and thermal conductivity and optical properties. This is a case of downward emergence (Section 3.3).

4.2 Topological effects

Physical systems characterised by topological ordering are strongly emergent because the relevant variables are non-local variables whose values are not determined by any local properties [94]. They occur at both micro and macro levels.

Quantum examples The Fractional quantum hall effect (FQHE) is an example of topological emergence [81] [94], where fractionally charged particles arise out of collective behaviour resulting from magnetic field interactions with 2-dimensional systems of electrons. Topological insulators - a vibrant field of current research - are another fascinating example of strong emergence in quantum physics with downwards effects [61] [125],

Polymers and colloids Soft matter physics [94] [93] deals *inter alia* with polymers and colloids, characterised by topological variables at a higher level than the electron level [81].

Knots and Knitting Knitting is an extraordinary operation where a 1-dimensional polymer chain is built into a 2-dimensional fabric (Figure 2) and then into a three-dimensional garment. Multiple levels of topological entanglement are thereby created: at the polymer level, at the fibre level, at the stitch level, at the fabric level, and at the garment level. The same is essentially true for knots in ropes, as used in sailing, climbing, and so on.

Figure 2: *Detail of structure of knitted fabric at stitch level (left) and outcome at fabric level (right). Irreducible higher level variables are introduced at each level, because they are topological.* Source: Wikipedia

.

4.3 Direction of time

The direction of time is a key property of macro physics, chemistry, and biology. However given all the details of positions and momenta of particles in a cylinder, you can't tell from that data in which direction of time entropy will increase, because coarse-graining a time symmetric micro theory in an isolated system necessarily results in time-symmetric macro physics. Whatever bottom-up proof you have via coarse graining that entropy S increases with time t: $dS/dt \geq 0$, will also prove that entropy increases in the opposite direction of time $t' := -t$ because the identical proof will show $DS/dt' \geq 0$. This is true both for the classical proof (Boltzman's H-theorem) and the quantum field theory version ([151]:150-151).

A global condition (the "Past Condition" [3]) is required to set initial conditions for local arrows of time, indicating *inter alia* in which direction of time entropy will necessarily increase. This is related to the cosmological *Direction of Time* [40] - a global variable deriving from the evolution of the universe, which can't be determined from lower level variables as it relates to the evolution of the Universe as a whole, and so is strongly emergent. *Inter alia* it determines the quantum mechanical arrow of time [35].

5 Digital computers

Digital computers are an excellent example of branching emergence and associated downward causation, because we can understand everything that goes on in them (since we built them!) Their structure enables the causal power of algorithms (Section 5.1). The link to the underlying physics occurs through logical branching via transistors(Section 5.2)

5.1 Causal power of algorithms

Digital computers are driven by the abstract logic of algorithms [80], initially coded in a specific high level language [1], and then chained down from the top level of the software hierarchy (the tower of virtual machines [139]) to machine code level by compilers [2] or interpreters. They are realised in a different language at each level, and at the machine language level control the hardware (transistors are "ON" or "OFF") through binary code ("0" or "1"). Because on the logical side arbitrary computations can be expressed this way (Turing's great discovery [63]), and on the hardware side the basic Boolean operations ("AND", "OR", "NOT") can be realised by suitable combinations of transistors,[2] this

[2]This is required in order that a Universal Turing Machine can function.

results in the ability to carry out any computable task. Hence algorithms can and do change the world [89].

> ***In digital computers****, the electrons in the transistors flow in accord with the logic of a chosen algorithm (translated into binary code). Physics does not determine those algorithms, precisely because they are logical in nature. They are abstract entities (logical procedures) that are causally effective.*

That this is a process of downward causation is obvious (the computer program is loaded by the operator at the macro level; different algorithms result in different electron flows at the transistor level). This downward causation is reflected in the multiple realisability of what happens at each downward step in the tower of virtual machines [139]; for example the Java Virtual Machine [86] enables Java to run on any hardware, and even the hardware/software distinction is mutable [139]. This enables the branching emergence whereby higher levels (such as a Word Processor program run at the top level) have real causal powers and control what happens at every physical level in the computer, including the electron level.

5.2 Logical branching via transistors

But how is this downward causation possible, leading to branching causation, given the alleged unitary nature of the underlying physics? On the one hand, algorithms control what gates are operated in what sequence hence the electrons at the bottom level do what they are told to do by the algorithms, all the while obeying Maxwell's and Newton's equations. On the other hand, the description on the microscopic level is based on a Hamiltonian for the ions and electrons ([123]:16):

$$\begin{aligned} H(\mathbf{R}_i, \mathbf{r}_i) &= -\sum_i \frac{\hbar^2}{2M_i}\nabla^2_{\mathbf{R}_i} - \sum_i \frac{\hbar^2}{2m_e}\nabla^2_{\mathbf{r}_i} + \sum_i\sum_{j>i} \frac{Z_i Z_j e^2}{4\pi\epsilon_0\,|\mathbf{R}_i - \mathbf{R}_j|} \\ &\quad - \sum_i\sum_j \frac{Z_i e^2}{4\pi\epsilon_0\,|\mathbf{R}_i - \mathbf{r}_j|} + \sum_i\sum_{j>i} \frac{e^2}{4\pi\epsilon_0\,|\mathbf{r}_i - \mathbf{r}_j|} \end{aligned} \tag{8}$$

where $\mathbf{R}_i$ are the positions of the ions with atomic weight Z_i and $\mathbf{r}_i$ the positions of the electrons. By itself, this would indeed lead to unitary dynamics.

However we must add to Hamiltonian a time dependent term $V(\mathbf{r}_i, t)$ due to the applied gate voltage $V(t)$ that turns the transistor ON or OFF [42]. This leads to a potential energy term in the Hamiltonian of the electrons:

$$H_V(t) = \sum_i eV(\mathbf{r}_i, t) \tag{9}$$

where the Level 5 (see the left-hand column of 2.2)) variable $V(t)$ determines the Level 2 variables $V(\mathbf{r}_i, t)$ in a downward way. This leads to a displacement of the electrons until a new equilibrium is reached where the electrical field created by the modified charge distribution cancels the electrical field due to the gate voltage. In order to calculate this new equilibrium, a self-consistent calculation based on the charge density due to doping, gate potential, and thermal excitation must be performed.

Thus if we regard the electrical field as a time-dependent constraint on the electrons in the transistor, this is a form of what is discussed in Section 3.1. The causal power of algorithms is realised via details of transistor design whereby the change in $V(t)$ enables currents to flow or not in the transistor, so it can act as part of a logical gate enabling higher level logic to emerge from the digital logic at the transistor level [95] [42].

6 Biological Examples

Biology emergence is based in contextual branching at each level, for example in gene regulatory networks and metabolic networks at the lower levels (Section 6.1). This enables emergence of physiological structures such as the heart (Section 6.2) and brain (Section 6.3). The link to the underlying physics is the contextual effects enabled via biomolecules, which enable top-down control of the underlying physics (Section 6.4).

6.1 Biology and Contextual Branching

The structure and function of biology are closely intertwined [25] [126] [129]

Structure The biological structural hierarchy is shown in Table 1 (§2.2). The cell is the crucial level: all living systems are made of cells, which are pluripotent to begin with but then (in multicelluar animals) are specialised to serve specific functions by developmental processes [51] [154] [52] (Section 3.3).

Function Causation at each level of the biological hierarchy tends to further the function α of a trait T through contextually informed branching dynamics so as to enhance the overall viability of the organism in its environment. As summarised by Hartwell *et al* [60]:

> "*Although living systems obey the laws of physics and chemistry, the notion of function or purpose differentiates biology from other natural sciences. Organisms exist to reproduce, whereas, outside religious belief, rocks and stars have no purpose. Selection for function has produced the living cell, with a unique set of properties that distinguish it from inanimate systems of interacting molecules. Cells exist far from thermal equilibrium by harvesting energy from their environment. They are composed of thousands of different types of molecule. They contain information for their survival and reproduction, in the form of their DNA*".

Because both the external environment and the internal milieu are continually changing, adaptation must take place on an ongoing basis (Section 3.5). Consequently, *Contextual Branching Dynamics* [41] is required to attain desired outcomes. A variety of systems enable this to happen, at both the macro and micro levels.

Systems At the macro level, all the physiological systems required for bodily function (structure) are controlled so as to respond appropriately to environmental circumstances (function). These systems include the Circulatory System, Immune System, Nervous System, Sensory Systems, and so on [126] [129]. Theye are emergent systems with the ability to respond appropriately at their emergent level according to their function, because that is what they are structured to do, thanks to evolutionary and developmental processes. They are supported at the molecular biology and cellular levels by

- **Metabolic Networks**, controlling production, distribution and use of matter and energy, and disposal of waste products [147]
- **Gene Regulatory Networks** controlling the reading of the genotype so as to produce proteins needed to comstruct the phenotype [147]
- **Cell signaling networks** conveying information that controls the gene regulatory networks and metabolic networks [14]

- **Developmental systems** that coordinate developmental programs generating an adult organism from a single cell [51] [154]

These often proceed on the basis of the lock and key molecular recognition mechanism of supramolecular chemistry [84] [85], with branching dynamics controlled by transcription factors, enzymes, and so on. In this case there is a selective response to a particular signalling molecule [14]. However they may also respond to physical stimuli, as for example in the case of voltage gated ion channels.

As a particular example, transcription factors may be ON (that is, able to bind to DNA) or OFF, in this way controlling transcription of DNA to messenger RNA and so to proteins needed for cell function. Thus if transcription factor TF_2 modulate synthesis of proteins in a metabolic pathway, it embodies branching logic of the form

$$IF\ \{TF_2\ on\},\ THEN\ \{X_2 \rightarrow X_3\},\ ELSE\ NOT \tag{10}$$

where X_A are metabolites [56]. This is what allows contextual control of gene expression [109] (Section2.2): regulatory processes determine what gene gets turned on where and when. This contextuality of branching represents top-down effects [109] [70].

Such processes are hierarchical and modular [128] [56]: a higher level regulator TF_1, sensitive to macro variables such as blood pressure or heart rate, can modulate the synthesis of intermediate enzymes and local transcription factors (such as TF_2 in (10), enabling top-down control of the process. The dynamics of such modules is multiply realizable: it does not matter what the internal variables and dynamics is, as long as the resultant genes or metabolites are what are required. Branching emergence occurs.

6.2 Physiology: The heart

The heart is a crucial macro system of the body (Section 3.3) which is an irreducible whole (Section 3.6), for example a key role in its functioning is played by a cardiac pacemaker which is a complexly connected set of cells controlling the heart rhythm.

The heart has been modeled in depth by Denis Noble [107], who calculated transmural pressure acting on coronary vessels due to myocardial stress: a downward effect from non-local variables to the level of coronary vessels, which respond to those stresses. His studies show how there is contextual regulatory control of lower level biological processes by higher level physiological states [109]. Noble and Noble [112] refer to the following data on how lifestyle choices influences RNAs and so control gene expressions:

- Bathgate et al (2018) have shown how "*RNA levels of control are changed by the lifestyle choices*" in identical twin studies [10].
- D'Souza et al (2017) investigated the fact that "*athletes have lower heart rates than non-athletes, which was once attributed to greater vagal tone. The changes have now been traced to microRNAs that downregulate expression of the HCN gene, so that the depolarizing current (if) produced in the sinus node cells is reduced by as much as 50%*" [36]

Thus these studies work out in detail how high-level choices produce change at the molecular level, thus demonstrating the downward causal power of higher level choices.

6.3 Structure and function of the brain

The power of thought (intentions, plans, equations) was discussed in section 2.6. The issue is how this can emerge from the underlying physics. The context is the Central Nervous System hierarchical structure [73], shown in Figure 3. This is the physical basis of consciousness. The neuron is the cellular level. The neocortex has cortical columns char-

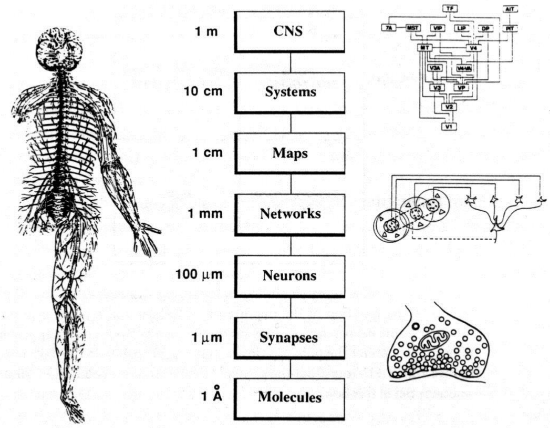

Figure 3: *The Central Nervous System* (T Sejnowski)

acterised by layered dense interconnections of neurons joined by synapses. The differences between people lie in the details of this neural network structure, which is determined by adaptation to the environment (neural plasticity underlies memory and learning). Action potentials propagate down neuron dendrites to the soma (nucleus) and then down axons to synapses where connection is made with other neurons by neurotransmitter diffusion across the synaptic cleft. Action potential spike chains are enabled by flow of ions in and out of axons via voltage gated ion channels, to create a current flow along the axon. This gives the Hodgkin-Huxley equations.

When built into neural networks linked by synapses in the neo-cortex, the resulting action-potential spike chains are the basis of logical thought and other mental phenomena. We do not know how thoughts are coded in action potential spike chains, nor do we know how consciousness arises; possibly by non-local synchronisation of neurons [74]. However Eric Kandel [70] gives a clear set of principles underlying what happens, as follows.

1. *All mental processes derive from operations of the brain.*

2. *Genes determine neuronal functioning.*

3. *Social and developmental factors contribute importantly to the variance in mental illness. These factors express themselves in altered gene expression.*

4. *Nurture is ultimately expressed as nature.*

5. *Altered gene expression induced by learning gives rise to changed patterns of neuronal connections, which give rise to different forms of thinking and behaviour.*

6. *Psychotherapy produces changes in long-term behaviour by learning which produces changes in gene expression, and hence changes in neuronal interconnection.*

These principles express top down action from the mental level to details of neural connections via gene regulatory networks (as in Figure 1). Brain plasticity at the macro level is enabled by changes in synaptic weights at the micro level, based in experience, via suitable gene expression. Logical emergence (Gödel's concern [149]) occurs [41].

6.4 The link to physics: contextual effects via biomolecules

The branching logical function that emerges in the brain is enabled at the molecular level by particular proteins, and again (as in Section 5) the issue is how is this branching dynamics compatible with the alegedly unitary underlying dynamics. I consider first Ligand gated ion channels, and then Voltage gated ion channels. Both control flow of ions

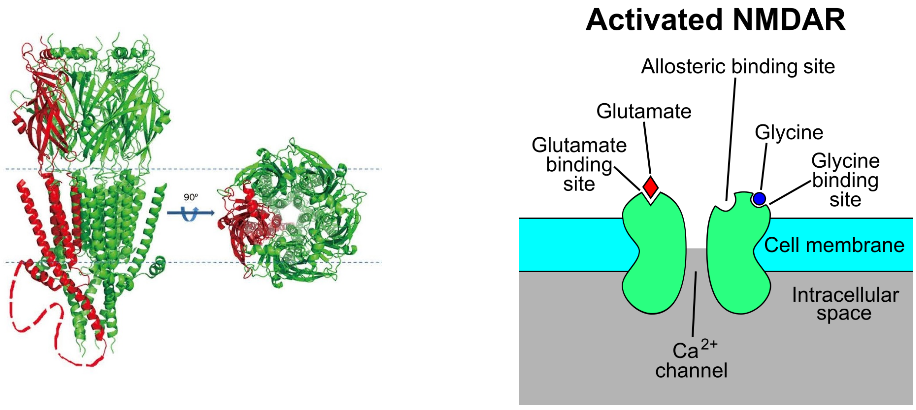

Figure 4: **Ligand gated ion channels** Left (a): *Basic structure of the cation-selective pentameric ligand-gated ion channels.* Source: [156] Right (b): *Stylized depiction of an activated N-methyl-D-aspartate (NMDA) receptor.* Source: Wikipedia.

across membranes, hence facilitating messaging in the neural system [90].

Ligand gated ion channels These are key to synaptic function [144], due to their speficic molecular structure (Figure 4(a)). They occur on the postsynaptic neuron in a synaptic cleft. Neurotransmitters released into the cleft by an excited presynaptic neuron binds to them if of the specific type they recognize, causing a conformational change which opens the ion channel. The resulting flow of ions across the cell membrane leads to either depolarization or hyperpolarization and so controls spike chain initiation. Thus time dependent molecular signals [14] reach down [109] [71] to change the conformation of biomolecules and alter outcomes. Recognition of the specific ligand by the receptor isdue to the lock and key molecular recognition mechanism of supramolecular chemistry [84].

In one specific case (Figure 4(b)), the ligand-gated ion channel is gated by the simultaneous binding of glutamate GLU and glycine GLY, thus it acts as an AND gate. The ion channel structure results in branching dynamics with the following logical structure:

$$\texttt{IF \{GLU AND GLY\} THEN \{allow ion flow\}, ELSE not} \tag{11}$$

This logical function is enabled by changes in the 3-dimensional conformation of the ion channels (Figure 4). This is the way the underlying physics (which determines possible molecule shapes) enables the logical (binary) outcome expressed in (11).

Thus control is via conformational change of proteins, which changes dynamics [57]. But how does this relate to the underlying Hamiltonian dynamics? The ligand binding changes molecular shape and hence alters the Hamiltonian and hence the dynamics, as in the case of the pendulum with varying length (Section 3.1). Karplus [75] states

"First, evolution determines the protein structure, which in many cases, though not all, is made up of relatively rigid units that are connected by hinges. They allow the units to move with respect to one another. Second, there is a signal, usually the binding of a ligand, that changes the equilibrium between two structures with the rigid units in different positions".

Working this out needs detailed quantum chemistry simulations with many-body Hamiltonians for electrons and nuclei. The Schrödinger equation for the ions and electrons comprising the biomolecules is again (8). Solving it in detail for a very large number of nuclei can be done by the CHARMM simulation suite of programs used widely for macromolecular mechanics and dynamics [22] [23]. However one can get a heuristic solution using the Born-Oppenheimer approximation [79] to determine the dependence of outcomes on nuclei distances $\rho_{IJ}(t)$, which can be regarded as time-dependent constraints dependent on binding molecules. This is how messenger molecules alter the outcomes of the underlying physics (which electrons flow where and when).

Voltage gated ion channels These ion channels [127] [27] are crucial to spike train propagation. When imbedded in axon and dendrite membranes they control the flow of potassium, sodium, and chloride ions across the membrane, leading to action potential spike chain propagation along the axons and dendrites. They implement the following branching logic:

$$\texttt{IF } \{V > V_0\} \texttt{ THEN } \{\texttt{allow ion flow}\}\texttt{, ELSE not} \tag{12}$$

via conformational changes of these molecules induced by the membrane potential V, for some threshold V_0 [41]. These proteins are selected in order to perform this function via Darwinian adaptive processes [147]. When built into neural networks with neurons linked by synapses in the neo-cortex, the resulting spike chains are the basis of logical thought and other mental phenomena [73] [72] [48]. Thus in this case, branching emergence supports logical emergence (cf. Section 2.4), as illuminatingly discussed by Paul Churchland in [30].

In both cases, conformational change of ion channels enables lower level branching dynamics [41]: biology causes physical branching by altering constraints at the molecular level in a time dependent way (Section 3.1). In this way unitary physics is conscripted to implement the branching logical dynamics of biology such as in Eqn.(11) and Eqn. (12), enabling branching emergence (Section 2.4) to occur.

7 Branching physics and supervenience

The previous sections give sound arguments firstly that strong emergence does indeed happen, and secondly regarding how it happens. However arguments based on supervenience together with the alleged causal completeness of physics at the lower levels claim this is not possible [77] [78], as discussed in depth in [50]. What answer can one give?

The premise is wrong. Physics is not causally complete firstly, because of quantum uncertainty (Section 7.1). Secondly, there are no isolated systems in the real world (Section 7.2). Thirdly, physics by itself is not causally complete because of the contextual effects discussed in previous sections (Section 7.3). Consequently while synchronic supervenience may be true, in most real world situations diachronic supervenience is not (Section 7.4); therefore arguments from supervenience fail to disprove strong emergence. The conclusion considers the implications of all the above for emergence and reductionism (Section 7.5).

7.1 Physics is not causally complete: quantum uncertainty

Physics is not causally complete at lower levels because it is not possible *in principle* to predict specific outcomes at the quantum level [87] [49]. This is proven *inter alia* by the foundational 2-slit experiment (Figure 5). Quantum uncertainty is irreducible in the

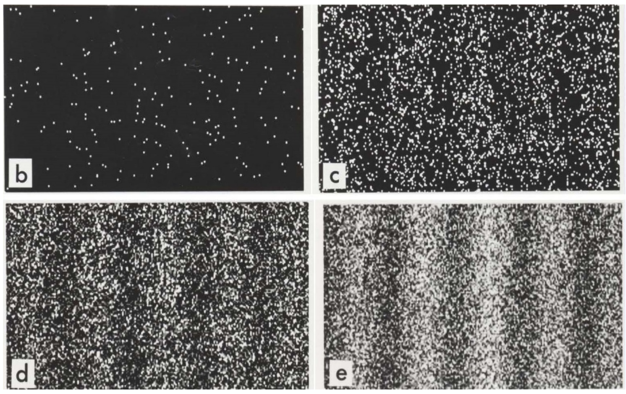

Figure 5: **Quantum uncertainty** *Double slit experiment performed by Dr. Tonomura showing the build up of an interference pattern of single electrons. The numbers of electrons are, (b) 200, (c) 6000, (d) 40,000, and (e) 140,000.*

one real world where we can carry out experiments. Statistical outcomes however are determinate.

The point then is that quantum uncertainty can get amplified to the macro level, and this happens for example in biology in the case of radiation damage to DNA caused by cosmic rays [119]; but the emission of a cosmic ray by an excited atom is a quantum event that is intrinsically unpredictable (there is no physics equation that tells when it will be emitted, or in what direction it will go). Cosmic ray damage to DNA has arguably been a significant effect in terms of the evolutionary history of life on Earth [135].

Amplification of quantum events also takes place in photomultipliers, CCDs, particle detectors, and so on; indeed one can relate this to Eqn.(6) for a variable length pendulum as follows: consider a Schrödinger-cat like setup, where a radioactive element emits particles received by a detector which each time sends a signal to a computer that uses it to alternately increase and decrease the length $L(t)$. Then the dynamical outcome of Eqn.(6) is in principle unpredictable: the motion of the pendulum is not determinate.

Thus the claim of causal completeness of physics, in the sense of being a unitary theory where specific outcomes are predicted by the initial data, is simply not correct.[3]

7.2 There are no isolated systems

The belief that physics leads to unitary dynamics is based on the combination of Hamiltonian dynamics with the concept of an isolated system. But while the isolated systems that would lead to unitary behaviour at the micro and hence macro level are a useful conceptual device to isolate causal mechanisms at work, they do not in fact exist in the real universe,

[3] I am discounting Many Worlds and Hidden Variable theories because they simply have no cash value for the physicist doing experiments in her laboratory. They do not predict, on the basis of the initial data, the outcomes she will measure in specific individual cases.

both in temporal and causal terms. They may however be a useful approximation in restricted circumstances for a limited timespan.

A first example is that a freely oscillating pendulum in a laboratory cannot have existed for all time: it will not have existed before it was manufactured. Furthermore, however excellent it is, it will in fact not be frictionless: it will gradually slow down due to friction, which is possible only because the laboratory is in touch with a heat sink (the dark night sky) into which it dissipates waste heat. The reason the night sky is dark is because the cosmological context of the expanding universe is such that Cosmic Background Radiation has a temperature of 2.7K [121].

A second example is that a digital computer has a non-zero error rate due to cosmic rays [157]. The abstract of [115] states

> *This paper presents a review of experiments performed by IBM to investigate the causes of soft errors in semiconductor memory chips under field test conditions. The effects of alpha-particles and cosmic rays are separated by comparing multiple measurements of the soft-error rate (SER) of samples of memory chips deep underground and at various altitudes above the earth. The results of case studies on four different memory chips show that cosmic rays are an important source of the ionizing radiation that causes soft errors*

As indicated in the previous section, the specific resultant errors that occur (at the macro level) due to ionizing effects (at the micro level) are not predictable even in principle.

Random environmental effects at the molecular level In practice, the environment for biological systems at molecular levels is highly random: they are subject to massive fluctuations due to random molecular motion. However molecular machines have evolved to give reliable outcomes in this context by harvesting the molecular storm, see *Life's Ratchet: How Molecular Machines Extract Order from Chaos* by Peter Hoffmann [64]. A key way higher level layers can extract order out of this chaos is by adaptive selection from the ensembles provided by random processes, from which they can select preferred outcomes according to higher level selection criteria and thereby harness stochasticity [111] (Section 3.5). In particular this plays a key role in brain function [54] [130]. Thus far from physics being unitary at the appropriate level, as envisaged in supervenience discussion, it is highly random at this level, and biology takes advantage of this feature.

7.3 Physics by itself is not causally complete

In addition, physics by itself is also not causally complete because of the contextual effects that determine outcomes, as discussed in this essay.

No physical system is isolated from its larger context Phyiscs *per se* is not causally complete because biological, psychological, social, and environmental processes affect what happens in the world according to higher level dynamics, thereby jointly shaping outcomes. They do this by altering the context within which specific physical outcomes occur. Physical forces do the work needed in this larger functional context. I will just mention two specific cases: Darwinian evolutionary processes, and the functioning of digital computers.

Darwinian evolution Evolution over geological timescales in the sense of an extended evolutionary synthesis [124] is a key process in biology [91] that cannot be comprehended in physics terms (neither 'animal' nor 'alive' are physics concepts). It is a form of adaptive selection (Section 3.4): higher level conditions select what happens at lower levels according to the selection criterion of successful reproduction. Now the point is that survival of individuals depends on factors like evolution of the social brain [37] [53] and development of the symbolic capability of the human mind [31]. These are both irreducible higher level factors that have lead to the extraordinary success of the human race, in particular enabling the emergence of technology and commerce and thereby altered physical outcomes across the world [21] [59]. The underlying physical interactions in the brain and outcomes in the world only proceed within this broad context of human evolutionary development, in which abstract elements such as national pride and societal issues such as technological competence play an important role. Gene-cuture co-evolution takes place [53]; physical causation is just one part of the overall story (with some key causal factors being abstract).

Digital computer outcomes Digital computers depend on the algorithms that drive them (Section 5). The algorithms deployed depend on goals for which computer programs are written, which are engineering, economic, and social purposes, for example including the development of search engines and social media. These in turn are crucially affected by the values and understanding of meaning that are driving them [140] [42], which are in fact the highest level variables shaping the outcomes of digital computers.

In summary : physics *per se* is not causally complete. In determining what actually happens, all these other factors need to be taken into account as part of the causal matrix determining specific physical outcomes. Physics alone cannot determine them [39].

7.4 Synchronic and diachronic Supervenience

It has been claimed [77] [78] that supervenience prevents strong emergence; arguments for and against are discussed in [50].

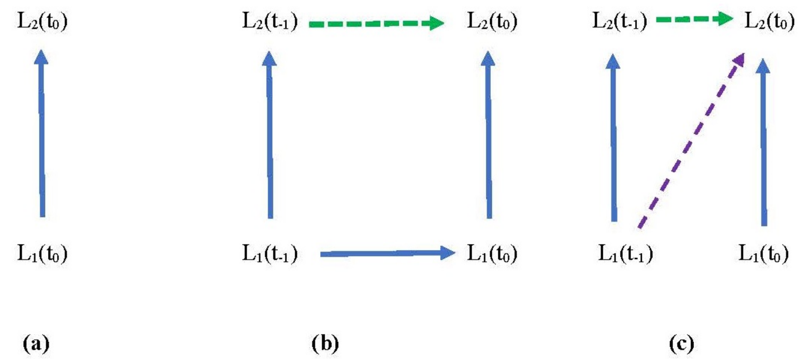

Figure 6: **Unitary supervenience**. *Left (a): Synchronic supervenience. The lower level state $L_1(t_0)$ uniquely determines the higher level state $L_2(t_0)$ at that time. Middle (b): The lower level state $L_1(t_{-1})$ uniquely determines the lower level state $L_1(t_0)$, which uniquely determines the higher level state $L_2(t_0)$ at that time. Right (c): The outcome: unitary diachronic supervenience. The lower level state $L_1(t_{-1})$ at time t_{-1} uniquely determines the higher level state $L_2(t_0)$ at the later time t_0.*

Unitary emergence Supervenience assumes the lower level state $L_1(t)$ synchronically (i.e. at each instant t_0) uniquely determines the higher level state $L_2(t)$: $L_1(t_0) \Rightarrow L_2(t_0)$. If $L_2(t_0)$ were different, $L_1(t_0)$ would be different (Figure 6 (a)). In the case of unitary emergence, physics is indeed causally complete at the lower level, so $L_1(t_{-1}) \Rightarrow L_1(t_0)$ (Figure 6 (b)), and then there is no freedom for higher levels *per se* to influence outcomes: unitary supervenience $L_1(t_{-1}) \Rightarrow L_2(t_0)$ occurs (Figure 6 (c)). Both synchronic and diachronic supervenience hold: the lower level state $L_1(t_{-1})$ determines the lower level state $L_1(t_0)$ and hence both the higher level states $L_2(t_{-1})$ and $L_2(t_0)$.

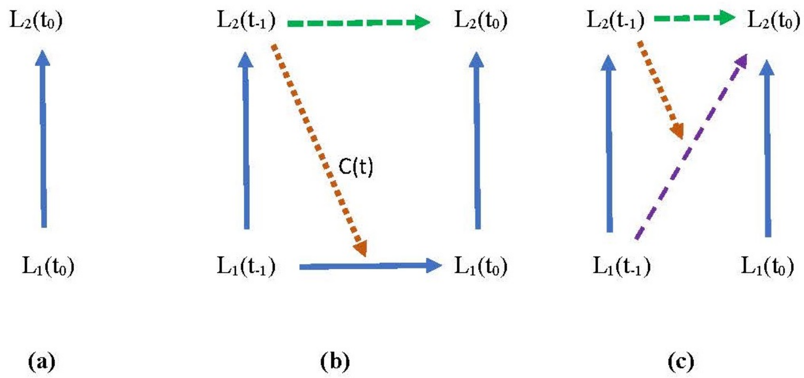

Figure 7: **Non-Unitary diachronic supervenience**. *Left (a): Synchronic supervenience. The lower level state $L_1(t_0)$ uniquely determines the higher level state $L_2(t_0)$ at that time. Middle (b): The lower level state $L_1(t_{-1})$ does not uniquely determine the lower level state $L_1(t_0)$, which is also influenced by the higher level state $L_2(t_{-1})$. Right (c): The outcome: non-unitary diachronic supervenience. The lower level state $L_1(t_{-1})$ at time t_{-1} does not uniquely determine the higher level state $L_2(t_0)$ at the later time t_0. In fact $L_2(t_{-1})$ determines the outcome.*

However as discussed in depth above, in living systems we have branching emergence rather than unitary emergence, and in the case of the brain we have logical emergence (Section 2.3). The outcome then is very different.

Branching emergence In the case of non-unitary emergence, physics is not causally complete at the lower level (Sections 7.2 and 7.3) and higher levels crucially influence outcomes, so non-unitary diachronic supervenience occurs. Synchronic supervenience means the lower level state $L_1(t_{-1})$ determines the higher level state $L_2(t_{-1})$ (Figure 7 (a)), and $L_1(t_0)$ determines $L_2(t_0)$ (Figure 7 (b)). However the lower level evolution $L_1(t_{-1}) \to L_1(t_0)$ depends on the higher level state $L_2(t_{-1})$ via time dependent constraints, downward emergence, and downward selection (Section 3). Thus the higher level state $L_2(t_{-1})$ influences $L_1(t_0)$ and hence $L_2(t_0)$ (Figure 7 (c)): the higher level state has causal power. Diachronic supervenience does not occur [133].

Example: Learning Many examples have been given above, and I will consider just one here: the brain of a student who knows Maxwell's equation at the time t_0. If we could reproduce in another brain in precise detail all her neural conections together with all the excitations of of those neurons, it is plausible that other brain would also know Maxwell's equations (Figure 7 (a)). However the brain is plastic: at a previous time t_{-1} she did not know Maxwell's equations; the question is how did they get coded in her synaptic connections at the later time?

This happened by a process of learning at the macro level. Through interactions with a teacher, textbooks, and her own internal thoughts, she learned those equations by time t_0. The learning process proceeded by downward causation controlling the reading of genes in synapses (as stated in Kandel's principles [70], see Section 6.3), and so altering her neural network connections: the lower level state $L_1(t_{-1})$ changed to $L_1(t_0)$. That change could not take place without the explicit learning process indicated by the downward arrow $C(t)$ in Figure 7(b) (see the discussion in [71]), which is a socially mediated process (no brain can meaningfully be regarded as living in isolation [34] [37]).

Finally one should note that in any case "supervenience is an inadequate device for representing relations between different levels of phenomena" (Humphreys [68]). Emergence involves much more.

7.5 Conclusion: Emergence and Reductionism

This essay has made the case that unitary strong emergence, branching strong emergence, and logical strong emergence all occur. Contextual choices are being made all the time at all levels of the emergent hierarchy of biology (Table 1), with real causal power residing at every level, including the the psychological and social levels. All levels are equally real (as emphasized by Denis Noble [109]). The result is amazing:

> "*We can give a general characterization of what it is for a system to be able to represent within itself some other system, and so can think of organisms in terms not of biochemistry or evolutionary biology but of information theory and formal logic. And from this point of view we can consider not only consciousness but self-consciousness, and a system that can represent within itself not just some other system but itself as well. There are a whole series of self-reflexive arguments* (Lucas [87]).

Top down action enables this, whereby the lower level physics is conscripted to fulfil higher level purposes via time dependent constraints, downward emergence, and downward selection. Thus the arguments against strong emergence based in diachronic supervenience do not hold. There is no violation of the underlying physical laws. Rather their operating context is shaped to obtain the desired outcomes.

In summary, the reductionist views Gödel opposed [149] are unjustified.

> **What questions reductionists cannot answer:** *Reductionists cannot answer why strong emergence (unitary, branching, and logical) is possible, and in particular why abstract entities such as thoughts and social agreements can have causal powers. The reason why they cannot answer these questions is that they do not take into account the prevalence of downward causation in the world, which in fact occurs in physics, biology, the mind, and society.*

Details of why this is so have been given above. Further support for this view comes from consideration of both quantum entanglement [117], and the emergence of classical physics from quantum theory [49] [42], which I do not consider here.